



TECHNISCHE UNIVERSITÄT MÜNCHEN

Fakultät für Chemie

Bayerisches NMR-Zentrum

Lehrstuhl für biomolekulare NMR-Spektroskopie

Structural and Biophysical Characterization of SPOP- Client Interactions Involved in Diabetes and Cancer

MICHAEL SEBASTIAN OSTERTAG

Vollständiger Abdruck der von der Fakultät für Chemie der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzende(r)

Prof. Dr. Bernd Reif

Prüfer der Dissertation:

1. Prof. Dr. Michael Sattler

2. Prof. Dr. Franz Hagn

Die Dissertation wurde am 26.09.2019 bei der Technischen Universität München eingereicht und durch die Fakultät für Chemie am 04.02.2020 angenommen.

Structural and Biophysical Characterization of SPOP-Client Interactions Involved in Diabetes and Cancer

(Strukturelle und biophysikalische Charakterisierung von SPOP-Substrat-Interaktionen im Zusammenhang mit Diabetes und Krebs)

Dissertation

zum Erlangen des akademischen Grades
Doktor der Naturwissenschaften (Dr. rer. nat)

vorgelegt von

Michael Sebastian Ostertag

geboren am 03.07.1991 in München



Technische Universität München

Fakultät für Chemie

Bayerisches NMR-Zentrum

Lehrstuhl für biomolekulare NMR-Spektroskopie

HelmholtzZentrum münchen

Deutsches Forschungszentrum für Gesundheit und Umwelt

Institut für Strukturbiologie

München, August 2019

Table of Contents

Abstract	3
Zusammenfassung	4
1 Introduction	5
1.1 Biological Background of the Speckle-type POZ Protein (SPOP).....	5
1.2 The Influence of SPOP-Pdx1 Interaction in Diabetes	8
1.3 The Impact of the SPOP-BET Interaction on cancer drug susceptibility	9
2 Methods	11
2.1 Protein Expression and Purification	11
2.2 Protein Crystallography and X-Ray Structure Determination	12
2.3 NMR Spectroscopy for Ligand Binding and Protein Dynamics Studies	17
2.4 Fluorescence Polarization Binding Assays	21
2.5 Biophysical Characterization using Isothermal Titration Calorimetry	23
3 Publications.....	25
3.1 The Structure of the SPOP-Pdx1 Interface Reveals Insights into the Phosphorylation- Dependent Binding Regulation.....	25
3.2 Structural Insights into BET Client Recognition of Endometrial and Prostate Cancer- Associated SPOP Mutants	26
4 Conclusions	27
I References.....	30
II Abbreviations	34
III Acknowledgements.....	36
IV Appendix.....	37

Abstract

SPOP is the substrate recognition subunit of the cullin-3-RING E3 ubiquitin ligase complex. It is directly responsible for client protein binding and triggers their degradation via the ubiquitin-proteasome system. SPOP is a key regulator in several pathological processes. Pdx1, a vital pancreatic transcription factor is a client of SPOP. Under diabetogenic conditions, Pdx1 levels in pancreatic β -cells are diminished, leading to loss-of-function and death of insulin secreting cells. Here, two co-crystal structures of SPOP and Pdx1 are presented which give detailed information on the interface between the two proteins. Compared to a previously published SPOP-binding consensus sequence, the SPOP-binding site of Pdx1 shows an altered primary sequence. Additionally, the binding interface of Pdx1 is extended to the N-terminus. The SPOP-Pdx1 interface is thus distinct from other SPOP-substrate interfaces. Pdx1 was previously shown to be subject to post-translational modification via phosphorylation. A biophysical characterization of the SPOP-Pdx1 complex using phosphorylated Pdx1 indicates that phosphorylation in the SPOP-binding site of Pdx1 dramatically reduces affinity to SPOP. Pdx1 phosphorylation is therefore considered a key regulatory mechanism for SPOP binding. It potentially serves as a rescue mechanism for cellular Pdx1 levels, whose decline causes β -cell death, which is a hallmark of type 2 diabetes

SPOP-client binding is also relevant in cancer pathology. Recent studies identified the BET protein family as SPOP clients. An oncogenomics paradox was proposed, suggesting that SPOP mutations in the same domain lead to opposing BET inhibitor drug susceptibility in endometrial and prostate cancer. The point mutations presumably in- or decrease BET affinity of SPOP in the tumors respectively. Four co-crystal structures are presented which characterize the SPOP-BET client interface on a structural level. It was previously reported that prostate cancer-associated SPOP point mutations hinder BET client binding, which is confirmed by biophysical experiments such as NMR spectroscopy and fluorescence polarization. The observed loss-of-function in prostate cancer-associated SPOP mutants is explained on a mechanistic level by the presented co-crystal structures. As opposed to previously reported findings, in this study, endometrial cancer-associated SPOP mutants did not show altered client binding behavior compared to wild-type SPOP. Yet, SPOP mutations directly affect cellular BET protein levels in tumor cells. Thus, SPOP mutation screening is a promising personalized medicine approach to increase the efficacy of antitumor therapies in which BET inhibitor drugs are used.

Zusammenfassung

SPOP ist die für Substraterkennung zuständige Untereinheit des Cullin-3-RING E3 Ubiquitin Ligase Komplexes. Es bindet Substrate direkt und initiiert deren Abbau durch das Ubiquitin-Proteasom-System. SPOP ist ein wichtiger Regler in diversen pathologischen Vorgängen. Pdx1, ein zentraler Transkriptionsfaktor in der Bauchspeicheldrüse (Pankreas), ist ein Substrat von SPOP. Unter diabetogenen Bedingungen ist der Pdx1-Spiegel in den Insulin-sekretierenden β -Zellen verringert, was zu deren Funktionsverlust und Absterben führt. Hier werden zwei Co-Kristallstrukturen von SPOP und Pdx1 präsentiert, welche detaillierte Informationen über das Interface der beiden Proteine darbieten. Die SPOP-Bindestelle von Pdx1 weicht im Vergleich zu einer publizierten SPOP-Binde-Konsensussequenz in ihrer Aminosäureabfolge ab. Zusätzlich ist die Bindestelle in Pdx1 zum N-Terminus hin verlängert. Hierin unterscheidet sich das SPOP-Pdx1-Interface von den Bindestellen anderer SPOP-Substrat-Komplexe. Es wurde bereits gezeigt, dass Pdx1 durch Phosphorylierung posttranslational modifiziert wird. Die biophysikalische Charakterisierung des SPOP-Pdx1-Komplexes unter Verwendung von phosphoryliertem Pdx1 zeigt, dass die Phosphorylierung von Pdx1 die Affinität zu SPOP stark reduziert. Sie wird deshalb als wichtiger Regulationsmechanismus der SPOP-Bindung gesehen. Sie dient als möglicher Rettungsmechanismus für einen sinkenden zellulären Pdx1-Spiegel, welcher das Absterben der β -Zellen – ein Hauptmerkmal von Typ-2-Diabetes – auslöst.

Die SPOP-Substratinteraktion ist auch in der Tumorpathologie von Relevanz. Aktuelle Studien identifizierten die Familie der BET-Proteine als SPOP Substrate. Ein onkogenomisches Paradoxon wurde postuliert, demnach Mutationen in derselben SPOP Domäne gegensätzliche Anfälligkeiten für BET Inhibitoren in Prostata- und Gebärmutter-schleimhautkrebs bewirken. Die Punktmutationen stehen im Verdacht die BET-Affinität von SPOP zu verringern bzw. zu erhöhen. Vier Co-Kristallstrukturen werden präsentiert, die das SPOP-BET-Interface auf struktureller Ebene charakterisieren. Es war bekannt, dass Prostatakrebs-SPOP-Mutanten die Bindung von BET-Substraten stören. Dies wird durch biophysikalische Methoden bestätigt. Der beobachtete Funktionsverlust in den Prostatakrebsmutanten wird anhand der gezeigten Co-Kristallstrukturen mechanistisch erklärt. Gegenätzlich zu publizierten Resultaten haben die Gebärmutter-schleimhautkrebs-SPOP-Mutanten kein verändertes Bindeverhalten im Vergleich zum Wildtyp gezeigt. Dennoch beeinflussen SPOP-Mutanten direkt die BET-Spiegel in Tumorzellen. Deshalb wird ein SPOP-Mutationsscreening als ein vielversprechender Ansatz in der personalisierten Medizin gesehen, um die Effektivität von Antikrebstherapien mittels BET-Inhibitoren zu steigern.

1 Introduction

In the effort to combat disease and devise new treatment strategies, it is necessary to dissect and understand the underlying biological processes in the affected organism. In a functional organism, all biological processes are tightly regulated, often by complex pathways. If these pathways are disrupted or altered, manifestation of a disease can be the consequence. Biological pathways often rely on protein-protein interactions, which trigger chemical reactions or conformational changes to convey signals. Structural biology focuses on the characterization of three-dimensional structures of biomolecular complexes in order to find mechanistic explanations. It combines different techniques for structure determination as well as methods for studying conformational dynamics and biophysical parameters of biomolecules in an integrated fashion. The determination of complex protein structures enables the strategy of structure-based drug discovery. It utilizes the obtained stereochemical and biophysical parameters of a protein surface in order to design specific ligands that alter the target proteins functionality. This is referred to as a target-driven approach in structure-based drug discovery, which hopefully leads to the development of candidate molecules which can be turned into drugs in the future process. The focus of this thesis is to structurally characterize ligand complexes of the speckle-type POZ protein (SPOP) and to elucidate their role in different pathologies. This is of interest as SPOP is a component of a key cellular regulatory mechanism, the ubiquitin proteasome system (UPS) and has been linked to a variety of diseases in the past¹.

1.1 Biological Background of the Speckle-type POZ Protein (SPOP)

The ubiquitin proteasome system is a key regulation system on the cellular level, which is responsible for protein catabolism in all eukaryotic organisms². Proteins destined for proteasomal degradation are tagged with the 8.5 kDa regulatory protein ubiquitin. This process is mediated by a series of enzymes (Fig. 1). First, the E1 ubiquitin-activating enzyme activates and covalently binds free ubiquitin. Subsequently, it transfers the ubiquitin to the E2 ubiquitin-conjugating enzyme. E2 can bind to several different E3 ubiquitin-ligase proteins, which recognize the target proteins. After such a complex is formed, E2 continues to transfer the ubiquitin to the target protein. Proteins destined for proteasomal degradation are usually polyubiquitinated².

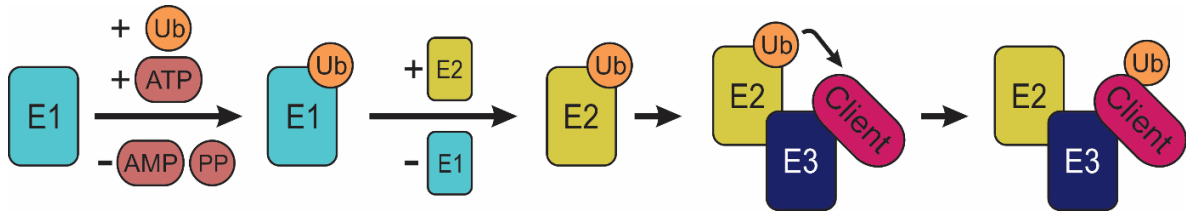


Figure 1: Schematic display of the ubiquitin transfer pathway to a client protein.

Ubiquitin (Ub) is first activated by an E1 ubiquitin-activating enzyme. Adenosine triphosphate (ATP) is consumed to form a reactive ubiquitin-adenosine monophosphate (AMP) intermediate, and subsequently a covalently bound E1-ubiquitin complex. Ubiquitin is then transferred to the E2 ubiquitin-conjugating enzyme. E2 is able to recognize a variety of E3 ubiquitin-ligase proteins. They serve as adapters to bring E2 in close contact with the client protein. Ubiquitin is ultimately transferred to the target protein by E2. Figure adapted from².

Comparative genome analyses indicated that in humans, the amount of E2 proteins exceeds the number of E1 proteins by a factor of approx. three. Moreover, there are approx. ten times more E3 than E2 enzymes³. This variety, especially in E2 and E3 proteins enables the UPS to specifically target certain protein for degradation, while others are left intact². Although a number of E3 ubiquitin ligases exist, some seem to be more prominently involved in pathogenic processes. The cullin-3-RING ligase complex (CRL) is such an E3 ligase complex which was shown to play critical roles in several physiological and pathological processes, acting e.g. as tumor suppressor in prostate cancer⁴.

As opposed to other E3 ligase complexes, in CRLs, a single, BTB (Broad-Complex, Tramtrack and Bric a brac) domain containing adapter protein links cullin-3 to the ubiquitination substrate (Fig. 2A). That adapter protein is directly recognizing and binding the client protein⁵.

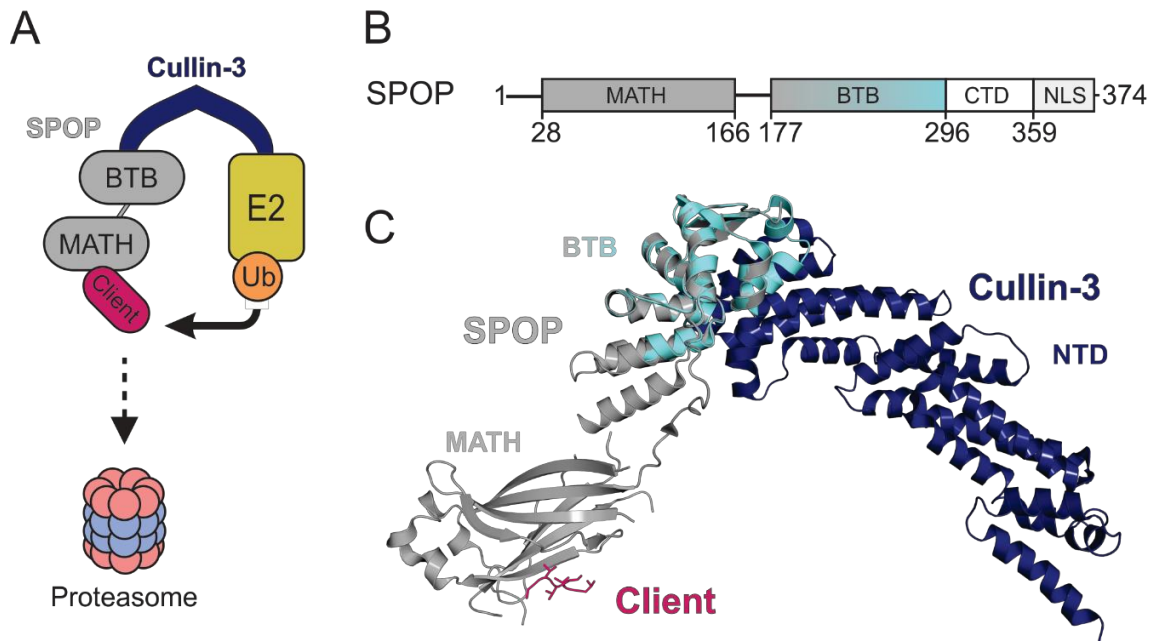


Figure 2: Overview of function and domain organization of the E3 ubiquitin-ligase adapter protein SPOP.

A: Schematic display of the SPOP-cullin-3-RING ligase complex.

In cullin-3-RING ligase complexes, cullin-3 is connected to the substrate protein via a single polypeptide chain. The SPOP protein is directly responsible for target recognition. It binds the client protein with its MATH domain. The BTB domain binds to cullin-3, which in turn binds the E2 ubiquitin-conjugating enzyme. The CRL thus brings E2 and the client protein in proximity, enabling ubiquitin transfer. The ubiquitinated client is targeted for proteasomal degradation.

B: Overview of SPOP domain organization.

SPOP contains three domains: the substrate-binding MATH domain, the cullin-3-binding BTB domain and the C-terminal domain (CTD) responsible for SPOP oligomerization. It also contains an N-terminal nuclear localization signal (NLS). Domain boundaries by residue numbers are indicated.

C: Overlay of co-crystal structures visualizing the SPOP-cullin-3 complex with ubiquitination target.

PDB 3HQ1 and PDB 4EOZ were aligned to create a model of the client-SPOP-cullin-3 complex using PyMOL⁶. PDB 3HQ1 contains the complex of SPOP MATH-BTB (grey) and a peptide from SPOP client Puc (pink). PDB 4EOZ contains the complex of SPOP BTB (cyan) and cullin-3 N-terminal domain (NTD, dark blue). SPOP BTB domains from both co-crystal structures are highly similar. The model demonstrates the adapter function of SPOP, connecting the client protein bound by the MATH domain to cullin-3, bound by the BTB domain.

A main subject of study in this thesis is the speckle-type POZ protein (SPOP), which was shown to be such a cullin-3 adapter protein⁷. It contains three domains (Fig. 2B), an N-terminal MATH (meprin and TRAF homology) domain for substrate recognition, an internal BTB domain, also known as POZ (poxvirus and zinc finger) domain for cullin-3 binding (Fig. 2C) and a C-terminal domain (CTD) which mediates SPOP oligomerization. The protein also contains a nuclear localization signal (NLS)⁸. Due to structural features of its CTD and BTB domain, SPOP is prone to oligomerize and form liquid nuclear speckles which increase protein concentration and thus ubiquitination turnover in these areas⁹. Recent findings suggest that the enzymatic activity is correlated with the liquid-liquid phase separation and that this process is disrupted by disease-linked mutations¹⁰.

The SPOP protein as subunit of CRLs was shown to be involved in a number of pathological processes such as kidney, prostate and endometrial cancer¹¹. Reports suggest that its role in tumorigenesis seems highly dependent on the tissue and cellular context as it acts as e.g. as a tumor suppressor in prostate cancer⁴, while having tumorigenic effects in kidney cancer due to mislocalization¹². Thus far, several SPOP substrates have been identified. In this study, two SPOP substrate proteins critically involved in pathogenic processes involved in diabetes and cancer were selected for biophysical and structural characterization to explain the underlying mechanisms on a molecular basis.

1.2 The Influence of SPOP-Pdx1 Interaction in Diabetes

Pdx1 (pancreas / duodenum homeobox protein 1) is a transcription factor vital for the formation of a functional pancreas during embryogenesis¹³. In adult organisms, Pdx1 regulates the expression of genes critical for glucose homeostasis such as insulin, IAPP (islet amyloid polypeptide) and GLUT2 (glucose transporter 2)¹⁴. It is also responsible for the maintenance and survival of the β -cells as such^{15,16}.

Diabetes mellitus is a metabolic disorder, where a combination of factors leads to loss-of-function in glucose homeostasis. In both type 1 and type 2 diabetes mellitus (T1D or T2D, respectively), progressive β -cell death is a hallmark of the disease. In T1D, β -cells are targeted by an autoimmune response of the body¹⁷, while T2D is characterized by insulin resistance and hyperglycemia, diabetogenic conditions which adversely affect pancreatic β -cells causing gradual loss-of-function¹⁸. The fate of the β -cells seems to be tightly linked to cellular Pdx1 protein levels, as studies have shown decreased Pdx1 expression levels in human islet cells under diabetogenic conditions¹⁹. Furthermore, a rescuing effect of preserved Pdx1 expression on β -cell failure in diabetic mice was observed²⁰. However, little is known about the molecular processes controlling Pdx1 protein turnover in the β -cells.

The SPOP protein was first identified as PCIF1 (Pdx1 C-terminal interacting factor), a protein with the ability to inhibit Pdx1 transcriptional activity²¹. Subsequently, it was shown that Pdx1 is ubiquitinated and targeted for proteasomal degradation by SPOP as part of a cullin-3 containing E3 ubiquitin-ligase complex²². SPOP is the component that forms the direct link between Pdx1 and the UPS. The SPOP-Pdx1 interaction is thus considered a critical step in controlling Pdx1 protein levels and ultimately β -cell survival and glucose homeostasis.

Further studies on Pdx1 post-translational modifications, beyond ubiquitination, have revealed that the protein is phosphorylated by kinases such as MST1 (mammalian sterile 20-like kinase-1)²³ or CK2 (casein kinase II)²⁴. In the case of CK2, the phosphorylation site was localized directly to the SPOP-binding consensus (SBC) of Pdx1, to residues Thr230 and Ser231. It was postulated that Pdx1 phosphorylation at these residues would increase the SPOP-binding rate, thus diminishing Pdx1 stability in its phosphorylated form²⁵.

A major aim of this thesis was to obtain previously unavailable structural data of the SPOP-Pdx1 interface and to analyze the interaction using biophysical techniques. It was also of key interest to study the effect of phosphorylation-dependent SPOP binding of Pdx1 on a mechanistic level, as the interaction of the two proteins could potentially prove to be a valuable therapeutic target for treatment or prevention of T2D.

1.3 The Impact of the SPOP-BET Interaction on cancer drug susceptibility

Serving as the substrate recognition subunit of larger E3 ligase complexes, SPOP targets a variety of ligands²⁶. Recently, the BET (bromodomain and extraterminal domain) family of proteins was identified as being SPOP clients²⁷⁻²⁹. The ubiquitously expressed BET proteins BRD2, BRD3, BRD4 and the testis-specific BRD-T are involved in chromatin-remodeling³⁰. They specifically bind to acetylated lysines, which was shown via NMR spectroscopy³¹. They mediate transcriptional activation by recruiting a variety of transcriptional activator proteins and transcription factors³². Due to their function as transcriptional coactivators, BET proteins were found to be involved in the pathology of tumors. BRD4 for example was shown to be part of a fusion oncogene, causing aggressive forms of carcinoma³³. BET proteins were also found to modulate transcription of oncogenes such as c-Myc³⁴. As they serve as oncogenic co-activators, small molecule inhibitors were developed to inhibit aberrant function of the BET proteins. Inhibitors such as JQ1 compete with BET proteins for acetyl-lysine binding and inhibit BET-regulated tumor proliferation³⁵. The efficacy of BET inhibitors was shown for a number of different cancer types such as hepatocellular carcinoma, colon cancer, pancreatic cancer and prostate cancer (for a detailed overview of the individual studies see³⁶). Yet, as BET inhibitors were tested in clinical trials, BET inhibitor-resistant tumors were detected³⁷.

In prostate cancer, it was postulated that BET inhibitor resistance is caused by frequent mutations in the SPOP protein, which normally targets BET proteins for proteasomal degradation²⁸. A loss-of-function in BET protein binding of prostate cancer-associated SPOP mutants was suggested, leading to the accumulation of BET proteins in the tumor cells^{28,29}.

SPOP mutations were also found in other cancer types, such as endometrial cancer. Interestingly, in this cancer type, reduced cellular BET protein levels, concomitant with increased BET inhibitor drug susceptibility were observed²⁹. This reveals an oncogenomic paradox, since mutations mapping to the SPOP MATH domain were postulated to cause opposing drug susceptibilities in two different cancer types²⁹. As these observations are based on data obtained from cell culture experiments, no mechanistic data on the SPOP-BET interaction was available.

A second aim of this thesis was to study the SPOP-BET interaction, especially of mutated SPOP protein, on a structural and mechanistic level. To find the mechanistic explanation of the observed effects, the BET binding behavior of 14 SPOP mutants, both endometrial and prostate cancer-associated, was assessed and compared to the wild-type protein. Several co-crystal structures of SPOP mutants with BRD3, a representative member of the BET protein family were obtained in order to find a structural basis for the postulated altered BET protein binding behavior of prostate and endometrial cancer-associated SPOP mutants. As SPOP regulates BET protein levels via the ubiquitin-proteasome system²⁷⁻²⁹, screening for SPOP mutations in cancer patients might prove a valuable approach to boost the efficacy of BET-targeted cancer treatment via personalized medicine.

2 Methods

2.1 Protein Expression and Purification

This section provides an overview of the general procedures used for expression and purification of proteins. Detailed procedures for each protein construct are given in the methods section of the respective publications.

Constructs of the SPOP and BRD3 proteins were cloned into the pETM-11 vector (Gunter Stier, EMBL Heidelberg). The constructs contained an N-terminal His₆-tag followed by a TEV (tobacco etch virus) protease cleavage site. Constructs of the Pdx1 protein were cloned into the pET_hSu vector, adding an N-terminal His₆-tag followed by the SUMO (Small ubiquitin-related modifier 5) fusion protein. The vectors were transformed into the *E. coli* strains BL21 (DE3) or Rosetta2 (DE3) for protein expression.

The bacterial expression cultures were grown in ZYM-5052 rich auto-induction medium³⁸ if the respective protein products did not require isotope labeling. If isotope labeling was required for NMR experiments, M9-based minimal media were used. The minimal media were supplemented with ¹⁵NH₄Cl for uniform labeling with ¹⁵N and additionally with ¹³C glucose for uniform ¹⁵N/¹³C labeling, e.g. for backbone resonance assignment experiments. The bacterial cultures were grown at 37°C until the OD₆₀₀ (optical density at a wavelength of 600 nm) reached a value of 0.7-1.0 and then transferred to 20°C. To M9-based media, 1 mM IPTG (isopropyl β-D-1-thiogalactopyranoside) was added to induce protein expression.

The proteins were purified from bacterial lysates via IMAC (immobilized metal affinity chromatography) using a Ni-NTA column. Lysates containing Pdx1 proteins were previously subjected to PEI (polyethyleneimine) precipitation of DNA (deoxyribonucleic acid), followed by protein precipitation via ammonium sulfate. After the first IMAC step, expression tags were removed from the target proteins by adding TEV protease or yeast SUMO hydrolase dtUD1, respectively. Proteins were then subjected to a second IMAC step, followed by a final purification via size exclusion chromatography. All purified protein samples were analyzed via SDS-PAGE (sodium dodecyl sulfate polyacrylamide gel electrophoresis) prior to their use in experimental procedures.

2.2 Protein Crystallography and X-Ray Structure Determination

Serving as “molecular machines”, proteins not only catalyze chemical reactions but also convey a multitude of other functions such as cellular signaling and regulation of gene transcription. For the elucidation of such biological mechanisms, it is of utmost importance to study the three-dimensional structure of the respective proteins. Protein crystallography and subsequent X-ray structure determination are invaluable techniques which provide such three-dimensional structures of individual proteins or protein-protein complexes with up to atomic resolution. In this thesis, several co-crystal structures were solved in order to study the binding interface of SPOP with different ligand proteins.

For X-ray structure determination, crystals of the respective protein complex must first be produced. Although a single protein crystal may be enough to solve a structure, this remains very much a limiting step, as not all proteins crystallize readily. In order to obtain crystals, an extremely pure and homogenous sample of the protein of interest must be produced. The sample protein should be highly concentrated, ideally close to its saturation limit in a favorable buffer solution. The solution is then brought to a state of supersaturation, exceeding the protein’s solubility limit. This can be achieved by manipulating conditions such as ionic strength, pH or temperature of the sample. Also, salts or organic precipitants can be added. To reestablish equilibrium in the sample after supersaturation, a solid phase is formed, ideally in the form of ordered protein crystals³⁹. These crystals may form a broad variety of shapes (Fig. 3). After harvesting and preserving the crystals in liquid nitrogen, they are irradiated with X-rays in order to obtain diffraction patterns from which the three-dimensional protein structure can be calculated.

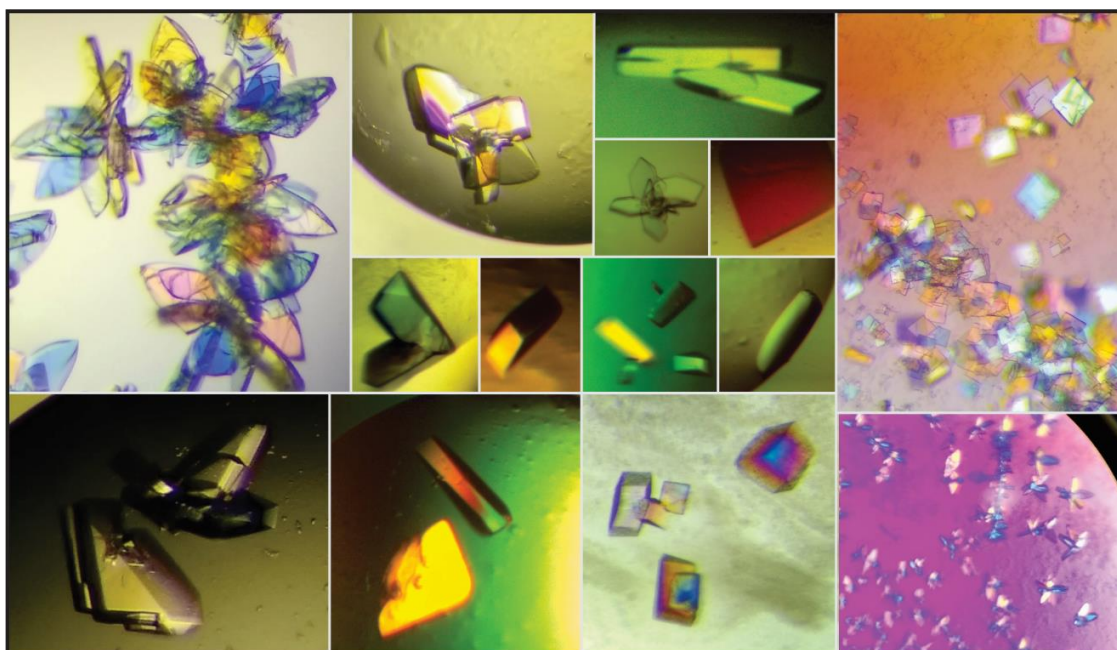


Figure 3: Compilation of photos taken from protein crystals created during this thesis.

To date, there are no means of predicting a protein's optimal crystallization conditions. It is rather a "try and error" approach, where lots of different conditions are tested systematically. From the different techniques available for protein crystallization experiments, the vapor diffusion setup in sitting-drop format was employed in this thesis (Fig. 4A). In this method, a small drop (approx. 200 nl) of buffered protein solution is placed in a container together with a large volume of reservoir solution. To the drop, an equal volume (e.g. 200 nl) of the reservoir solution is added, allowing the testing of numerous different conditions depending on which reservoir solution is used. The mixed drop and the remaining reservoir solution in the container are physically separated, e.g. by placing the drop on an elevated pedestal (Fig. 4A). The container is sealed, so that no air exchange with the surrounding environment is possible. Initially, the drop contains a high protein concentration and low salt and/or precipitant concentration. The reservoir solution contains a high concentration of salt or other precipitant (i.e. polyethylene glycol). Over time, the solvent of the protein drop evaporates and diffuses towards the reservoir solution. This gradually increases the protein concentration in the drop until it reaches a state of supersaturation. To reequilibrate, a solid phase, ideally consisting of ordered protein crystals is formed. These crystals are brittle and very susceptible to drying. As they show relatively high solvent contents (on average about 50%), they possess a rather gel-like state, where all protein molecules

are surrounded by a hydrate shell. Solvent molecules can diffuse freely through the large intermolecular spaces between the protein bodies or even into accessible pockets on the protein surface. This property allows for so-called soaking, a method in which existing protein crystals are incubated with a ligand which is then taken up and incorporated into a complex structure. In this thesis however, both components, protein and ligand were co-crystallized, meaning that they were both present in the initial sample used for the crystallization setup.

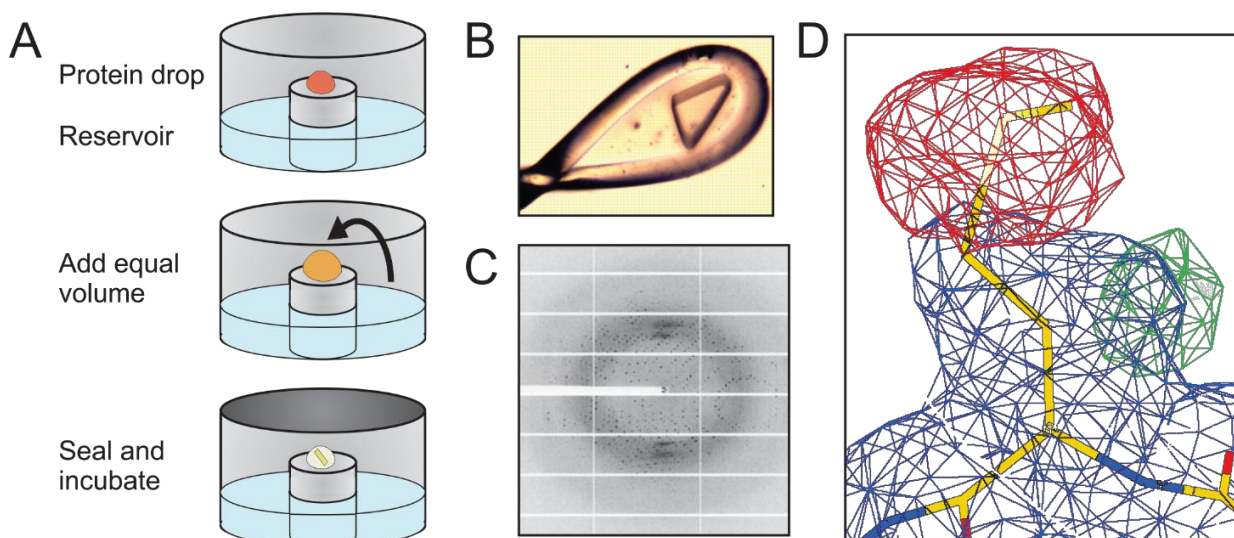


Figure 4: Examples of protein crystallization setup, diffraction data acquisition and structural data visualization. A: Setup of vapor-diffusion, sitting drop protein crystallization experiments.

A drop of highly concentrated protein solution is placed on a separated pedestal in a container with crystallization solution. An equal volume of this reservoir solution is added to the protein drop. The container is sealed and incubated. Over time the drop's solvent vaporizes and the solution becomes supersaturated. Crystals may form as a solid phase is generated in order to reequilibrate.

B: Loop containing a flash-frozen protein crystal mounted on X-ray beamline.

Protein crystals are fished using a nylon loop and flash frozen in liquid nitrogen. At the beamline, the loop containing the crystal is mounted onto a goniometer for exact positioning. Liquid nitrogen cooling must be maintained at all times.

C: Diffraction pattern image obtained from protein crystal.

After X-ray exposure of the crystal, the reflections are recorded by a detector. Spots closer to the midpoint indicate higher resolution of the obtained data. As the crystal is rotated incrementally, several hundred or thousands of such images are recorded, which are then used to calculate an electron density map.

D: Graphical representation of electron density map with modeled protein coordinates.

The electron density map is used as model to fit the crystallized protein's residues. The model is then iteratively refined by the user to fulfill geometrical, chemical and sterical restraints. For this, the coot software⁴⁰ was used in this thesis. The modeled protein residues (without hydrogens) are shown as sticks. Carbon atoms are shown in yellow, nitrogen atoms are shown in blue, oxygen atoms are shown in red, sulfur atoms are shown in bright yellow. The electron density map is shown as mesh. Blue color indicates that the model fits well to the experimental data. Red color indicates that the model includes atoms for which no density is present in the recorded data. Green color indicates that the experimental data contains density which has not been (sufficiently) modeled. The picture shows data obtained on SPOP mutant M117V. As the wild-type amino acid sequence was used for initial modeling, methionine is now present in the structure. The electron density map shows that the experimental data does not support a methionine sidechain at this position. The map also shows unmodeled (green) density to the right, corresponding to the missing valine's methyl group. The residues have to be exchanged in the model in order to match the experimental data.

The tendency of proteins to form ordered crystals over random aggregates is highly dependent on the surrounding conditions. By mixing the initial sample with different reservoir solutions, a vast variety of conditions can be tested. These may include different types of salts, precipitants or additives in different concentrations. The pH of the sample may also be iterated, as well as the surrounding temperature. To conveniently screen a multitude of different crystallization conditions, commercial screens are available, which were used in this thesis in a 96-well plate setup. The use of different solutions may have drastic effects on the outcome of the crystallization experiment. If crystals are at all obtained, their quality in the later data collection process may differ strongly based on the conditions in which they were produced.

Protein crystallization is a dynamic process, as crystals grow over time from a nucleation core, and may also degrade again in given time. If a crystal large enough for data collection is produced, it must be harvested and conserved. For this, the crystal is fished from its container using a tiny nylon loop (Fig. 4B). It is then mixed with a cryoprotectant such as glycerol and immediately frozen in liquid nitrogen, where it must remain until data collection.

The data later used for structure calculation are obtained from protein crystals via X-ray diffraction. In this method, a single protein crystal (frozen in the fishing loop) is mounted in a goniometer under a liquid nitrogen stream. This minimizes radiation damage to the crystal. The goniometer allows for accurate positioning in the X-ray beam used for measurement. When the focused X-ray beam hits the crystal, a diffraction pattern is produced that can be recorded by a detector (Fig. 4C). The crystal is then rotated in small increments in the X-ray beam, in order to get a multitude of diffraction patterns from different angles. The compendium of images is then used to create a three-dimensional electron density map via a series of complex mathematical operations such as indexing (identification of the unit cell dimensions and determination of the crystal space group) and merging (identification of peaks recorded in two or more images and scaling of their intensities). The collected data is the representation of the electron density map in reciprocal space. In order to generate the real space map, two parameters are required. The amplitude and phase of a wave function. The amplitude is given by the intensities obtained from the diffraction, but the phase information is not included in the data output. In this thesis, molecular replacement (MR) was used to solve this problem. For this method, a crystal structure similar to the one being solved is required. As structures of SPOP-ligand complexes were solved in this thesis, available structures of SPOP from the PDB were utilized as search models for molecular replacement. The MR software takes such a structure and tries to relate the experimental information to it in order to create an electron density model that best fits the experimental data.

The output of the MR routine is an initial model, which is then iteratively refined to fit better and better to the recorded electron density. In between each refinement step, the electron density map and the therein fitted protein atoms are visualized (Fig. 4D). The fit between the proteins' coordinates and the electron density map should then be manually improved. This process is repeated until the model and the diffraction data achieve a high correlation. Geometrical and sterical constraints for the individual amino acids such as root-mean-square deviation (RMSD) ranges for bond lengths and angles should be fulfilled to create a sensible model. The most important statistical value is the so-called R-factor. It represents the overall agreement of the modeled structure to the actual diffraction data. The lower this value is, the better the correlation between the model and the recorded data. The range for the R-factor is between 0 for a (theoretical) perfect fit and 0.63 for a totally random set of atoms fitted to the diffraction pattern. Typical values are around 0.20, as this value is affected by the large portion of unstructured solvent molecules⁴¹. A second statistical value, R-free is used to avoid overfitting bias which is introduced due to the way that the R-factor is calculated. For the determination of R-free, 10% of the observed data is excluded from the data set prior to refinement. R-free is then calculated to determine in which quality the refined model is able to predict the previously excluded data. For a good, not over-fitted dataset, the R-free value should be similar to the R-factor⁴¹.

A reliable crystal structure is invaluable for the study of biological processes. In this thesis, crystal structures were solved in order to study the interaction of SPOP with different ligand proteins. These co-crystal structures revealed insights into the interactions formed between the proteins on an atomic level. The presence of hydrogen bonds (H-bonds), hydrophobic interactions or ionic interactions is revealed and gives information about the nature of the interface. Based on crystal structures, ligand binding sites or the impact of possible mutations in certain regions can be determined. To date, X-ray crystallography remains the gold standard in the determination of three-dimensional protein structures.

2.3 NMR Spectroscopy for Ligand Binding and Protein Dynamics Studies

As opposed to X-ray crystallography, nuclear magnetic resonance spectroscopy is able to provide structural data and information about interactions and conformational dynamics in solution. It can be used for a multitude of studies, e.g. to distinguish structured and flexible regions of a protein or to characterize a possible interaction of two biomolecules.

NMR spectroscopy relies on the magnetic properties of certain atoms which can be used in order to generate visualized spectra. As the behavior of nuclei is highly dependent on the chemical environment, detailed information can be derived.

An important value in NMR spectroscopy is I , the nuclear spin quantum number. This value is dependent from the proton and neutron count of the respective atom. If an atom has a spin of 0, caused by an equal proton and neutron count, it is inactive in NMR spectroscopy. All atoms with unequal proton/neutron counts do have nuclear spins and could technically be used for NMR. In practice, nuclei used in biomolecular NMR spectroscopy have a spin of $\frac{1}{2}$, as their magnetic states have a much longer lifetime than those of atoms with higher spin numbers^{42,43}. These commonly used nuclei are ^1H , ^{15}N , ^{13}C and to a lesser extent ^{19}F and ^{31}P .

Such nuclei possess a directional magnetic moment, characterized as μ (Fig.5A). This value is dependent from a nucleus specific value, the gyromagnetic ratio γ . In NMR spectroscopy, the nuclei in a sample are exposed to a strong external magnetic field B_0 , which influences the orientation of the nuclei, similar to how two bar magnets would reorient each other when brought into proximity. In the external magnetic field, the magnetic moments of the nuclei are oriented in either one of two distinct states, along or opposed the external field (Fig. 5B). The orientation along the external magnetic field is the lower energy state, which is slightly preferred. The orientation opposed to the external magnetic field is the higher energy state. This splitting into two states is called the *Zeeman* effect (Fig. 5C). While in the external magnetic field, the magnetic moment of the nuclei still moves in a circular fashion around B_0 . The angular frequency of this movement is described as ω_0 , the *Larmor* frequency^{42,43}. The relation of the mentioned quantities can be expressed as follows, whereas \hbar is the reduced Planck constant:

$$\Delta E = \hbar\omega_0 = \gamma\hbar B_0$$

See ⁴³ for derivation of the equation and further reading.

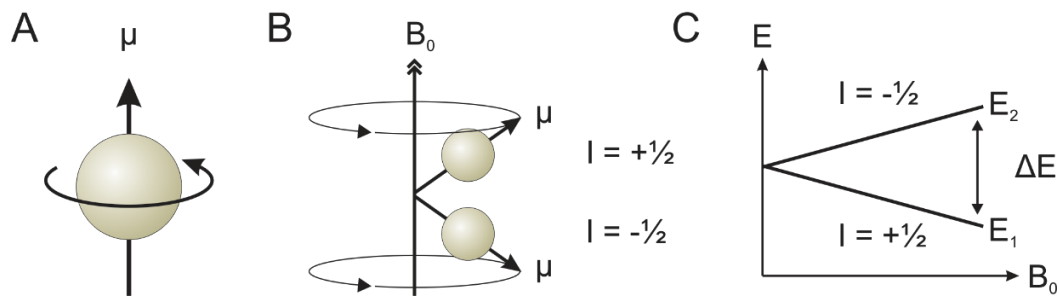


Figure 5: Principles of nuclear magnetism⁴³.

A: Simplified model of a spinning nucleus.

An NMR active nucleus ($I \neq 0$) can be visualized as a rotating sphere with a magnetic moment μ .

B: Behavior of a spinning nuclei in an external magnetic field.

If an external magnetic field B_0 is applied, the magnetic moments of the nuclei reorient. The magnetic moments align either along or opposed to the magnetic field, which splits them into two groups (*Zeeman* effect). Nuclei with a quantum number of $I = \frac{1}{2}$ (e.g. ^1H) can assume two modes of orientation. The orientation along B_0 with $I = +\frac{1}{2}$ is energetically slightly more favored over the opposed orientation ($I = -\frac{1}{2}$).

C: Energy distribution of spins in a magnetic field.

The energy difference between the two states, ΔE is dependent from B_0 and the *Larmor* frequency of the respective nucleus, which can be used to affect the energy distribution of the sample. This gives rise to the NMR signal.

In an NMR measurement, a second magnetic field B_1 is applied to the sample at radiofrequency wavelengths. If the applied wavelength matches the *Larmor* frequency of the observed nuclei, they are able to absorb energy in a process called resonance⁴³. Nuclei in the low energy state can transition to the high energy state, resulting in a detectable NMR signal. As the local chemical and electronical propensities affect the resonance frequencies of each nucleus, the same type of nucleus may show slightly different resonance frequencies reflecting its direct environment. This effect is called chemical shift, an important readout in NMR spectroscopy^{42,43}.

In protein NMR spectroscopy, two- or more-dimensional techniques are usually preferred, meaning that magnetization is transferred from one NMR-active nucleus to another in between excitation and detection. This is possible via scalar couplings mediated through chemical bonds. A commonly used two-dimensional NMR technique is the HMQC (heteronuclear correlation through multiple quantum coherence) spectrum. For this experiment, a protein sample uniformly labeled with ^{15}N is required. During the experiment, ^1H nuclei are first excited, then the magnetization is transferred to ^{15}N nuclei and subsequently transferred back to ^1H for detection⁴³. The resulting spectrum shows one signal per ^{15}N - ^1H (amide) pair. Such a spectrum recorded on a protein is sometimes referred to as “fingerprint” of the respective protein. With some exceptions, it shows one peak per amino acid contained in the protein. The distinct shape of the spectrum, representing the chemical environments of each amide pair can be used for identification of a known protein, quality and purity control of a sample or ligand binding studies (Fig. 6).

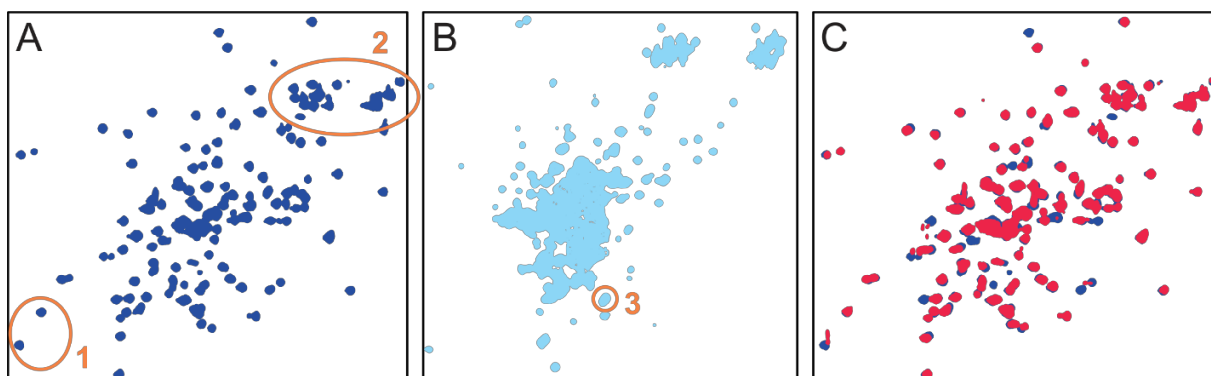


Figure 6: Examples and applications of ^1H - ^{15}N HMQC NMR spectra of different proteins.

A: ^1H - ^{15}N HMQC spectrum of a well-folded protein (SPOP).

The spectrum serves as a characteristic “fingerprint” of the protein. With the exception of proline, each residue produces at least one peak. Certain amino acids produce additional peaks due to side chain amides, which appear in specific regions indicated in orange. (1) Trp sidechain region. (2) Asn and Gln sidechain region, peaks typically appear as pairs. The spectrum can be used for identification and quality control of a protein sample.

B: ^1H - ^{15}N HMQC spectrum of an unfolded protein sample.

This spectrum is characterized by lots of peaks overlapping in the center region. Overlapping peaks indicate similar chemical environments of the corresponding residues. This is the case for residues that are solvent exposed, without presence of intramolecular contacts. The spectrum thus indicates a (largely) unfolded protein sample.

Yet, as in most HMQC spectra, the N-terminal residue produces a strong signal in a specific region (3).

C: Overlay of two ^1H - ^{15}N HMQC spectra of SPOP protein to visualize ligand binding.

By overlaying two HMQC spectra of the same protein, the impact of different sample conditions on a protein can be studied. Here, the blue spectrum was recorded on a protein without ligand, while the red spectrum was recorded on a protein sample with excess of a peptide ligand. The overlay visualizes changes between the spectra such as peak disappearances (line broadening) or peak shifts upon ligand addition, which clearly indicates binding. After resonance assignment, the ligand binding site could be identified by analyzing which peaks/residues show line broadening or chemical shifts.

In order to use the full potential of this NMR technique, backbone chemical shift assignment is necessary. This means that each visible peak in the spectrum is unambiguously assigned to the amide pair of a specific amino acid in the primary protein sequence. Note, that proline is not visible in ^1H - ^{15}N correlation spectra. Also, amides from amino acid sidechains are able to produce visible peaks. The resonance assignment for the two-dimensional ^1H - ^{15}N HMQC spectrum can be obtained by employing three-dimensional NMR techniques such as HNCACB⁴⁴ or CBCACONH⁴⁵. These techniques require protein samples uniformly labeled with ^{15}N and ^{13}C . In these experiments, magnetization is transferred from ^1H nuclei via ^{13}C nuclei and ^{15}N nuclei to amide ^1H nuclei for detection, meaning that the readout contains information about the carbon atoms the respective amide proton is coupled to. Due to the high sensitivity of the chemical shifts towards the chemical environment, the ^{13}C atoms contained in different types of amino acids show distinct shifts. As information about the adjacent residue is contained, the peaks can be correlated and subsequently assigned. Assigned HMQC spectra can be used e.g. to determine the binding site of a ligand, or study the involvement of certain residues or protein regions in binding processes.

Two- or three-dimensional NMR techniques can be used to study protein dynamics in solution. The ^{13}C shift values (secondary chemical shifts) e.g. from HNCACB experiments can be used to study the secondary structure of a protein. As the chemical shifts are strongly affected by the chemical environment of the respective nucleus, the shifts of ^{13}C atoms oriented in alpha-helices differ significantly from the carbon shifts observed in beta-sheets or random coil motives. By comparison to reference values for each residue type, regions that form either of these three structural motifs can be identified.

The experiments described so far use through-bond magnetization transfer via scalar couplings. In some cases however, it is feasible to use through-space magnetization transfer, especially to obtain restraints for structure calculation or to study the flexibility of different regions in the protein. Such a through-space magnetization transfer occurs between nuclei in close spatial proximity (approx. 3-6 Å), based on the nuclear Overhauser effect (NOE). The ^1H - ^{15}N heteronuclear NOE (hetNOE) experiment uses this principle to gain information about the flexibility of residues in the sample protein and helps to discern unstructured, flexible regions from rigid areas forming secondary structure elements.

In this thesis, several proteins were studied using the described NMR techniques. After obtaining a backbone resonance assignment, HMQC spectra or the related HSQC (heteronuclear single quantum coherence) spectra were used to prove biomolecular interactions and to study ligand binding behavior. The secondary structure features and the intrinsic flexibility of proteins were studied using secondary chemical shifts and heteronuclear NOE spectra, respectively. Generally, NMR spectroscopy is a versatile tool that can provide information on many different aspects of biomolecular structure and interaction in their native liquid state.

2.4 Fluorescence Polarization Binding Assays

Assays based on fluorescence polarization (FP) or fluorescence anisotropy (FA) provide a fast and cost-efficient way of characterizing biomolecular interactions. In a simple case, samples containing two components suffice to determine their binding affinity by measuring a binary titration curve. The assay can also be adjusted for competitive measurements, e.g. to test if an added compound is able to disrupt a pre-formed complex. FP assays can be performed in 384-well plates, which enables usage in high-throughput situations and provides excellent comparability when titrations with different binding partners are prepared on a single plate. So far, a diverse range of applications for FP assays has been described, such as substrate and cofactor binding studies on the bacterial carbamoyl-phosphate synthetase (CPS) protein⁴⁶, or the quantitative characterization of RNA (ribonucleic acid) synthesis of influenza virus polymerase⁴⁷. Sometimes, intrinsic fluorescence of tyrosine residues in proteins can be exploited to characterize interactions⁴⁸. However, in most cases it is feasible to introduce a chemical fluorophore as reporter, e.g. by using a fluorescently labeled small molecule or peptide ligand to study its binding to another macromolecule such as a protein or nucleic acid.

Such a sample of two binding partners, one of them fluorescently labeled, is then exposed to linearly polarized light, generated by an excitation polarizer. After passing the sample, the light is detected by an analyzing polarizer, with fixed orientations relatively to the excitation polarizer, e.g. parallel and perpendicular (Fig. 7A). Thus, two fluorescence intensities are measured. They are subtracted and normalized to obtain the fluorescence polarization value P ⁴⁹. The contribution of the sample itself to the total polarization value is dependent on the molecular motion of the fluorophore during the fluorescence lifetime⁵⁰. If the fluorophore molecules reorientate randomly within the excitation lifetime, the total sample polarization should be 0. If a net orientation of the fluorophores remains, a total polarization of the sample is observed. This usually indicates binding of the fluorophore-containing molecule to a biological macromolecule as increased molecular weight greatly reduces the molecular motion of the complex⁴⁹.

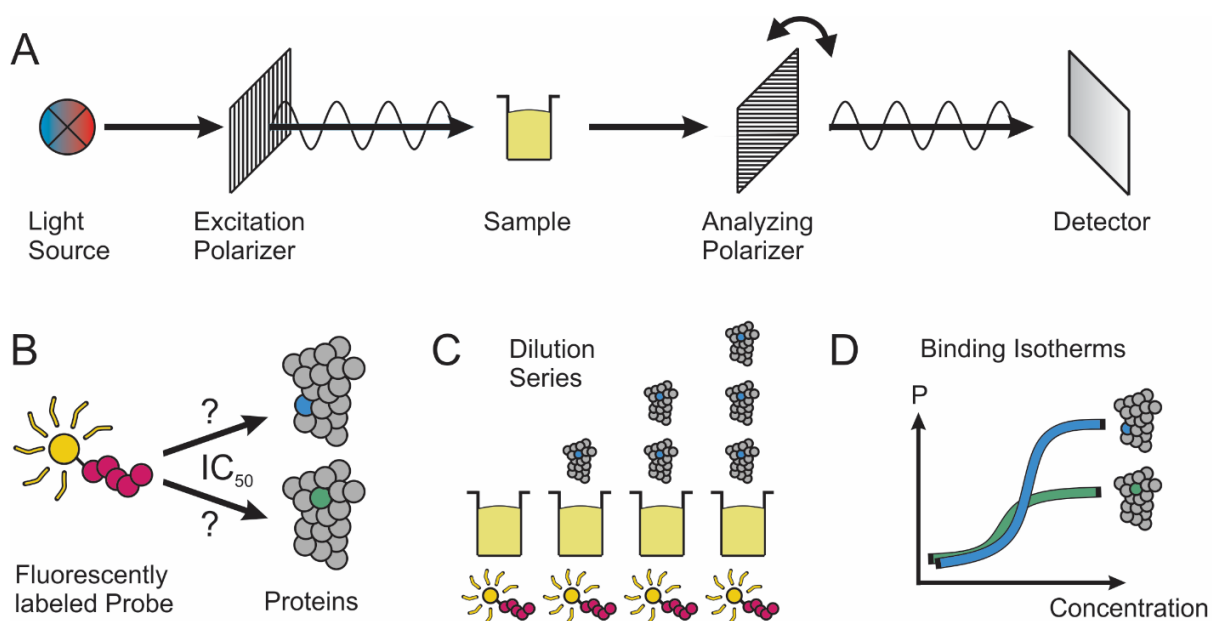


Figure 7: Instrumental and experimental setup of fluorescence polarization (FP) based binding assays.

A: Scheme of an assembly for measuring of fluorescence polarization.

Light is emitted at specific wavelengths for fluorescence excitation in the sample. The light is polarized by an excitation polarizer prior to passing the sample. Depending on the fluorophore environment, the sample may alter the polarization of the light. The emitted light is analyzed by a second polarizer which can be oriented in different angles relative to the excitation polarizer (usually parallel and perpendicular). The difference in detected light intensity between the two polarizer orientations is the fluorescence polarization value P .

B: Scenario for the utilization of a fluorescence polarization assay.

The binding strength (e.g. IC_{50}) of one molecule to another, such as a peptide or small molecule to a protein can be determined via fluorescence polarization assays. This requires labeling one component (usually the smaller one) with a fluorescent dye.

C: Preparation of binary fluorescence polarization assay.

Several samples with a constant concentration of fluorescently labeled probe are prepared. The binding partner is then added in a serial dilution, usually on assay plates. The binding behavior of the probe to different binding partners can be tested parallel.

D: Experimental readout of binary fluorescence polarization binding assays.

The assay readout P for each well is plotted over the binding partner concentration (or molar ratio compared to the probe). This generates binding isotherms for each tested molecule pair. These can be used for the determination of IC_{50} and enable easy comparison of binding intensities. Here, the blue curve indicates stronger binding compared to the green one.

In this thesis, FP assays using a fluorescein-tagged peptide were employed to quantitatively study the binding behavior of BET proteins to different mutants of the SPOP protein (Fig. 7B). In the experimental setup, the fluorescently labeled peptide was kept at a constant concentration, while different SPOP proteins were added in separate dilution series (Fig. 7C). By plotting the fluorescence polarization readout of each sample over the respective protein concentration, binding isotherms for each respective peptide-protein combination were obtained (Fig. 7D). They allowed easy comparison of the binding strength of different SPOP mutants to the BET proteins.

2.5 Biophysical Characterization using Isothermal Titration Calorimetry

Isothermal titration calorimetry is a technique for biophysical characterization of interactions between molecules. It is commonly used to determine binding affinities and often referred to as the gold standard in the determination of the dissociation constant (K_D). The method is able to provide a variety of parameters such as the binding stoichiometry and thermodynamics of an interaction (Fig. 8A). To date, more and more automated systems are available which reduce sample consumption and improve cleaning procedures. A basic ITC device consists of two cells, one sample cell and one reference cell (Fig. 8B). These cells are kept at a pre-defined temperature by the system. The titrant is located in a syringe that injects fixed volumes to the sample cell in regular intervals. If the titrant interacts with the contents of the sample cell, the temperature in the sample cell changes. Heat is released if the reaction is exothermic, or taken up if the reaction is endothermic. Regardless, the device re-equilibrates the temperatures of both cells and records the required power (DP).

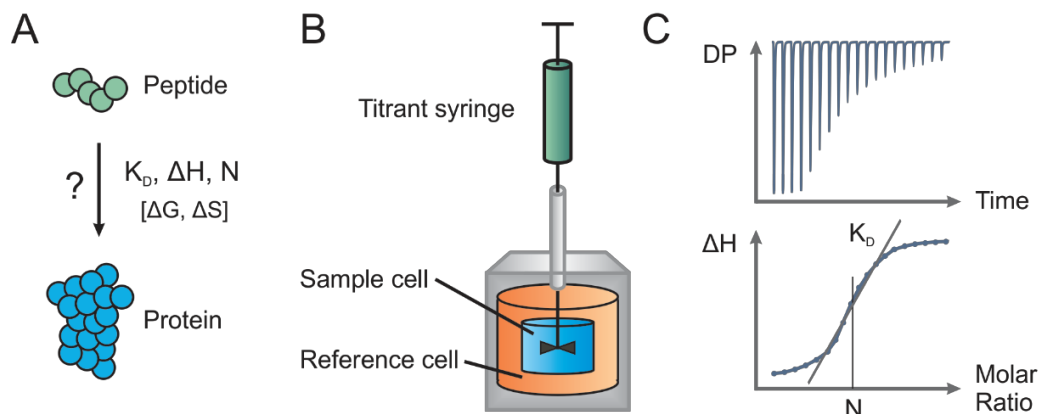


Figure 8: Experimental setup and data readout of the isothermal titration calorimetry (ITC) method.

A: Design of an ITC device for binary titration.

A basic device contains two cells. The sample cell contains a component, whose binding parameters are to be tested. The reference cells contains distilled water, which is kept at a defined temperature. Via a syringe, a specified volume of titrant is injected into the sample cell in regular intervals. If the titrant interacts with the sample in the cell, heat is released to or taken up from the surroundings according to the nature of the interaction. After each injection, the machine re-equilibrates the temperature difference between sample and reference cell and records the required power.

B: Ascertainable biophysical parameters of ITC experiments.

ITC experiments provide a variety of biophysical parameters. Different molecules such as proteins, peptides, nucleic acids and chemical compounds can be titrated to each other to study their binding behavior. Direct readouts from the experiment are the dissociation constant K_D , the binding enthalpy ΔH and the binding stoichiometry N . Thermodynamic parameters such as the entropy ΔS and the free enthalpy ΔG can be subsequently calculated.

C: Data readout of ITC experiments.

The direct readout of the experiment is the differential power DP supplied to maintain a temperature equilibrium between the reference and the sample cell over the course of the experiment (top). Via integration and normalization in respect to the concentrations, a binding isotherm is obtained (bottom). After fitting to an appropriate binding model, the binding stoichiometry N , the binding enthalpy ΔH and the dissociation constant K_D can be obtained from the fit.

This power difference is plotted over the time course of the experiment (Fig. 8C, top). After integration and normalization of the curve, a binding isotherm is obtained (Fig. 8C, bottom). This curve is fitted to an appropriate binding model e.g. according to the number of binding sites, which reveals the dissociation constant K_D , the binding stoichiometry N and the binding enthalpy ΔH of the two titrated components. If the appropriate binding model is not obvious from the curve shape, orthogonal methods such as static light scattering (SLS) of the studied complex might provide the missing information. From the determined parameters, thermodynamic values such as the entropy ΔS and the free enthalpy ΔG can be calculated. Isothermal titration calorimetry is therefore a valuable technique that allows comprehensive characterization of molecular interactions.

In this thesis, ITC experiments were mostly used to determine and compare the binding strength (via the dissociation constant K_D) of different protein-ligand setups. For this purpose, ITC was used complementary to other biophysical techniques such as NMR spectroscopy or fluorescence polarization assays, which were previously used to qualitatively study binding behavior.

Among the experiments performed for this thesis, titrations of the same protein titrant to either short peptides or long protein constructs containing the same binding consensus sequence were performed in several cases. For titrations involving peptides (with a range of 10-20 residues) it was noticed that late injections, where the component in the sample cell should already be saturated with titrant, still cause a rather high heat release. This effect has a certain influence on the determination of parameters e.g. the dissociation constant K_D . It is assumed, that this heat release is caused by unspecific binding of the unstructured peptides to surface-exposed areas of the protein, leading to an under-estimation of the determined K_D values in protein-peptide titrations.

3 Publications

The results of this thesis have been reported in two publications. The original scientific studies published in international, peer-reviewed journals are summarized below.

3.1 The Structure of the SPOP-Pdx1 Interface Reveals Insights into the Phosphorylation-Dependent Binding Regulation

The article, reference ⁵¹ can be accessed at [<https://doi.org/10.1016/j.str.2018.10.005>]. The author of this doctoral thesis, Michael Ostertag, is first author of the publication, and was majorly involved in the conception, performance and analysis of all experiments as well as the preparation of the manuscript. Citations of the article itself are omitted in the following summarizing section.

The publication focuses on the structural and biophysical characterization of the interaction of Pdx1, a vital pancreatic transcription factor^{14,15,20} and SPOP, the substrate-recognition subunit of an E3 ubiquitin-ligase complex^{8,26} involved in its proteasomal degradation pathway^{21,22}. The publication shows two novel co-crystal structures of the SPOP-Pdx1 interface, which give important insights into how binding of the two proteins is facilitated on a molecular level. This is especially of interest, since the SPOP-binding sequence of Pdx1 differs from a previously defined SPOP-binding consensus²⁶. The SPOP-Pdx1 co-crystal structures show an extended interface compared to published SPOP-ligand co-crystal structures and show that SPOP can accommodate ligands with lower amino acid conservancy. The study further contains biophysical data on the SPOP-Pdx1 interface. Different constructs of the SPOP and Pdx1 proteins were used in the determination of thermodynamic parameters such as the dissociation constant (K_D) or binding stoichiometry using ITC experiments. NMR spectroscopy was used to study dynamics of Pdx1 in solution, determining that the SPOP-binding site in Pdx1 is a largely unfolded motif which is not influenced by the adjacent, highly structured DNA-binding domain of Pdx1. The study provides further insights into the regulation of the SPOP-Pdx1 interaction by post-translational modification. Pdx1 was previously shown to be phosphorylated at residues Thr230 and Ser231 by kinase CK2^{24,25}. These residues are part of the SPOP-binding sequence of Pdx1. The binding capabilities of phosphorylated Pdx1 to SPOP were assessed with ITC and NMR experiments. Both single and double phosphorylation of Pdx1 was shown to clearly disrupt SPOP binding, a finding that is backed by the presented structural data, yet opposes previously published studies²⁵. Still, phosphorylation poses an important regulation mechanism on SPOP-Pdx1 binding.

3.2 Structural Insights into BET Client Recognition of Endometrial and Prostate Cancer-Associated SPOP Mutants

The article, reference ⁵² can be accessed at [<https://doi.org/10.1016/j.jmb.2019.04.017>]. The author of this doctoral thesis, Michael Ostertag, and Wiebke Hutwelker contributed equally to the publication. M.O. was majorly involved in the conception, performance and analysis of NMR, ITC and crystallographic experiments as well as the preparation of the manuscript. Citations of the article itself are omitted in the following summarizing section.

The publication contains novel co-crystal structures of the SPOP-BRD3 interface, representative for other BET proteins. They are a family of potentially oncogenic transcription regulators and were recently found to be SPOP clients, which is the substrate-recognition subunit of an E3 ubiquitin-ligase complex^{8,26}. The co-crystal structure gives insights into molecular interactions formed during SPOP-BET protein binding. The publication aims to study a previously reported oncogenomics paradox, where mutations in the same SPOP domain cause opposing BET inhibitor drug susceptibility in endometrial and prostate cancer, presumably based on altered BET binding of mutated SPOP²⁷⁻²⁹. To elucidate their binding behavior, a total of 14 prostate and endometrial cancer-associated SPOP mutants were expressed, and tested in biophysical assays. Prostate cancer-associated SPOP mutation sites are located in the ligand binding groove of the SPOP MATH domain. A series of NMR, ITC and FP experiments confirmed that these mutations critically impair BET protein binding. This is rationalized by our co-crystal structure, which shows that prostate cancer-associated mutation sites are located in SPOP residues involved in key ligand interactions.

As opposed, endometrial cancer-associated SPOP mutation sites are found in regions of the MATH domain outside the ligand binding groove. The BET protein binding behavior of such mutants was assessed with NMR, ITC and FP experiments. When a small fragment of BRD3 comprising the core SPOP-binding sequence was used, all tested endometrial cancer-associated SPOP mutants showed binding behavior highly similar to the wild-type protein. Three co-crystal structures of different endometrial cancer-associated SPOP mutants and BRD3 confirm that the interface remains unchanged. When a larger BRD3 construct (43 kDa) was used as binding partner, all but one mutant showed binding behavior identical to wild-type SPOP. For SPOP M117V, a small increase in BRD3 affinity was observed. However, based on our data, the disease phenotypes observed in endometrial cancer *in vivo*²⁹ are more likely caused by higher-level regulatory mechanisms, than alterations in the core SPOP-BET protein interface.

4 Conclusions

The SPOP protein is a key component of the ubiquitin-proteasome system^{8,26,53}, which is a critical cellular recycling mechanism responsible for protein catabolism². For normal function, tight control over the UPS is required, and failures in its regulation were shown to have pathological effects⁵⁴. SPOP is the substrate-recognition subunit of the cullin-3-RING ligase complex⁴ and binds substrates with its MATH domain²⁶. A variety of proteins were reported to be SPOP substrates such as the pancreatic transcription factor Pdx1^{21,55} or members of the BET family of transcriptional co-activators such as BRD3²⁷⁻²⁹. SPOP serves as adapter protein which enables the binding and subsequent ubiquitination of many different client proteins by the same E2 ubiquitin-conjugating enzyme⁴.

In this study, the interaction of SPOP with the two above mentioned ligands, Pdx1 and BRD3 was studied on a structural and biophysical level. A number of co-crystal structures of SPOP in complex with either ligand were solved. The SPOP-Pdx1 co-crystal structures can be accessed at the Protein Data Bank (PDB) with the accession numbers 6F8F and 6F8G. SPOP-BRD3 co-crystal structures can be found in the PDB with the accession codes 6I41, 6I5P, 6I68 and 6I7A. A comparison of these co-crystal structures reveals that SPOP binds both ligands at the same position, a flat binding groove in the center of the MATH domain. This concurs with previous literature describing binding positions of several SPOP ligands, where a general SPOP-binding consensus of nonpolar-polar-Ser-Ser/Thr-Ser/Thr (found e.g. as VTSTT in Puc phosphatase) was defined based on the analysis of four different SPOP client proteins²⁶. Interestingly, neither the SPOP binding sites of BRD3 nor Pdx1 conform to this SBC. While BRD3 and other proteins of the BET family contain residues with highly similar properties (ADTTT), the SBC in Pdx1 (VTSGE) deviates significantly in the sterical and physical properties. Still, SPOP is able to accommodate this ligand, based on extended contacts distal from the core binding site⁵¹. Serving as an adapter protein, it is sensible that SPOP is able to recognize and bind to a broad variety of ligands.

This study also includes biophysical data on SPOP interaction of the two studied ligands Pdx1 and BRD3, such as the determination of dissociation constants and binding stoichiometry. Previous reports indicate that a number of SPOP clients contain multiple binding sites to increase affinity and ultimately the ubiquitination turnover²⁶. The obtained data on full-length Pdx1 indicate that this protein contains only one SPOP binding site with an affinity of $\sim 60 \mu\text{M}$ ⁵¹. Similar affinities were obtained a large BRD3 construct containing both bromodomains (SPOP affinity $\sim 60 \mu\text{M}$). Affinities of different SBCs determined in previous studies range from ~ 4 - $260 \mu\text{M}$, depending on the primary sequence. Proteins that contain only one SBC site such as MacroH2A were reported

to have highly similar K_{Ds} ($\sim 60 \mu\text{M}$)²⁶ to the ligands studied here. This suggests that SPOP binding of Pdx1 and BRD3 follows similar properties as other well-studied SPOP clients.

Apart from the affinity, differential regulation mechanisms of SPOP-client binding are likely, and have been shown e.g. for Pdx1. Several studies indicated that Pdx1 is a phosphoprotein, and is post-translationally modified at several positions²³⁻²⁵. In this thesis, it was shown that the phosphorylation of Pdx1 at residues Thr230 and Ser231 in the SBC site drastically reduces SPOP binding⁵¹. This strongly suggests that phosphorylation serves as a direct regulation mechanism of SPOP binding for Pdx1. As reduced SPOP binding presumably reduces Pdx1 ubiquitination and proteasomal degradation, it is conceivable that this serves as a rescue mechanism against declining Pdx1 levels in the pancreatic β -cells, which causes their dysfunction and death. The SPOP-Pdx1 interaction could therefore be exploited as therapeutic target in the combat against T2D.

Several studies suggest that SPOP is involved in cancer pathology^{11,12,56,57}. On a mechanistic level, both mutation and mislocalization have been suggested in literature. In this thesis, SPOP point mutations were studied regarding their BET protein binding behavior. Concomitant with previous literature²⁷⁻²⁹, it was found that prostate cancer-associated SPOP mutants show significantly reduced affinity or are unable to bind the BRD3 ligand at all. These findings are rationalized by the obtained co-crystal structures of the interface, which show that the prostate cancer-associated SPOP mutation sites are found in residues forming direct contacts with the ligand. Mutations in the ligand binding groove of SPOP are very likely to affect binding of all SPOP clients in some manner, as they rely on similar contacts and interacting residues^{26,51}.

In endometrial cancer, other SPOP mutations were found, located at positions distal from the ligand binding groove²⁹. As SPOP mutations were repeatedly observed in endometrial cancers^{11,56,58}, it is likely that they are linked to the pathology. A recent study suggested that the SPOP-BET protein interaction was etiological in some endometrial cancers²⁹. In this thesis, different endometrial cancer-associated SPOP mutants were studied. Here, no difference in ligand binding affinity was observed in the majority of cases, and co-crystal structures show that the core binding interface of the mutated SPOP MATH domain and the BRD3 SBC remains unaltered⁵². As BET protein binding seems to be unaffected by the SPOP mutations in endometrial cancer, it is conceivable that they affect other properties of the protein, leading to the observed *in vivo* effects²⁹.

Generally, SPOP is considered a tumor suppressor protein because it targets oncogenes for proteasomal degradation e.g. in breast cancer⁵⁹. In other tumors such as kidney cancer, SPOP seems to promote tumorigenesis, because it is mislocalized from the nucleus to the cytosol¹².

Recently, it was shown that SPOP mutations disrupt liquid-liquid phase separation of the protein, which correlates to loss of function¹⁰. As the data presented in this thesis do not seem to show an alteration of the BRD3 binding interface of endometrial cancer-associated SPOP mutants, it is conceivable that other regulation mechanisms such as mislocalization or an altered protein expression level are causative here.

As the SPOP mutation landscape of a tumor seems to have direct impact on the susceptibility towards certain anticancer drugs²⁷⁻²⁹, screening for SPOP mutations may be a valuable personalized medicine approach in anticancer therapy.

In summary, this study elucidates the biological background of the SPOP-mediated degradation of Pdx1 and BRD3, which are ligands whose turnover directly affects pathogenic processes^{29,60}. The characterization of the binding interfaces of the complexes enables structure-based drug discovery approaches for the development of ligand molecules to combat type 2 diabetes as well as prostate and endometrial cancer. Furthermore, the data suggests a phosphorylation-driven rescue mechanism for Pdx1 levels⁵¹, which could be exploited as a therapeutic approach in the treatment or even prevention of diabetes if ways to manipulate Pdx1 phosphorylation *in vivo* are revealed in future studies.

I References

- 1 Mani, R. S. The emerging role of speckle-type POZ protein (SPOP) in cancer development. *Drug Discov Today* **19**, 1498-1502, doi:10.1016/j.drudis.2014.07.009 (2014).
- 2 Nandi, D., Tahiliani, P., Kumar, A. & Chandu, D. The ubiquitin-proteasome system. *J Biosci* **31**, 137-155 (2006).
- 3 Semple, C. A., Group, R. G. & Members, G. S. L. The comparative proteomics of ubiquitination in mouse. *Genome Res* **13**, 1389-1394, doi:10.1101/gr.980303 (2003).
- 4 Cheng, J. *et al.* Functional analysis of Cullin 3 E3 ligases in tumorigenesis. *Biochim Biophys Acta Rev Cancer* **1869**, 11-28, doi:10.1016/j.bbcan.2017.11.001 (2018).
- 5 Pintard, L., Willems, A. & Peter, M. Cullin-based ubiquitin ligases: Cul3-BTB complexes join the family. *EMBO J* **23**, 1681-1687, doi:10.1038/sj.emboj.7600186 (2004).
- 6 The PyMOL Molecular Graphics System (Schrödinger LLC., Version 1.8.6.0).
- 7 Furukawa, M., He, Y. J., Borchers, C. & Xiong, Y. Targeting of protein ubiquitination by BTB-Cullin 3-Roc1 ubiquitin ligases. *Nat Cell Biol* **5**, 1001-1007, doi:10.1038/ncb1056 (2003).
- 8 van Geersdaele, L. K. *et al.* Structural basis of high-order oligomerization of the cullin-3 adaptor SPOP. *Acta Crystallogr D Biol Crystallogr* **69**, 1677-1684, doi:10.1107/S0907444913012687 (2013).
- 9 Marzahn, M. R. *et al.* Higher-order oligomerization promotes localization of SPOP to liquid nuclear speckles. *EMBO J* **35**, 1254-1275, doi:10.15252/embj.201593169 (2016).
- 10 Bouchard, J. J. *et al.* Cancer Mutations of the Tumor Suppressor SPOP Disrupt the Formation of Active, Phase-Separated Compartments. *Mol Cell* **72**, 19-36 e18, doi:10.1016/j.molcel.2018.08.027 (2018).
- 11 Wei, X. *et al.* Functional roles of Speckle-Type Poz (SPOP) Protein in Genomic stability. *J Cancer* **9**, 3257-3262, doi:10.7150/jca.25930 (2018).
- 12 Li, G. *et al.* SPOP promotes tumorigenesis by acting as a key regulatory hub in kidney cancer. *Cancer Cell* **25**, 455-468, doi:10.1016/j.ccr.2014.02.007 (2014).
- 13 Jonsson, J., Carlsson, L., Edlund, T. & Edlund, H. Insulin-promoter-factor 1 is required for pancreas development in mice. *Nature* **371**, 606-609, doi:10.1038/371606a0 (1994).
- 14 Ashizawa, S., Brunnicardi, F. C. & Wang, X. P. PDX-1 and the pancreas. *Pancreas* **28**, 109-120 (2004).
- 15 Fujimoto, K. & Polonsky, K. S. Pdx1 and other factors that regulate pancreatic beta-cell survival. *Diabetes Obes Metab* **11 Suppl 4**, 30-37, doi:10.1111/j.1463-1326.2009.01121.x (2009).
- 16 Kaneto, H. & Matsuoka, T. A. Role of pancreatic transcription factors in maintenance of mature beta-cell function. *Int J Mol Sci* **16**, 6281-6297, doi:10.3390/ijms16036281 (2015).

- 17 Cnop, M. *et al.* Mechanisms of pancreatic beta-cell death in type 1 and type 2 diabetes: many differences, few similarities. *Diabetes* **54 Suppl 2**, S97-107 (2005).
- 18 Robertson, R. P. Chronic oxidative stress as a central mechanism for glucose toxicity in pancreatic islet beta cells in diabetes. *J Biol Chem* **279**, 42351-42354, doi:10.1074/jbc.R400019200 (2004).
- 19 Guo, S. *et al.* Inactivation of specific beta cell transcription factors in type 2 diabetes. *J Clin Invest* **123**, 3305-3316, doi:10.1172/JCI65390 (2013).
- 20 Yamamoto, Y. *et al.* Preserving expression of Pdx1 improves beta-cell failure in diabetic mice. *Biochem Biophys Res Commun* **483**, 418-424, doi:10.1016/j.bbrc.2016.12.128 (2017).
- 21 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* **24**, 4372-4383 (2004).
- 22 Claiborn, K. C. *et al.* Pcif1 modulates Pdx1 protein stability and pancreatic beta cell function and survival in mice. *J Clin Invest* **120**, 3713-3721, doi:10.1172/JCI40440 (2010).
- 23 Ardestani, A. *et al.* MST1 is a key regulator of beta cell apoptosis and dysfunction in diabetes. *Nat Med* **20**, 385-397, doi:10.1038/nm.3482 (2014).
- 24 Meng, R. *et al.* CK2 phosphorylation of Pdx-1 regulates its transcription factor activity. *Cell Mol Life Sci* **67**, 2481-2489, doi:10.1007/s00018-010-0348-0 (2010).
- 25 Klein, S., Meng, R., Montenarh, M. & Gotz, C. The Phosphorylation of PDX-1 by Protein Kinase CK2 Is Crucial for Its Stability. *Pharmaceuticals (Basel)* **10**, doi:10.3390/ph10010002 (2016).
- 26 Zhuang, M. *et al.* Structures of SPOP-substrate complexes: insights into molecular architectures of BTB-Cul3 ubiquitin ligases. *Mol Cell* **36**, 39-50, doi:10.1016/j.molcel.2009.09.022 (2009).
- 27 Zhang, P. *et al.* Intrinsic BET inhibitor resistance in SPOP-mutated prostate cancer is mediated by BET protein stabilization and AKT-mTORC1 activation. *Nat Med* **23**, 1055-1062, doi:10.1038/nm.4379 (2017).
- 28 Dai, X. *et al.* Prostate cancer-associated SPOP mutations confer resistance to BET inhibitors through stabilization of BRD4. *Nat Med* **23**, 1063-1071, doi:10.1038/nm.4378 (2017).
- 29 Janouskova, H. *et al.* Opposing effects of cancer-type-specific SPOP mutants on BET protein degradation and sensitivity to BET inhibitors. *Nat Med* **23**, 1046-1054, doi:10.1038/nm.4372 (2017).
- 30 Shi, J. & Vakoc, C. R. The mechanisms behind the therapeutic activity of BET bromodomain inhibition. *Mol Cell* **54**, 728-736, doi:10.1016/j.molcel.2014.05.016 (2014).
- 31 Dhalluin, C. *et al.* Structure and ligand of a histone acetyltransferase bromodomain. *Nature* **399**, 491-496, doi:10.1038/20974 (1999).
- 32 Rahman, S. *et al.* The Brd4 extraterminal domain confers transcription activation independent of pTEFb by recruiting multiple proteins, including NSD3. *Mol Cell Biol* **31**, 2641-2652, doi:10.1128/MCB.01341-10 (2011).
- 33 French, C. A. *et al.* BRD4-NUT fusion oncogene: a novel mechanism in aggressive carcinoma. *Cancer Res* **63**, 304-307 (2003).

- 34 Delmore, J. E. *et al.* BET bromodomain inhibition as a therapeutic strategy to target c-Myc. *Cell* **146**, 904-917, doi:10.1016/j.cell.2011.08.017 (2011).
- 35 Filippakopoulos, P. *et al.* Selective inhibition of BET bromodomains. *Nature* **468**, 1067-1073, doi:10.1038/nature09504 (2010).
- 36 Perez-Salvia, M. & Esteller, M. Bromodomain inhibitors and cancer therapy: From structures to applications. *Epigenetics* **12**, 323-339, doi:10.1080/15592294.2016.1265710 (2017).
- 37 Rathert, P. *et al.* Transcriptional plasticity promotes primary and acquired resistance to BET inhibition. *Nature* **525**, 543-547, doi:10.1038/nature14898 (2015).
- 38 Studier, F. W. Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* **41**, 207-234 (2005).
- 39 McPherson, A. & Gavira, J. A. Introduction to protein crystallization. *Acta Crystallogr F Struct Biol Commun* **70**, 2-20, doi:10.1107/S2053230X13033141 (2014).
- 40 Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501, doi:10.1107/S0907444910007493 (2010).
- 41 RCSB PDB-101: Educational Portal of PDB: R-value and R-free, <<https://pdb101.rcsb.org/learn/guide-to-understanding-pdb-data/r-value-and-r-free>> (2019).
- 42 Cavanagh, J. *Protein NMR Spectroscopy: Principles and Practice*. 2nd edn, (Elsevier Science & Technology, 2007).
- 43 Lambert, J. B. *Nuclear Magnetic Resonance Spectroscopy: An Introduction to Principles, Applications and Experimental Methods*. (John Wiley & Sons, Inc., 2019).
- 44 Grzesiek, S. & Bax, A. An efficient experiment for sequential backbone assignment of medium-sized isotopically enriched proteins. *J. Magn. Reson.* **99**, 201-207, doi:10.1016/0022-2364(92)90169-8 (1992).
- 45 Grzesiek, S. & Bax, A. Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR. *J Am Chem Soc* **114**, 6291-6293, doi:10.1021/ja00042a003 (1992).
- 46 Kasprzak, A. A. & Villafranca, J. J. Interactive binding between the substrate and allosteric sites of carbamoyl-phosphate synthetase. *Biochemistry* **27**, 8050-8056 (1988).
- 47 Reich, S., Guilligay, D. & Cusack, S. An in vitro fluorescence based study of initiation of RNA synthesis by influenza B polymerase. *Nucleic Acids Res* **45**, 3353-3368, doi:10.1093/nar/gkx043 (2017).
- 48 Takahashi, M., Sakumi, K. & Sekiguchi, M. Interaction of Ada protein with DNA examined by fluorescence anisotropy of the protein. *Biochemistry* **29**, 3431-3436 (1990).
- 49 Owicki, J. C. Fluorescence polarization and anisotropy in high throughput screening: perspectives and primer. *J Biomol Screen* **5**, 297-306, doi:10.1177/108705710000500501 (2000).

- 50 Lea, W. A. & Simeonov, A. Fluorescence polarization assays in small molecule screening. *Expert Opin Drug Discov* **6**, 17-32, doi:10.1517/17460441.2011.537322 (2011).
- 51 Ostertag, M. S., Messias, A. C., Sattler, M. & Popowicz, G. M. The Structure of the SPOP-Pdx1 Interface Reveals Insights into the Phosphorylation-Dependent Binding Regulation. *Structure*, doi:10.1016/j.str.2018.10.005 (2018).
- 52 Ostertag, M. S., Hutwelker, W., Plettenburg, O., Sattler, M. & Popowicz, G. M. Structural Insights into BET Client Recognition of Endometrial and Prostate Cancer-Associated SPOP Mutants. *J Mol Biol*, doi:10.1016/j.jmb.2019.04.017 (2019).
- 53 Errington, W. J. *et al.* Adaptor protein self-assembly drives the control of a cullin-RING ubiquitin ligase. *Structure* **20**, 1141-1153, doi:10.1016/j.str.2012.04.009 (2012).
- 54 Senft, D., Qi, J. & Ronai, Z. A. Ubiquitin ligases in oncogenic transformation and cancer therapy. *Nat Rev Cancer* **18**, 69-88, doi:10.1038/nrc.2017.105 (2018).
- 55 Liu, A., Oliver-Krasinski, J. & Stoffers, D. A. Two conserved domains in PCIF1 mediate interaction with pancreatic transcription factor PDX-1. *FEBS Lett* **580**, 6701-6706, doi:10.1016/j.febslet.2006.11.021 (2006).
- 56 Zhang, P. *et al.* Endometrial cancer-associated mutants of SPOP are defective in regulating estrogen receptor-alpha protein turnover. *Cell Death Dis* **6**, e1687, doi:10.1038/cddis.2015.47 (2015).
- 57 Barbieri, C. E. *et al.* Exome sequencing identifies recurrent SPOP, FOXA1 and MED12 mutations in prostate cancer. *Nat Genet* **44**, 685-689, doi:10.1038/ng.2279 (2012).
- 58 Le Gallo, M. *et al.* Exome sequencing of serous endometrial tumors identifies recurrent somatic mutations in chromatin-remodeling and ubiquitin ligase complex genes. *Nat Genet* **44**, 1310-1315, doi:10.1038/ng.2455 (2012).
- 59 Li, C. *et al.* Tumor-suppressor role for the SPOP ubiquitin ligase in signal-dependent proteolysis of the oncogenic co-activator SRC-3/AIB1. *Oncogene* **30**, 4350-4364, doi:10.1038/onc.2011.151 (2011).
- 60 Guo, S. *et al.* Inactivation of specific beta cell transcription factors in type 2 diabetes. *J Clin Invest* **123**, 3305-3316, doi:10.1172/JCI65390 (2013).

II Abbreviations

AMP	adenosyl monophosphate
ATP	adenosyl triphosphate
B_0	(external) magnetic field
BET	bromodomain and extraterminal domain
BTB	Broad-Complex, Tramtrack and Bric a brac
CK2	casein kinase II
CPS	carbamoyl-phosphate synthetase
CRL	cullin-3-RING ligase complex
CTD	C-terminal domain
DNA	deoxyribonucleic acid
ΔE	Energy difference
EMBL	European Molecular Biology Laboratory
FA	fluorescence anisotropy
FP	fluorescence polarization
γ	gyromagnetic ratio
ΔG	Gibbs free energy
GLUT2	glucose transporter 2
ΔH	binding enthalpy
H-bond	hydrogen bond
hetNOE	^1H - ^{15}N heteronuclear NOE
HMQC	heteronuclear correlation through multiple quantum coherence
HSQC	heteronuclear single quantum coherence
I	nuclear spin quantum number
IAPP	islet amyloid polypeptide
IC_{50}	half maximal inhibitory concentration
IMAC	immobilized metal affinity chromatography
IPTG	isopropyl β -D-1-thiogalactopyranoside
ITC	isothermal titration calorimetry
K_D	dissociation constant
μ	magnetic moment
MATH	meprin and TRAF homology

MR	molecular replacement
MST1	mammalian sterile 20-like kinase-1
<i>N</i>	binding stoichiometry
NLS	nuclear localization sequence
NMR	nuclear magnetic resonance
NOE	nuclear Overhauser effect
NTD	N-terminal domain
ω_0	<i>Larmor</i> frequency
OD_{600}	optical density at a wavelength of 600nm
PEI	polyethyleneimine
PP	pyrophosphate
RNA	ribonucleic acid
ΔS	entropy
SBC	SPOP-binding consensus
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SPOP	speckle-type POZ protein
SLS	static light scattering
SUMO	Small ubiquitin-related modifier 5
T1D	type 1 diabetes mellitus
T2D	type 2 diabetes mellitus
TEV	tobacco etch virus
Ub	ubiquitin
UPS	ubiquitin proteasome system

III Acknowledgements

First and foremost I would like to thank my doctorate supervisor Prof. Michael Sattler for the great opportunity to work in his institute and for being part of a group of excellent scientists.

Special thanks goes to Dr. Grzegorz Popowicz, with whom I worked together very closely. I have learned a lot from you, and I appreciate your care both on the scientific and personal side.

I would like to thank all my past and present colleagues – especially Charlotte Softley, Roberto Fino and Dr. Miriam Sonntag – both for help in the laboratory and the friendly conversations.

Also, I would like to acknowledge my co-authors Dr. Ana Messias, Wiebke Hutwelker and Prof. Oliver Plettenburg who helped realizing the publications that this thesis is based on.

Further thanks to all collaboration partners that contributed to my projects, especially my thesis committee member Dr. Kerstin Stemmer as well as Dr. Alexandra Harger, Dr. Robert Janowski and Dr. Kenji Schorpp.

Special mention goes to Astrid Lauxen and Dr. Arie Geerlof. Without your diligent work and constant effort to maintain a nice lab environment, my work would not have been as pleasant.

Last but not least, a big thank you to all my friends and family who accompanied me on my path and to my wife Katharina – for everything.



IV Appendix

	Type	Details
Appendix 1	Publication	Ostertag et al., The Structure of the SPOP-Pdx1 Interface Reveals Insights into the Phosphorylation-Dependent Binding Regulation. <i>Structure</i> 2019 https://doi.org/10.1016/j.str.2018.10.005
Appendix 2	Publication	Ostertag, Hutwelker et al., Structural Insights into BET Client Recognition of Endometrial and Prostate Cancer-Associated SPOP Mutants. <i>J. Mol. Biol.</i> 2019 https://doi.org/10.1016/j.jmb.2019.04.017