universität wien

MASTERARBEIT | MASTER'S THESIS

Titel | Title

Fragment screening of CaBdf1 BD2 for lead compound identification and theoretical analysis of NOESY peak intensities for the structure elucidation of protein ligand complexes.

verfasst von | submitted by

Colin Marc Schmoll B.Sc.

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien | Vienna, 2024

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 066 865
Studienrichtung lt. Studienblatt | Degree programme as it appears on the student record sheet:	Masterstudium Molecular Biology
Betreut von | Supervisor:	Assoz. Prof. Dipl.-Ing. Dr. Julien Orts M.Sc. PhD

Acknowledgements

First and foremost I would like to thank Ass.-Prof. Dr. Julien Orts for the opportunity to perform this thesis in his lab and his guidance throughout the project.

Further, I would like to thank Dr. Nicolas Coudevylle for his help in the lab and support in technical and theoretical matters alike.

I would also like to thank all lab members of the group for always being available for questions and general support.
Special thanks to Marie Jose Abi Saad, who oversaw most of my work. Thank you for your patience, guidance, support and advice.

List of Contents

List of Figures

List of Tables

Abbreviations

1D – One dimensional
2D – Two dimensional
3D – Three dimensional
Arg – Arginine
Asn – Asparagine
BD – Bromodomain
Bdf1 – Bromodomain containing factor 1
BET – Bromo- and Extra-Terminal domain
C. albicans – Candida albicans
CP – Cross-peak
CPMG – Carr-Purcell-Meiboom-Gill
CSP – Chemical shift perturbation
DMSO – Dimethyl sulfoxide
DP – Diagonal-peak
DSI – Diamond, SGC and iNEXT
DTT – Dithiothreitol
Etc – et cetera
FBDD – Fragment-based drug discovery
FBS – Fragment-based screening
FID – Free induction decay
Gln – Glutamine
GMP-PNP – Guanosine-5'-[β,γ-imido]triphosphate
GDP – Guanosine Diphosphate
GTP – Guanosine Triphosphate
HIV – Human Immunodeficiency Virus
HSQC – Heteronuclear single quantum coherence
HT – HisTrap
HTRF – Homogenous time-resolved fluorescence
IPTG – Isopropyl β-D-1-thiogalactopyranoside
K-RAS – Kirsten Rat Sarcoma
LW – Line width
NMR – Nuclear Magnetic Resonance
NOE – Nuclear Overhauser Effect
NMWL – Nominal molecular weight limit
PDB – Protein Data Bank
R2 – Transversal relaxation rate
RMSD – Root mean square distance
SEC – Size exclusion chromatography
STD – Saturation-transfer difference
TCEP – Tris(2-carboxyethyl)phosphine
TEV – Tobacco Etch Virus
Trp – Tryptophan
Tyr – Tyrosine

Abstract

Candida albicans, an opportunistic yeast present in 40-60% of healthy adults, can become pathogenic in immunocompromised individuals, causing candidiasis with a high mortality rate. The Bromodomain-containing factor 1 (CaBdf1) is crucial for *C. albicans* survival, playing a key role in gene regulation. As a member of the BET family, CaBdf1 contains two bromodomains with binding pockets slightly wider than their human homologs, allowing selective inhibition. The first project of this thesis utilizes Nuclear Magnetic Resonance (NMR) fragment-based screening using Saturation Transfer Difference (STD), Carr-Purcell-Meiboom-Gill (CPMG), and $^{1}H,^{15}N$ HSQC techniques to identify eight promising fragments binding to the orthosteric's site of BD2 domain of CaBdf1. A fragment-based pharmacophore was developed and validated against the structure-based pharmacophore derived from the published structure, supporting these fragments as foundations for promising binder development.

The second project is about the Nuclear Overhauser Effect (NOE), an NMR phenomenon where spatially close nuclei interact through cross-relaxation, and is crucial for studying protein-ligand interactions. A 2D-[^{1}H-^{1}H]-NOESY spectrum was simulated for K-Ras bound to 7-fluoro-N, N-dimethyl-1-benzofuran-2-carboxamide under varying conditions of mixing times, temperatures, and total protein/ligand concentrations. Simulations showed realistic behavior of diagonal- and cross-peak intensities as a function of mixing time, validating the model. Temperature-dependent simulations also align with the experimental data for a different K-ras ligand complex, showing increased diagonal-peak and decreased cross-peak intensities at higher temperatures. These results provide a robust foundation for further analysis and validation of protein-ligand interaction simulation.

Zusammenfassung

Candida albicans ist ein opportunistischer Hefepilz, der in 40-60% gesunder Erwachsenen vorkommt. Bei immungeschwächten Personen kann *C. albicans* in ein pathogenes Stadium verfallen und eine Candidose mit einer hohen Sterblichkeitsrate verursachen. Der Bromodomänen-enthaltende Faktor 1 (Bdf1) ist entscheidend für das Überleben von *C. albicans* und spielt eine Schlüsselrolle bei der Genregulation. Als Mitglied der BET-Familie enthält CaBdf1 zwei Bromodomänen, deren Bindungstaschen etwas breiter sind als bei ihren menschlichen Homologen, was eine selektive Hemmung ermöglicht. Im ersten Projekt dieser Arbeit wurde ein Fragment basiertes Screening mittels Kernspinresonanzspektroskopie (NMR) unter Verwendung von Saturation Transfer Difference (STD), Carr-Purcell-Meiboom-Gill (CPMG) und ^{1}H, ^{15}N HSQC-Techniken durchgeführt, um acht vielversprechende Fragmente zu identifizieren, die an die orthosterische Stelle der BD2 von CaBdf1 binden. Es wurde ein Liganden-basierter Pharmakophor berechnet und mit dem strukturbasierten Pharmakophor verglichen, der aus der veröffentlichten Struktur abgeleitet wurde. Dieser Vergleich zeigte alle wichtigen strukturellen Merkmale des Liganden-basierten Pharmakophores auch im strukturbasierten Pharmakophor, was diese Fragmente als Grundlage für die Entwicklung eines vielversprechenden Inhibitors unterstützt.

Das zweite Projekt befasst sich mit dem Nuclear Overhauser Effect (NOE), einem NMR-Phänomen, bei dem räumlich nahe beieinander liegende Kerne durch Kreuzrelaxation interagieren und das für die Untersuchung von Protein-Ligand-Wechselwirkungen entscheidend ist. Ein 2D-[1H-1H]-NOESY-Spektrum wurde für K-Ras, das an 7-Fluor-N,N-dimethyl-1-benzofuran-2-carboxamid gebunden ist, unter verschiedenen Bedingungen wie Mischzeiten, Temperaturen und Gesamtkonzentrationen von Protein und Ligand simuliert. Die Simulationen zeigten ein realistisches Verhalten der Diagonal- und Kreuzpeak-Intensitäten als Funktion der Mischungszeit, was das Modell bestätigt. Die temperaturabhängigen Simulationen stimmen auch mit den experimentellen Daten für einen anderen K-ras-Ligandenkomplex überein und zeigen erhöhte Diagonalpeak- und verringerte Cross-Peak-Intensitäten bei höheren Temperaturen. Diese Ergebnisse bilden eine solide Grundlage für die weitere Analyse und Validierung der Simulation von Protein-Ligand-Interaktionen.

Chapter 1 – Screening of CaBdf1 BD2

1.1 Introduction

Candida albicans (C. albicans) is a common yeast that typically resides harmless in the human microbiome, found in 40-60% of healthy adults.[1,2] However, under certain conditions, such as immune suppression or extensive antibiotic use, *C. albicans* can shift into a pathogenic state, causing candidiasis, yielding a mortality rate of 30-40% .[3,4]
This shift makes *C. albicans* a significant concern, especially in hospital environments where individuals with compromised immune systems, such as human immunodeficiency viruses (HIV) or those undergoing intensive antibacterial treatment, are more susceptible. [1,2]
Additionally, *C. albicans* is the most frequently isolated fungal species in biofilms that form on medical devices or human tissue[5,6], further complicating treatment efforts in clinical settings.

The Bromodomain-containing factor 1 (Bdf1) is a critical protein in *C. albicans* and is vital in regulating gene transcription. It is the only bromo- and extra-terminal domain (BET) family member in *C. albicans.*[7] In *Saccharomyces cerevisiae*, Bdf1 acts as a global transcriptional regulator, influencing the expression of over 500 genes and playing an essential role similar to that of C. albicans.[8] Inhibition of CaBdf1 would effectively stop the growth of *C. albicans,* making it a promising target for antifungal therapies. Interestingly, while inhibition of both bromodomains (BD1 and BD2) of the Bdf1 results in substantial growth defects, targeting only one bromodomain shows milder growth defects[9]. This suggests that dual inhibition is more effective at stopping the organism's growth.[9]

Bromodomains (BDs) are highly conserved structural modules that recognize and bind to acetylated lysine residues, playing critical roles in chromatin structure, histone recognition, and transcription.[10] Bromodomains are composed of four left-twisted amphipathic alpha helices (Z, A, B, and C)[11] (Figure 1A), which the ZA and the BC loop link to form a hydrophobic binding pocket (Figure 1B) responsible for recognizing and binding acetylated lysine residues.[10,12]

The CaBdf1 contains two bromodomains - BD1 and BD2 - crucial for its function in gene regulation.[13] Structurally, CaBdf1 BDs share 31-46% sequence identity with human BET BDs and exhibit a high affinity for peptides containing multiple acetylation marks.[9] Crystal structures of CaBdf1 BDs (PDB: 5N15, 5N13) reveal the canonical BDs fold and their human counterparts (PDB: 3MXF, 2OUO). This includes the conserved α-helices arrangement and the ZA and BC loops' acetyl-lysine binding pocket.[9]

Mietton *et al.* (2017)[9] demonstrated that CaBdf1 BDs exhibit relative insensitivity to human BET inhibitors, such as JQ1, with a 2-3-fold decrease in sensitivity. This reduced sensitivity is due to critical differences in the residues involved in the recognition of these inhibitors. In particular, 5 of the 16 key residues in BD1 and 4 in BD2 are divergent in CaBdf1, including the WPF-shelf Trp residue and the Leu residue opposite the shelf. These substitutions result in smaller side chains, which slightly widen the binding pocket in CaBdf1 compared to human BET BDs[9]. This structural difference makes it feasible to develop selective inhibitors targeting CaBdf1 BDs.

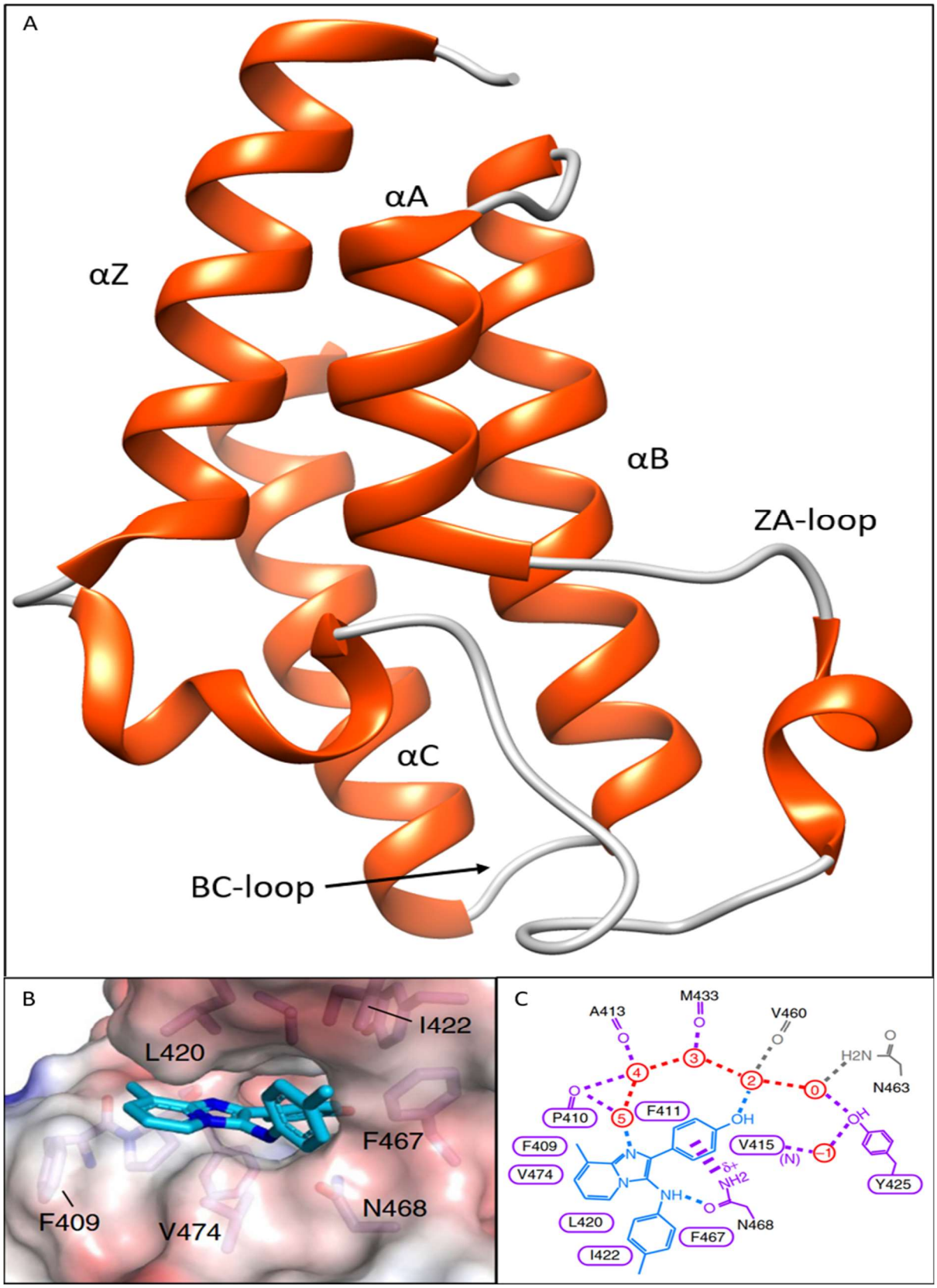

Figure 1: (A) 3D structure of the second Bromodomain BD2 from *C. albicans* in the unbound form (PDB: 5N13) with labeled helices and the loops defining the binding pocket. **(B)** Close-up surface representation of the orthosteric site of CaBdf1 BD2 bound to an imidazopyridine (PDB: 5N18). **(C)** Schematic summary of interactions. Hydrogen bonds are shown as dashed lines. Water molecules are in red. A thick dashed line shows the cation–pi type interaction between the partially charged N468 amino group and the phenol ring. Residues mediating van der Waals contacts with imidazopyridine are indicated by labels within a cartouche. **(B)** and **(C)** were copied from Mietton *et al.* (2017).[9]

Identifying selective inhibitors for CaBdf1 BDs remains a significant challenge in drug discovery. Mietton *et al.* (2017)[9] is the only study to have conducted a systematic screening for selective inhibitors of CaBdf1 BDs. They employed a homogenous time-resolved fluorescence (HTRF) inhibition assay to screen a large library of ~80.000 compounds, identifying several hundred preliminary hits for BD1 and BD2. Through a dose-response analysis, they refined these hits to 44 compounds that showed selective inhibition for CaBdf1 BD2. Among these, an imidazopyridine compound showed a particularly strong selectivity, leading to further structural characterization. The crystal structure of CaBdf1 BD2 bound to this imidazopyridine (PDB: 5N18, Figure 1B-C) revealed key residues within the binding pocket that interact with the compound directly or via water-mediated interactions, including F409, P410, F411, A413, V415, L420, I422, Y425, M433, V460, N463, F467, N468, V474 (Figure 1C).[9] These findings provided a preliminary structural framework for the design of selective BD inhibitors.

Despite these advances, the limited scope of Mietton *et al.'s* screening and the structural similarity between BD1 and BD2 suggest a need for further exploration. Starting with their findings, I seek to identify novel inhibitors, by using the state-of-the-art technique, Nuclear Magnetic Resonance (NMR) spectroscopy, addressing Mietton's screening and providing a more comprehensive foundation for the design of selective inhibitors.

1.1.1 Fragment-based-drug discovery

Fragment-based-drug-discovery (FBDD), or fragment-based-screening (FBS), is a method where libraries of small molecules, or fragments, are screened against a target protein to explore a vast chemical space for potential activators/inhibitors.[14] This approach is often used for identifying lead compounds, where initial hits are further optimized, linked to other fragments, or chemically grown to enhance binding affinity and interaction strength.[15] A positive hit, depending on the assay used, is indicated by a measurable change in a target parameter, such as inhibition of the target molecule or an increase in the relaxation rate of the fragment[15, 14]
Due to their small size, fragments typically have low binding affinity (ranging from 10^{-3} M to 10^{-6} M). FBDD combines biophysical and biochemical techniques with molecular modeling programs to identify the most promising ones, ultimately guiding the development of high-affinity lead compounds.[14]

1.1.2 Nuclear Magnetic Resonance (NMR) and its techniques

NMR spectroscopy is based on manipulating atomic nuclei with non-zero nuclear spin angular momentum (or "spin") in an external magnetic field. When an external magnetic field is applied, these nuclei align with the field and create a measurable nuclear magnetic moment. This alignment can be influenced by radio frequency (RF) radiation, causing transitions between energy states depending on the nucleus's properties and the field strength. Each NMR-active nucleus's resonance frequency depends on its unique chemical environment, which allows NMR spectroscopy to reveal structural and dynamic information about complex

molecules in solution. NMR is widely used to explore the structure, composition, and dynamics of biochemical molecules.[16]

This thesis used ligand-observed NMR techniques, such as Carr-Purcell-Meiboon-Gill (CPMG) and Saturation-Transfer-Difference (STD), for hit identification, followed by cross-validation using 2D-[^{1}H-^{15}N]-Heteronuclear Single Quantum Coherence (HSQC) NMR and pharmacophore modeling with LigandScout.

Carr-Purcell-Meiboom-Gill (CPMG) pulse sequence

The CPMG is a relaxation-edited, ligand-observed method commonly used in fragment hit screening and validation. The pulse sequence is a series of repeated 180° pulses, or "spin-echoes". A spin-echo causes the bulk magnetization to refocus and to return to its original state. Due to the transversal relaxation rate R2, the overall signal intensity will decrease after multiple repetitions of said spin-echoes.[17]
Since R2 is proportional to the rotational tumbling time τ_c and the size of the molecule, a smaller molecule relaxes way slower than a bigger molecule. Hence, the protein resonances will decay much faster than the ligand resonances. However, if a compound interacts with the protein, the R2 of the ligand will adapt to the same relaxation as the protein, causing the signal to decay faster.[18]
In screening, ^{1}H-CPMG measurements are performed at two relaxation times to compare the relaxation rate of the small molecule resonance with and without the protein. A significant signal decay in the protein-ligand mixture suggests that the ligand interacts with the protein, identifying it as a "positive hit".
However, a drawback of this 1H CPMG NMR technique is that it may struggle with weak binding interactions, as very fast exchange rates can blur the distinction between free and bound states, requiring a sufficiently large population of the bound state for reliable detection. Additionally, this method is ligand-focused and does not yield structural or dynamic information about the protein.[19]

Saturation-Transfer-Difference NMR

Saturation-Transfer-Difference (STD) NMR is another commonly used ligand-observed technique for screening and validating fragment hits. It can define the binding epitope of a ligand when bound to its receptor.[20]
Two spectra are recorded: a reference one called "off-resonance", where the irradiation is far away from any ligand or protein signal (40 ppm), and an "on-resonance", in which a selective saturation of the protein signal is irradiated (-0.5 ppm).[20,21] Only ligand signals that received magnetization transfer from the protein appear in the difference spectrum (STD_{diff}), as their intensities decrease in the on-resonance spectrum due to spin diffusion. The saturation transfer occurs due to spin diffusion from the protein, which is mediated by the Nuclear-Overhauser-Effect (NOE).[20,21] If peaks appear in the difference spectrum, this indicates that the ligand has interacted closely with the protein long enough to receive magnetization transfer, marking it as a "positive hit". As the NOE rapidly decreases with distance ($NOE \propto d^{-6}$), residues showing intensities in the differences spectrum can be used for epitope

mapping.[20] Epitope mapping will describe the protein-ligand interactions, giving a better understanding of which ligands' atoms are close to the protein's side chains.[20]

However, many limitations can occur when using the STD NMR method. If the compound is accidentally irradiated, the signal detection can be interfered with, leading to false results. Furthermore, strong binders with low dissociation constants (K_D) may be underrepresented in the STD signal due to their slow off-rates, which reduces their contributions to bulk magnetization. Like CPMG NMR, STD NMR also provides limited information about the protein structure and dynamics, focusing only on ligand interactions.[19]

Heteronuclear Single Quantum Coherence (HSQC)

The HSQC experiment is a two-dimensional NMR experiment that transfers magnetization between two different NMR-active nuclei, typically ^{1}H (proton)and ^{15}N (nitrogen), creating a spectrum with peaks corresponding to specific chemical environments.[16]
In a ^{1}H-^{15}N HSQC spectrum of a protein, also known as the "fingerprint" spectrum, each peak corresponds to an amide bond in the protein backbone, except prolines residues, as well as N-H groups in certain side chains (e.g., Asn, Gln, Trp, and Arg). This spectrum provides a distinct set of peaks that reflect the unique chemical environment of each amide group. [22,23]

In a screening context, ^{1}H, ^{15}N HSQC can detect ligand binding by comparing the spectrum of the protein alone (apo) with that of the protein-ligand complex (holo).[24] Ligand binding causes peak shifts in the HSQC spectrum, which indicate changes in the chemical environment at specific residues. This phenomenon is called chemical shift perturbations (CSPs). Overlaying the apo and holo spectra allows the identification of binding sites based on these CSPs. If the protein's peaks have been assigned to specific residues, the affected areas can reveal the precise protein regions involved in ligand interactions.

These CSPs are quantified by a weighted Euclidean distance (Equation 1.1), where δ_X represents the shift change of nucleus X and α is a weighing factor, typically 0.14 for ^{15}N, accounting for the difference in chemical shifts ranges between 1H and ^{15}N.[23]

$$d = \sqrt{\frac{1}{2} * [\delta_H^2 + (\alpha * \delta_N^2)]} \qquad (1.1)$$

However, limitations of this technique include potential CSPs due to pH shifts caused by the ligands or interactions with solvent molecules, and it does not provide direct information about the ligand itself.[19]

1.1.3 Pharmacophore Modeling

A pharmacophore describes the arrangement of electronic and steric features and their spatial orientation essential for a molecule's biological activity. Standard pharmacophore modeling features include hydrogen-bond acceptors, hydrogen-bond donors, hydrophobic

sites, electrostatic interaction sites, and pi-stacking interactions. Pharmacophore models can be derived from either structure-based or ligand-based pharmacophores. [25, 26]
Structure-based pharmacophores identify binding pocket features from one or multiple 3D protein structures.[26] Meanwhile, ligand-based pharmacophores capture common features from one or multiple fragments. This can be helpful in case the binding site is unknown.[26]
While both methods of pharmacophore calculation are useful on their own for the characterization of binding pockets or summarizing the activity of a group of ligands, the comparison of the two models allows insights into the binding mode or binding orientation of ligands sharing electronic and steric features within a binding pocket of a biological target. It can further be useful when trying to grow or fuse promising compounds.[27]

1.1.4 Objective

In this work, FBDD will be applied by screening a library of 896 fragments against the BD2 domain of the CaBdf1 protein using NMR spectroscopy techniques. The goal is to identify compounds capable of inhibiting our target protein. To this end, a series of ligand-based and protein-based NMR techniques will be employed, and a pharmacophore mapping will be generated to support the FBDD approach.

1.2 Materials and Methods

1.2.1 Fragment library

A library of 896 DSI-poised mini-fragments was purchased from Enamie[28]. A poised fragment was defined as "a fragment synthesised from a robust and general synthetic reaction such that rapid elaboration of the fragment hit into a library of analogues can be done using parallel chemistry."[28] The DSI library was diluted from 500 mM to 100 mM stock solutions and stored in DMSO-d6 at -80 °C until used for the screening.
While in use, the plates were opened under argon to minimize the oxidation of the compounds.

1.2.2 Protein production

The DNA sequence encoded BD2 of the Bdf1 protein of *C. albicans* (residues 381-491) was synthesized by GenScript (The Netherlands) with an N-terminus affinity tag and cloned in a pETM-11 expression vector. The expression protocol was adapted from Mietton *et al.* (2017)[9].

The fusion protein encoded for His-6x tag with a linker, tag-tobacco etch virus (TEV) cleavage site-BD2, was expressed in *E. coli*-RIL cells and purified.

The bacteria were grown at 37 °C in a labeled ^{15}N-M9 medium supplemented with kanamycin and chloramphenicol. When A_{600} reached ~0.8, the temperature was lowered to 16 °C, induced with 0.5 mM isopropyl β-d-1-thiogalactopyranoside (IPTG), and left overnight.

Cells were harvested in Linx6000 at 4 °C at a speed of 6000 rpm. After centrifugation, the resuspended cells were lysed through the microfluidizer (LM10) at 8000 psi.

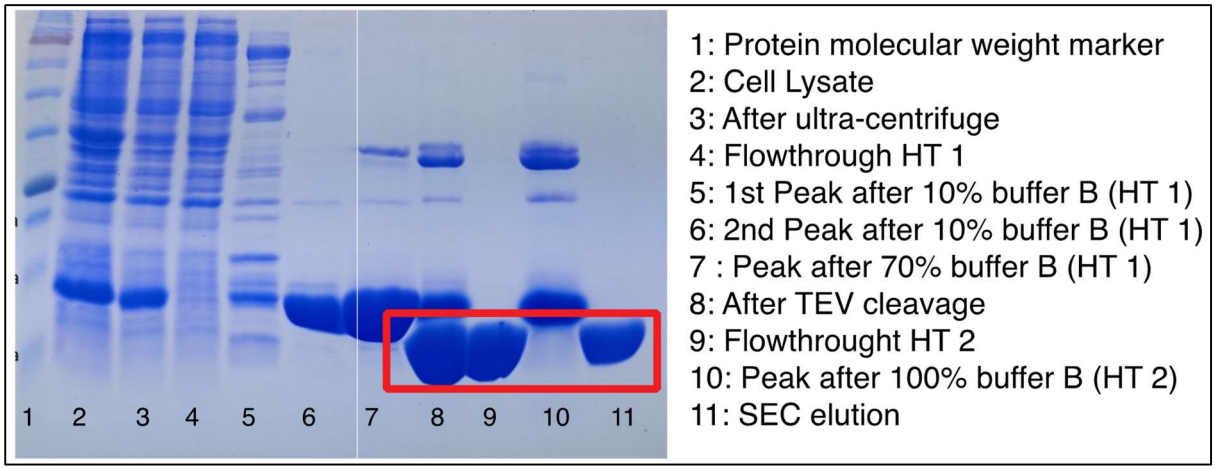

Figure 2: SDS-PAGE gel showing the purification of CaBdf1 BD2. The red box marks the bands containing the cleaved protein.

The cell lysate was passed through a HisTrap (HT) HP column (Cytivia) and eluted with 60% 50 mM Tris, 500 mM NaCl, 5 mM β-Mercaptoehtanol, and 300 mM imidazole with a pH of 7.4. Then, TEV protease, produced in-house, was added to the fractions containing this protein, and the mixture of 1:20 TEV-protein was dialyzed overnight in 50 mM Tris, 500 mM NaCl, and 1 mM DTT at 4 °C, followed by another pass-through the HT HP column. The relevant fractions

were concentrated, applied to a Sephadex G75 gel filtration column (Cytivia), and eluted with a 50 mM NaPi and 50 mM NaCl buffer. The purification protocol was adapted from Mietton *et al.* (2017)[9].

The yield was 5 mg/L, and the protein was stable at room temperature during the NMR experiments. For this work, 32.5 L of expression medium was prepared, resulting in ~162.5 mg of protein. The composition of all buffers used for the expression and purification is shown in Tables S3-9.

1.2.3 Recording of NMR data

The CaBdf1 BD2 protein was prepared in 100 mM NaPi (pH 6.5) and 50 mM NaCl in H_2O.

All NMR spectra were recorded at 298 K using a 500 MHz ASCEND equipped with an AVANCE III HD console and a 5mm Prodigy TCI-CryoProbe. Data processing was done with Bruker TopSpin 4.1.4, and HSQC analyses were performed using CCPNmr (v. 3.1.0). A screening box of 96 5 mm NMR tubes, each filled with 550 µL of sample (10% D_2O), was used for the experiments.

The screening used ligand-based NMR techniques, such as STD and CPMG, alongside protein-based ^{1}H-^{15}N HSQC NMR. The same experimental conditions were maintained across all methods.

The 1D proton spectra utilized Bruker's pulse program zggpw5.[29] The ligands' 1D proton spectra were acquired using 128 scans, 8k data points, and an acquisition time of 0.57 s, with a total experiment time of about 5 minutes.

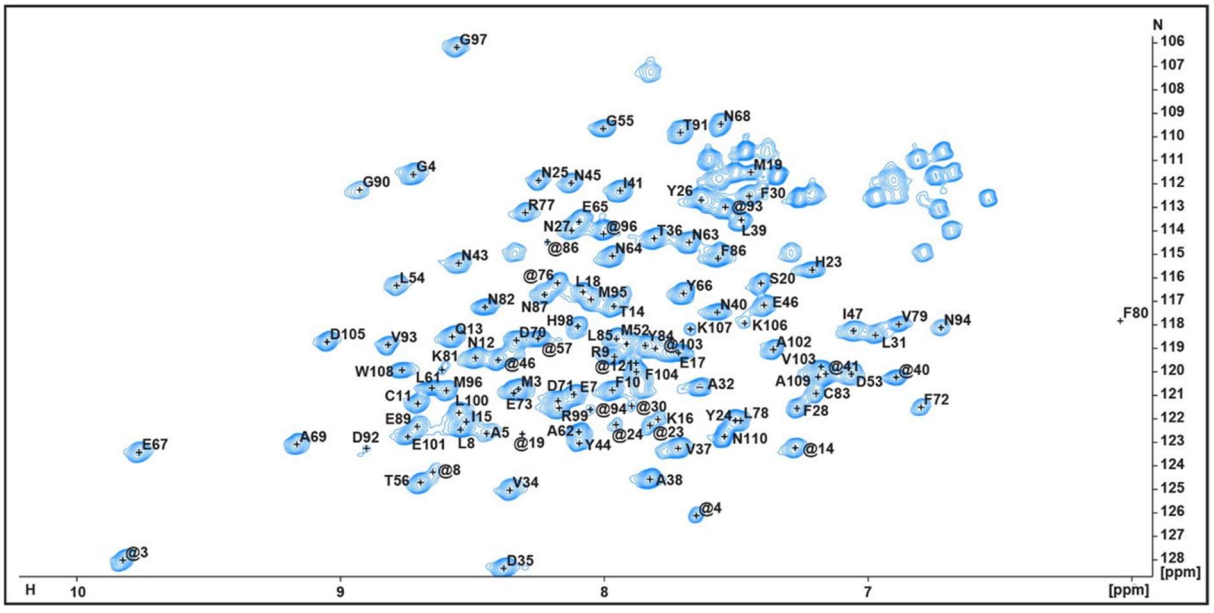

Figure 3: ^{1}H-^{15}N-HSQC spectrum of 0.6 mM ^{15}N labeled CaBdf1 BD2 in 50 mM NaPi with 10% D_2O at pH 7.4 at 303 K with carbon decoupling. It was measured at 500 MHz with Prodigy CryoProbe. Peaks labeled with a capital letter and the residue number are assigned peaks. Peaks labeled with a @ are unassigned. Copied from Avramovic (2023).[33]

The STD measurements utilized Bruer's pulse program stddiffgp19.3[20,30,29]. For the STD NMR experiments, parameters included 128 scans, 8k points, and an interscan delay of 4 s, with 0.57 s acquisition time. Protein saturation was achieved using a train of 50 ms Gaussian-shaped pulses, centered at -0.5 ppm, with a total saturation time of 4 s, and each experiment took approximately 25 minutes. Hits were identified by an STD NMR signal with a minimum signal-to-noise ratio of 2.5. The STD amplification factor determined epitope mapping.[31]

In the ^{1}H CPMG[32] experiments, relaxation delays were set at 8 and 40 ms, with an interscan relaxation delay of 2 s, 128 scans, and a 0.57 s acquisition time with 8k points.

The ^{1}H-^{15}N HSQC experiments utilized Bruker's standard pulse program hsqcgpph. Each 2D-[^{1}H, ^{15}N]-HSQC spectrum was acquired with a relaxation delay of 1 s, 16 scans, and 143 ms acquisition time, using 2k and 128 points.

All measurements were performed in the same sample tube to ensure consistent conditions. The purified CaBdf1 BD2 protein's quality check was assessed from the 2D-[^{1}H, ^{15}N]-HSQC spectrum (Figure 3).

A mixture of 2 to 3 ligands was used per tube (Tables S10-12), and the procedure was conducted in three steps (Figure 4):

1. A control experiment was performed with 100 µM of ligand mixture, and ^{1}H-NMR, STD, and CPMG spectra were recorded.
2. 4µM of protein was added to the 100 µM ligand mixture, and STD spectra were recorded.
3. An additional 96 µM of protein was introduced, and ^{1}H-NMR, CPMG, and ^{1}H-^{15}N HSQC spectra were recorded.

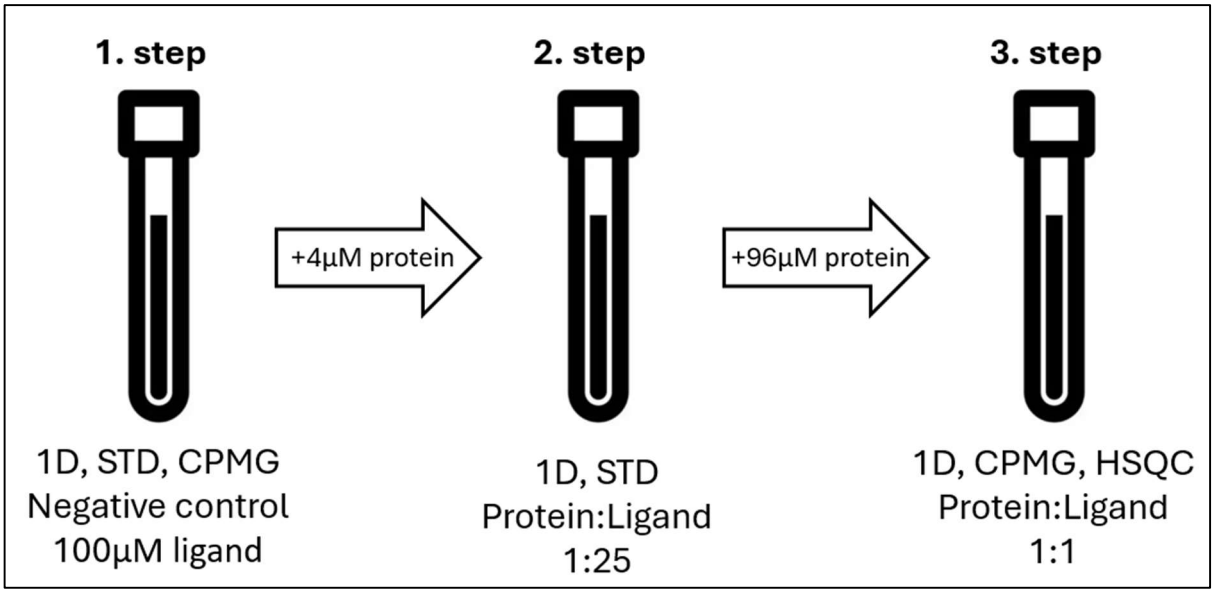

Figure 4: Schematic summarization of the screening process.

For each screening box, reference measurements were conducted under identical conditions (buffer, DMSO-d6, and 4 and 100 µM protein concentrations).

The backbone assignment for CaBdf1 BD2, used for the chemical shift perturbations (CSPs) analysis, was 85.7% complete (excluding prolines) and was previously assigned by Maja Avramovic (2023) (Figure 3).[33]

1.3 Results

A library of 896 fragments was pre-sorted based on solubility, eliminating non-soluble compounds (clogS < -2) to prevent micelle and aggregate formation, resulting in 479 fragments. The remaining fragments were then screened against the BD2 of the CaBdf1 protein.

The STD analysis identified 200 positive hits from the selected soluble fragments. However, several fragments also displayed STD signals in control measurement, indicating aggregation.[19] Fragment aggregation, likely at high concentrations, can occur even at nanomolar concentrations, leading to nonspecific protein binding and potential false positives across STD, CPMG, and CSP assays.[19,34]

For these tested fragments, an increase in STD signal intensities upon protein addition was used to confirm positive hits. Only 5 fragments showed strong STD signals without interference in the control spectrum, suggesting specific binding (Figure 5).

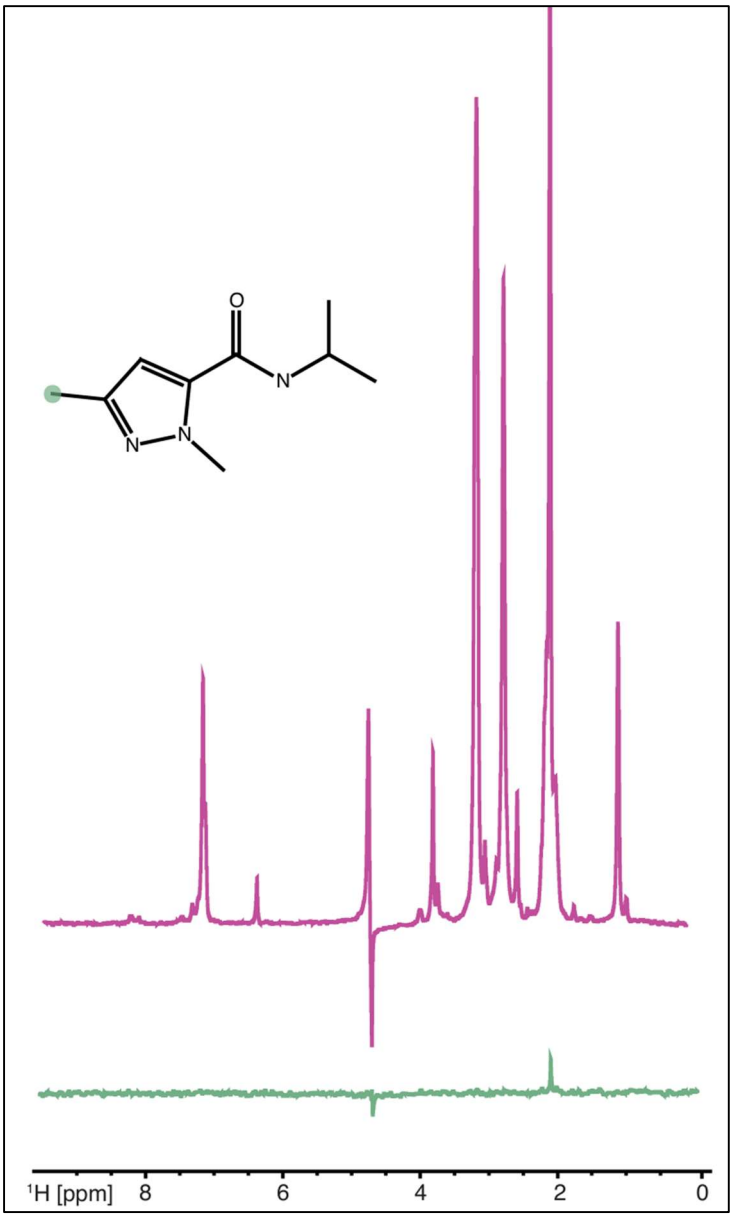

Figure 5: The ^{1}H NMR spectrum shows 100 µM of ligands 3E6 and 4E6 (**purple**) alongside the STD difference spectrum for 4 µM CaBdf1 BD2, 100 µM 3E6, and 100 µM 4E6 (**green**). All measurements were conducted in a buffer solution containing 100 mM NaPi, 50 mM NaCl, pH 7.4, and 298 K, with 10% D_2O included. The data was collected at 500 MHz with Prodigy CryoProbe. The interacting ligand, 4E6, is shown in the upper right corner, with the epitope mapping highlighted in green.

The CPMG data identified 222 positive hits out of 479 screened fragments, with 125 showing at least one resonance with an intensity decrease of 50% or more upon protein addition (Figure 6). Some fragments displayed contradictory behavior, with certain resonances decreasing in intensity while others increased, suggesting varied interaction dynamics.

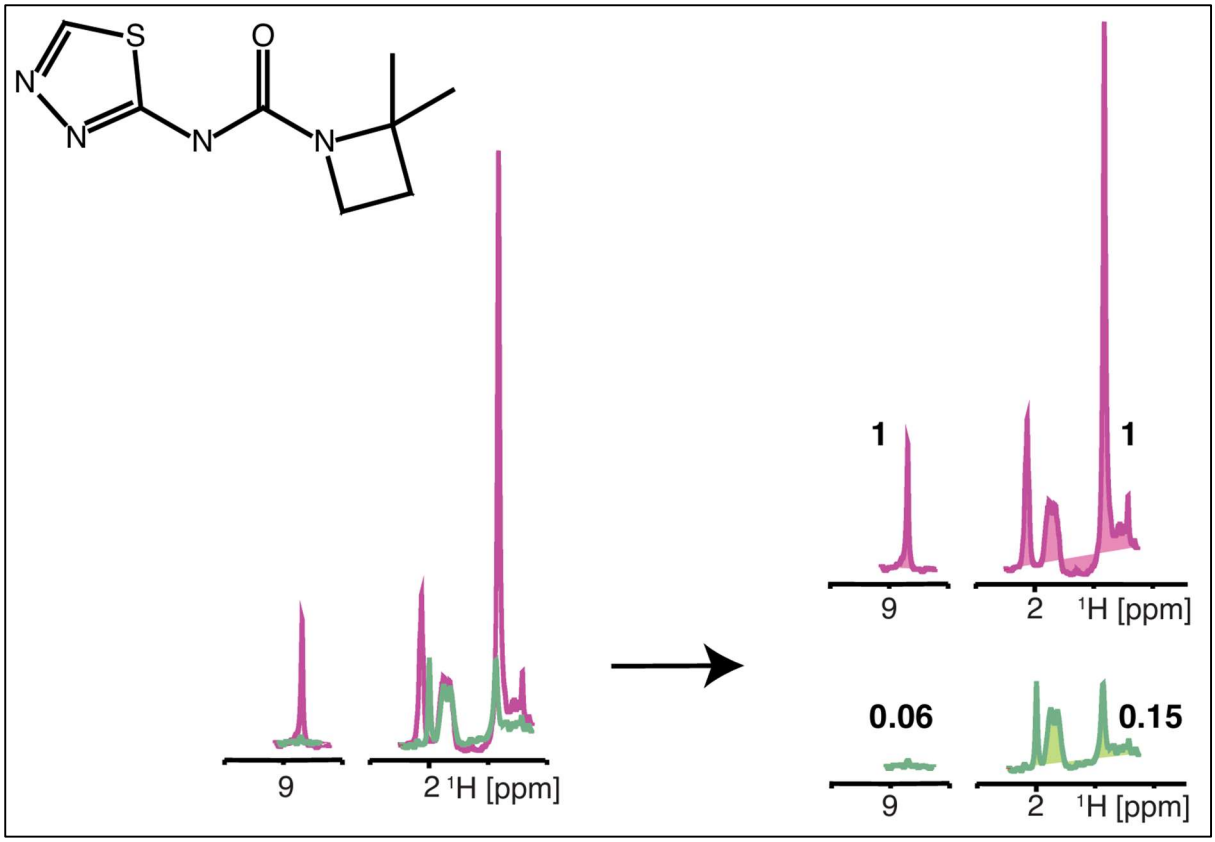

Figure 6: ^{1}H CPMG experiment showing the transverse relaxation rate for 100 µM 10E2, 11D6, and 12E2 (**purple**) and in the presence of 100 µM CaBdf1 BD2 (**green**). Both represented spectra were recorded on a 500 MHz with a 5 mm TCI-prodigy cryoprobe, with a relaxation time of 400 ms. All spectra were Measured in a buffer of 100 mM NaPi, 50 mM NaCl, pH 7.4, 298 K, and 10% D_2O. The overlaid spectra are shown on the left. The signal reduction is displayed on the right, where the signal intensity was normalized to the purple spectrum. The structure of the interacting ligand, 11D6, is shown in the upper right corner.

In both STD and CPMG analysis, 122 fragments were identified as positive hits, with 71 of these displaying a significant (>50%) intensity decrease in at least one resonance.

As an orthogonal technique for validation and to gain information about interacting residues of the protein, 2D-[^{1}H-^{15}N]-HSQC spectra were used to observe the chemical shift perturbations (CSPs) in the protein-ligand mixtures. The fragments were grouped into three categories based on CSPs: large, small, and none. Out of 479 fragments, 165 showed large CSPs (Figure 7A), 233 showed small CSPs (Figure 7B), and 81 showed no CSPs (Table S13-15).

Among the 165 fragments with large CSPs, 56 fragments showed positive STD and CPMG signals, while 34 showed only STD signals, 37 only CPMG signals, and 38 showed neither no positive signals in either STD or CPMG (Table S13).
Ligand-based screening aims to deconvolute mixtures by comparing individual ligand ^{1}H spectra with STD and CPMG signals from positive hits. However, some signals were difficult to assign due to overlapping resonances, multiple interactions per tube, and inconsistencies in reference conditions, limiting unambiguous identification.

Additional ^{1}H-^{15}N-HSQC spectra were obtained to refine these results with individual ligands for fragments showing large CSPs. This yielded 82 out of the 479 ligands with significant CSPs, equating to a hit rate of 17,05%.

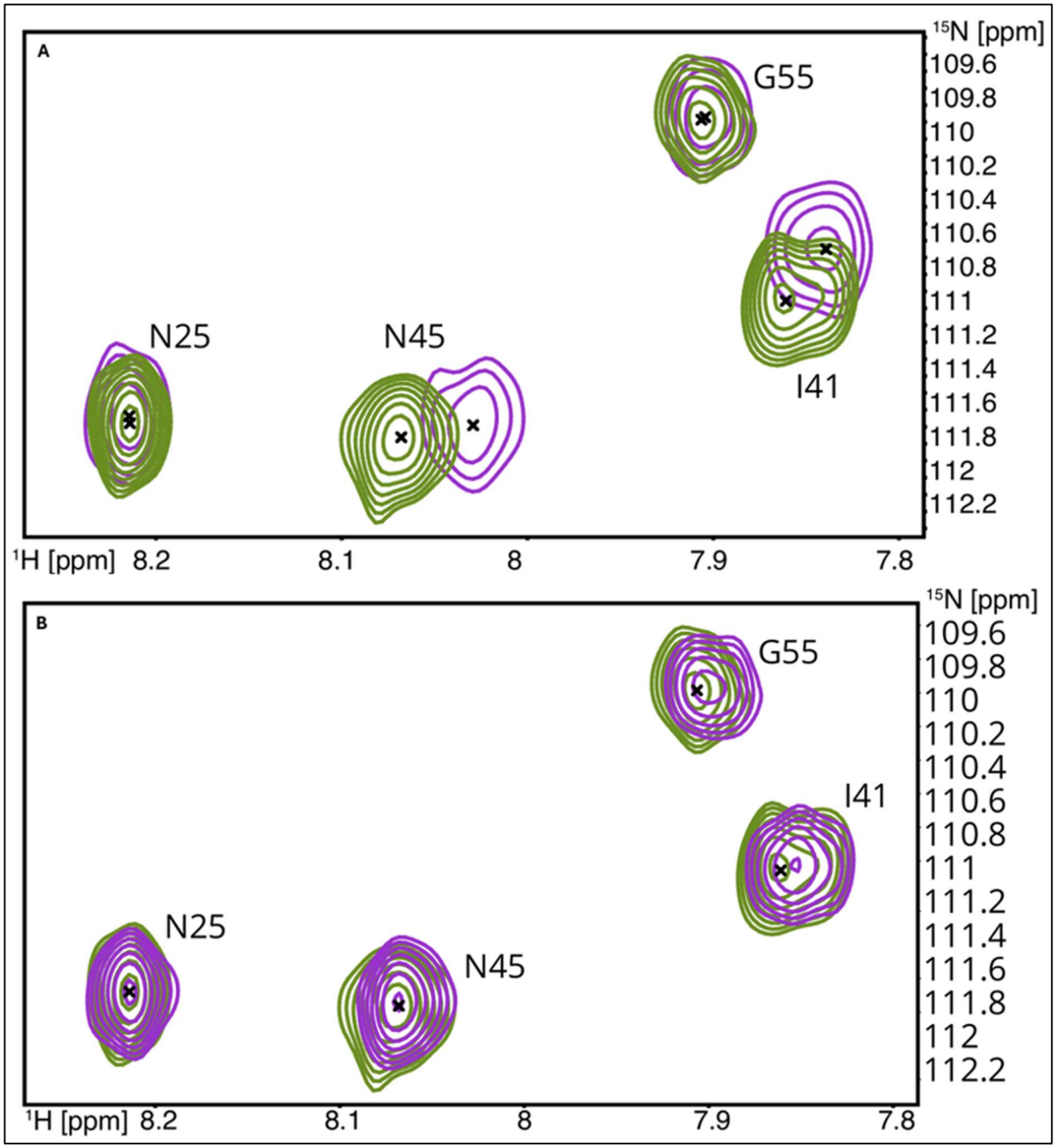

Figure 7: (A) Close-up of the backbone resonances N25, I41, N45 and G55 in the overlaid 2D ^{1}H-^{15}N HSQC spectra of 100 µM CaBdf1 BD2 +DMSO-d6 without (**green**) and with 100 µM 10E2, 11D6 and 12E2 (**purple**). Example for fragments showing large CSPs (see N45 and I41). **(B)** Close-up of the backbone resonances N25, I41, N45 and G55 in the overlaid 2D ^{1}H-^{15}N HSQC spectra of 100 µM CaBdf1 BD2 +DMSO-d6 without (**green**) and with 100 µM 10A3, 11A3 and 12A3 (**purple**). Example for fragments showing small CSPs (see I41 and G55). All spectra were measured in a buffer composed of 100 mM NaPi, 50 mM NaCl, pH 7.4, 298 K, and 10% D_2O on a 500 MHz with a 5 mm TCI-prodigy cryoprobe.

For further selection, ^{1}H and ^{15}N shift changes were calculated to determine the weighted average Euclidean distance, with a cutoff at 0.04 ppm.[34] Using this threshold, the original 82 ligands were refined to 8 positive hits (11D6, 4B5, 4E5, 5E9, 1E7, 5F8, 2B8, and 4F6) (Figure 8), corresponding to a hit rate of 1,67%.

Among these, three (1E7, 2B8, 4E6) showed positive STD and CPMG signals, two (4B5, 11D6) were positive only in CPMG, and three (5E9, 4F6, 5F8) showed neither positive STD nor CPMG signals.

Most observed CSPs aligned with the previously reported interacting or neighboring residues[9], except for ligands 2B8, 4F6, and 5F8. These three ligands displayed a more diverse CSP pattern (Figure S18), likely due to the aggregation effect. This is supported by the fact that all three fragments showed signals in the STD control spectrum without added protein. 4F6 and 5F8 also showed increasing signal intensities after the addition of protein in their respective CPMG spectra. For the other ligands, several residues showed consistent and significant shifts, including an unassigned residue @1, along with V415 (shifted in 7 out of 8 fragments), D416 (shifted in 5 fragments), T417 (shifted in 4 fragments), N421 (shifted in 4 fragments), I422 (shifted in 6 fragments), and D473 (shifted in 4 fragments) (Figure 8, Figure S18). These ligands indicate possible specific binding within the binding site.

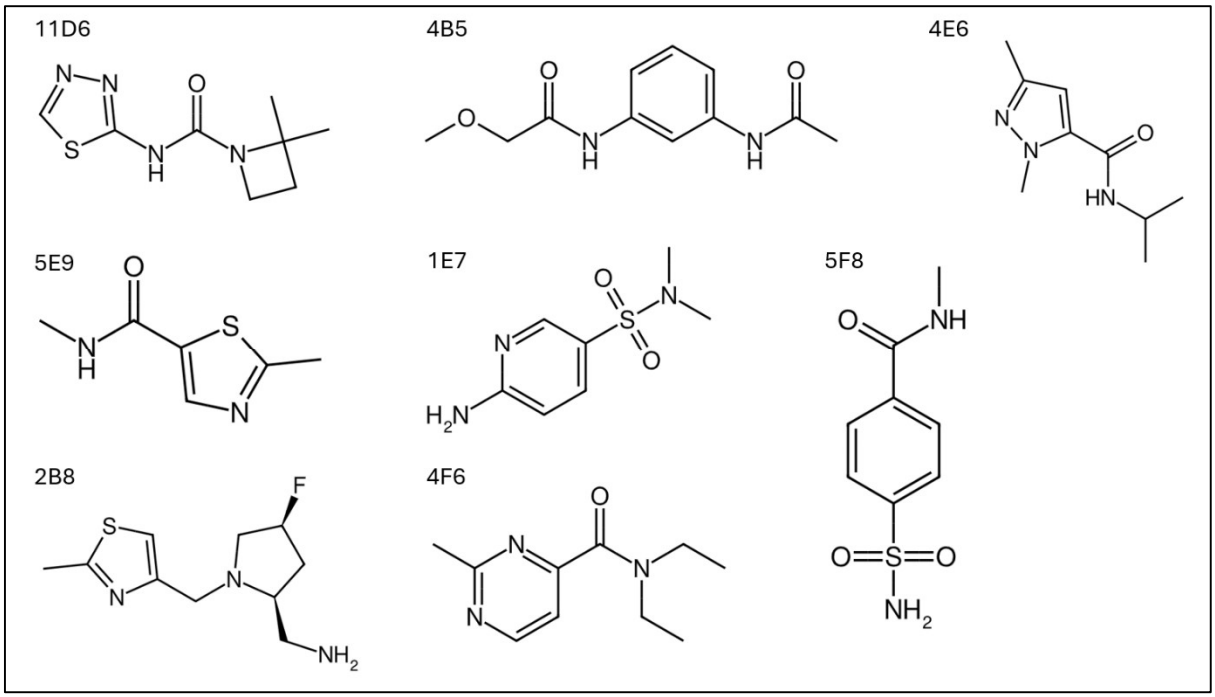

Figure 8: 2D structure of the 8 interacting ligands after the final refinement of the CSPs analysis.

A structure-based pharmacophore model was generated based on the published CaBdf1 BD2 structure complex with an imidazopyridine (PDB: 5N18)[9] (Figure 9A, B). This pharmacophore includes a hydrogen-bond donor and acceptor with the functional group of the molecule and hydrophobic interaction between the aromatic rings and methyl groups. The pharmacophore interacts with amino acids and water molecules 616, 633, and 640 lodged within the binding pocket. Key interacting residues are Val415, Leu420, Ile422, Phe467, Asn468, and Val474.

Using the 8 fragments shown in Figure 8, a ligand-based pharmacophore model was developed (Figure 9C). This model features hydrophobic sites, a hydrogen bond acceptor, and an aromatic ring system, which match some of the features of the structure-based pharmacophore.

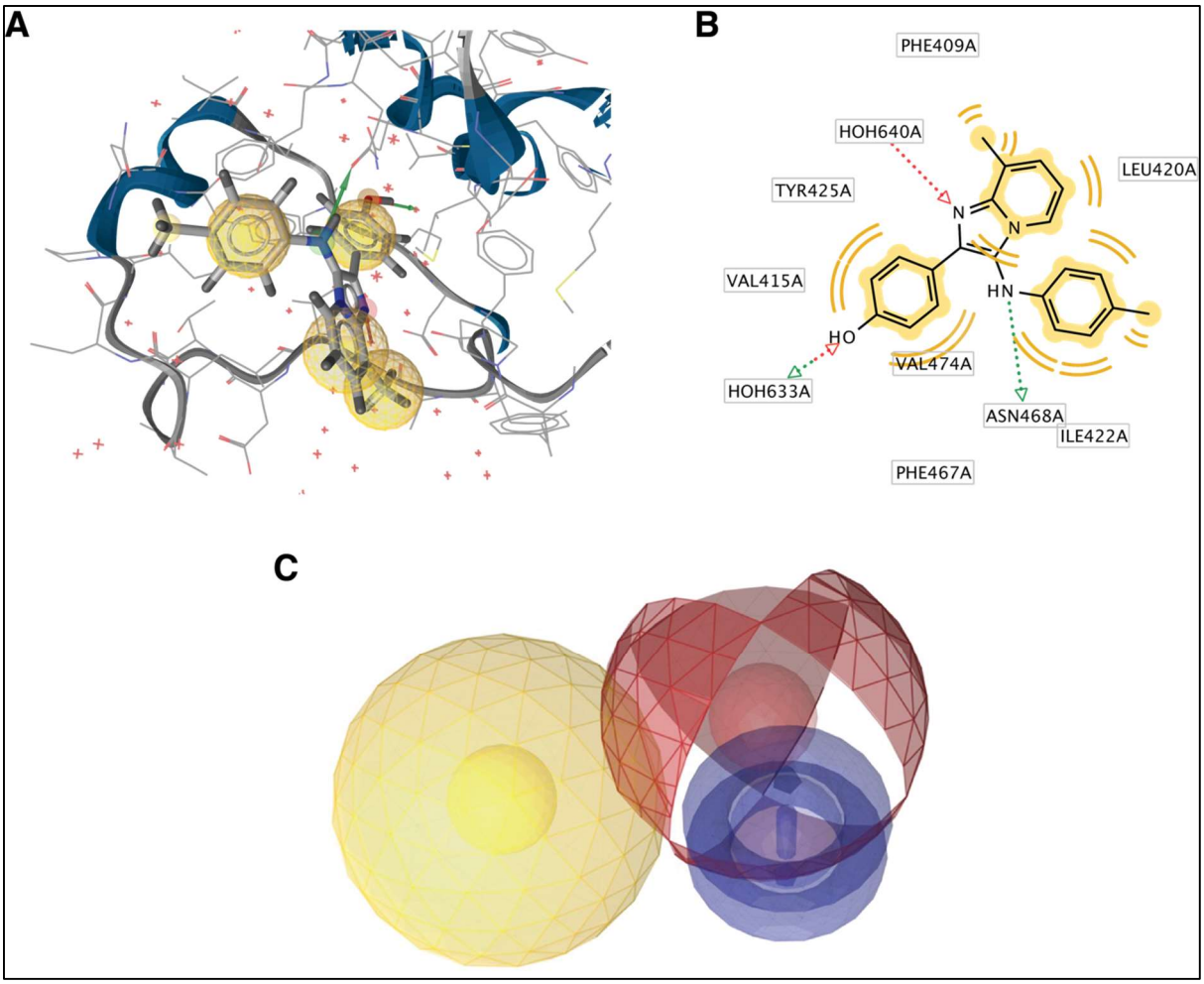

Figure 9: (A-B) A structure-based pharmacophore was created using LigandScout (v.4.5) from the X-ray complex structure of CaBdf1 BD2 in complex with an imidazopyridine (PDB: 5N18). This pharmacophore is represented in the 3D form ***(A)*** and the 2D where the compound is aligned to it ***(B)***. **Yellow** represents hydrophobic interactions, **Green** represents hydrogen bond donors, and **Red** represents hydrogen bond acceptors. ***(C)*** 3D representation of the ligand-based pharmacophore using the 8 hits. **Yellow** represents hydrophobic interactions, **Red** represents a hydrogen bond acceptor, and **Purple** represents interacting aromatic rings (pi-pi-stacking interactions).

1.4 Discussion

Fragment-based drug discovery (FBDD) has greatly risen in popularity in recent years and has led to the discovery and clinical trials of more than two dozen potential drugs.[34] Using libraries of low molecular weight compounds with usually relatively low affinity to the target protein, a large chemical space can be screened to gather information about possible binding sites and the chemical properties of potential binders.[14,18] Using small, interactive fragments does, however, have certain pitfalls. Besides compound stability, low-level impurities, and reactive groups on fragments, aggregation is one of the most significant causes of problems and misinterpretations. Despite its advantages, FBDD is prone to certain challenges, particularly aggregation, which can lead to false positives and other measurement artifacts.[19,34]

Given the low affinity and dynamic binding of fragments, highly sensitive techniques like NMR spectroscopy are especially suitable for FBDD. NMR offers the advantage of observing interactions from both ligands and protein perspectives.[18] However, the accuracy of NMR results may be compromised by limitations such as shifts in pH caused by acidic or basic fragments, which can affect resonance frequencies.[19] In STD experiments, shifts that coincide with the irradiation frequency can result in ligand irradiation artifacts. Additionally, STD experiments conducted with excess ligand concentration may underrepresent strong binders with low dissociation constants (K_D), as these binders have prolonged binding that reduces their representation in the bulk ligand population.[19]

This thesis screened CaBdf1 BD2 against a library of 879 fragments using various NMR techniques, including STD and ^{1}H-CPMG for hit identification and 2D ^{1}H,^{15}N HSQC for cross-validating. To carry out these experiments, 162mg of ^{15}N-labeled CaBdf1 BD2 was successfully expressed and purified.

Initial data analysis was performed using CPMG spectra with a relaxation time of 400 ms. This time point was chosen because, at 8 ms, substantial protein signals remained because the signals did not have enough time to fully relax, making the analysis very difficult due to overlapping signals. As the fragments are much smaller than the protein, they relax more slowly, making a relaxation time of 400 ms more suitable for analyzing the relaxation behavior of the ligand resonances.

Some CPMG measurements showed unexpected behavior, such as increased signal intensity for certain resonances upon adding the protein. This anomaly could indicate the presence of aggregates or small-molecule impurities in the sample. Aggregation refers to the clustering of molecules due to intermolecular interactions (such as van-der-Waals forces or hydrophobic interactions), which can create artifacts in the data.[19,34]
Mixing different ligands may also influence their behavior in solution, further promoting aggregation. Artifacts, likely caused by aggregation, were also observed in negative control STD spectra, suggesting non-specific protein interactions that could cause false positives across STD, CPMG, and HSQC data.[34]

As an orthogonal screening technique, a ^{1}H-^{15}N-HSQC spectrum was measured for all fragments. To refine the data, fragments were categorized into no, minor, and large CSPs. This qualitative grouping, performed by visual assessment, provided a basis for initial hit

identification but has limitations of subjectivity. For A more precise ranking, a quantitative approach such as K_D determination would be necessary.[23] Nonetheless, this method sufficed for preliminary screening.

The comparison across techniques revealed an 11,7% hit rate, with 56 fragments showing positive signals across all three techniques. Out of the 165 fragments showing large CSP, 56 fragments showed positive STD and CPMG signals, 34 showed only STD signals, 37 only CPMG signals, and 38 showed no signals in either. Further deconvolution with one ligand per tube in HSQC experiments identified 82 ligands with significant CSPs, representing a refined hit rate of 9,15% of the totality of the 896-fragment library. Applying a CSP cutoff of 0.04 ppm[34] further narrowed the hits to eight fragments, ensuring that potential binding events were not likely due to nonspecific interactions caused by aggregation. The 6 residues that shifted the most between the 8 positive hits were then marked in the three-dimensional structure of the published inhibitor of CaBdf1 BD2[9] (PDB: 5N18) (Figure 10).

Of these eight fragments, three (1E7, 2B8, 4E6) showed positive signals in all three techniques, while two (4B5, 11D6) showed positive CPMG only, and three (5E9, 4F6, 5F8) had neither positive STD nor CPMG signals. Further analysis revealed that residues V415, D416, and I422 shifted most frequently, aligning with known binding interactions in CaBdf1 BD2, suggesting a possible orientation for ligand binding within the pocket.

A structure-based pharmacophore was created from the CaBdf1 BD2 complexed with an imidazopyridine (PDB: 5N18) for comparison (Figure 9A, 9B). In addition, the 8 identified hits from previous experiments were used to generate the ligand-based pharmacophore (Figure 9C).

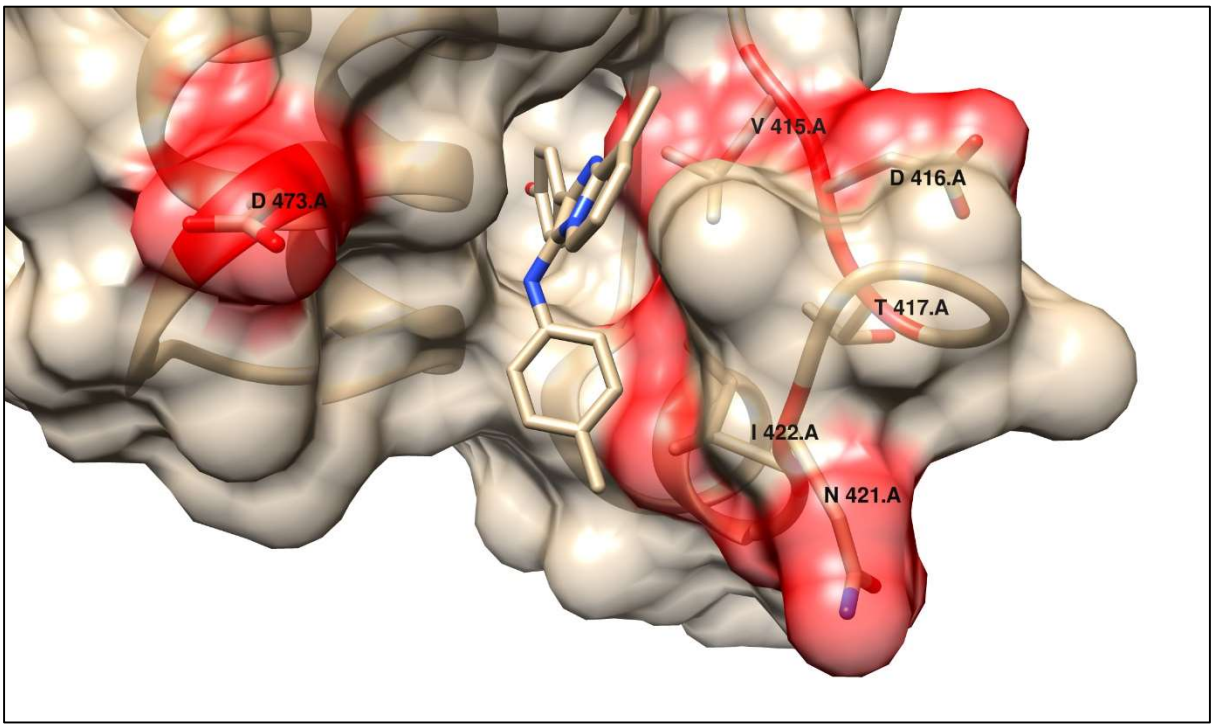

Figure 10: Closeup surface representation of the orthosteric site of CaBdf1 BD2 in complex with an imidazopyridine (PDB: 5N18). The six residues shifted most often in the ^{1}H-^{15}N-HSQC spectra of the fragments after the final refinement are marked in red and labeled.

All properties of the ligand-based pharmacophore calculated using the 8 fragments shown in Figure 8 were represented in the structure-based model, supporting their relevance for potential binding. Due to the small size of the fragments, multiple binding orientations or solvent effects could still influence binding behavior, as smaller fragments leave space within the pocket for water molecules.

In standard screening protocols, unlabeled protein is typically produced, and initial hit identification is performed using fluorescence assays, followed by cross-validation with NMR spectroscopy. However, we deviated from this established workflow in this study, opting to identify hits directly using HSQC-based NMR experiments. This approach, while unconventional, resulted in increased material consumption and necessitated repeating certain steps, ultimately slowing down the analysis and complicating the workflow. Sticking to the traditional screening sequence may have streamlined the process and enhanced efficiency.
Future work could focus on improving the sensitivity of STD and CPMG measurements, potentially by using deuterated buffers to eliminate protein resonance interference, thereby enhancing data clarity and comparability. Additionally, incorporating a non-ionic detergent like Triton X-100 could help reduce aggregation[34], thus minimizing false positives. Extending screening efforts to CaBdf1 BD1 or chemically modifying identified ligands through fragment merging or growth strategies may increase binding affinity and specificity, ultimately contributing to the development of potent CaBdf1 inhibitors.

1.4.1 Conclusion

In summary, this chapter highlights the efforts to screen BD2 of CaBdf1 with a library of 896 fragments. From the initial 896 compounds, 479 fragments were selected after eliminating the non-soluble ones. Using the NMR screening methodology, 56 fragments appear to interact with our receptor. To minimize false positives due to aggregation or other artifacts, the 8 most promising fragments were selected for detailed analysis. These fragments showed significant CSPs in six key residues (V415, D416, T417, N421, I422, and D473). Based on the published orthosteric site, these residues were present there. These findings suggest that these 8 promising ligands might be binding to one cavity of the orthosteric site.

Further validation through structure-based and ligand-based pharmacophore comparisons reinforced the relevance of these fragments. All essential features in the ligand-based pharmacophore were reflected in the structure-based model, highlighting these fragments as candidates for further development.

In conclusion, this modified FBDD approach validates these fragments as promising candidates for developing selective inhibitors. A follow-up for this work by deriving the affinity of these ligands, growing or combining fragments, and extending the screening efforts to CaBdf1 BD1 would lead to the potential development of new inhibitors targeting CaBdf1.

Chapter 2 – ^{1}H-^{1}H NOESY spectrum simulation

2.1 Introduction

The Nuclear-Overhauser-Effect (NOE) is a powerful phenomenon observed in nuclear magnetic resonance (NMR) spectroscopy that provides critical insights into molecular structure and dynamics. It occurs when two NMR-active nuclei, located close to each other in space, influence each other's signal intensities. This interaction happens without the nuclei needing to be connected by chemical bonds (J-coupling) or even belong to the same molecule; they only need to be in spatial proximity. NOE is a valuable tool in structural biology, especially for determining distances between atoms in proteins, nucleic acid, and other macromolecules.[16] As the interacting nuclei do not have to be part of the same molecule one can distinguish between intramolecular and intermolecular cross peaks.[35] For the analysis of protein-ligand interaction the most interesting peaks are the intermolecular cross peaks since it is possible to derive the distances between protein and ligand nuclei from the build-up of said cross peaks.[36,35] Unfortunately, intermolecular cross peaks can be difficult to observe in experimental spectra due to the faster decay of the broad resonances from the protein.[35]

The strength of this effect is proportional to the cross-relaxation rate, a parameter governed by the dipole-dipole relaxation mechanism, which describes the interactions between local magnetic fields. The magnitude and sign of the cross-relaxation rate are influenced by several factors, including the nature of the interacting nuclei, the strength of the magnetic field, the distance between the nuclei, and the rotational tumbling time of the molecules involved. One key aspect of the NOE is its sensitivity to internuclear distance: the cross-relaxation rate decreases rapidly as the distance between nuclei increases, following a $1/r^6$ dependence.[16]

Additionally, the rotational tumbling time of molecules plays a significant role in the NOE. For small molecules, the cross-relaxation rate is positive, whereas for larger molecules, it becomes negative. This behavior makes NOE an essential method for studying molecular size and three-dimensional (3D) structures and distinguishing between different states of molecular interaction.[16]

2.1.1 Matrix representation of a 2-dimensional ^{1}H-^{1}H-NOESY spectrum

The peak intensities in a two-dimensional (2D) ^{1}H-^{1}H-NOESY spectrum of a protein-ligand complex, in exchange with its free ligand and free protein states, are determined by the extent of the magnitudes of magnetization transfer. This magnetization transfer is described by equations (2.1) and (2.2), with equation (2.2) representing the solution to equation (2.1).[37,35]

$$\frac{dM(t)}{dt} = -(R + K) * (M(t) - M_0) \qquad (2.1)$$

$$M(\tau_{mix}) = \exp(-(R + K) * \tau_{mix}) * (M(0) - M_0) + M_0 \qquad (2.2)$$

Here, R is the relaxation matrix, K is the kinetic matrix, M(0) is the initial magnetization, M_0 is the equilibrium magnetization, and τ_{mix} the mixing time.[37]

Assuming a two-component, three-species equilibrium system described in equation (2.3), we derive the differential equations for the rate of change of the free ligand, free protein, and protein-ligand complex, as described in equations (2.5-2.7). In these equations, [L], [P], and [PL] represent the concentrations of the ligand, protein, and protein-ligand-complex, respectively. k_{on} and k_{off} are the association and dissociation rate constants.[35,37,38]

$$L + P \underset{k_{off}}{\overset{k_{on}}{\leftrightharpoons}} PL \tag{2.3}$$

$$K_D = \frac{k_{off}}{k_{on}} = \frac{[Protein]*[Ligand]}{[Protein{-}Ligan \quad complex]} \tag{2.4}$$

$$\frac{d[L]}{dt} = -k_{on} * [L][P] + k_{off} * [PL] \tag{2.5}$$

$$\frac{d[P]}{dt} = -k_{on} * [L][P] + k_{off} * [PL] \tag{2.6}$$

$$\frac{d[PL]}{dt} = k_{on} * [L][P] - k_{off} * [PL] \tag{2.7}$$

The kinetic matrix K is constructed based on equations (2.4 – 2.7) and can be written as follows in equation (2.8):[37,38]

$$K = \begin{bmatrix} k_{on}[P] * I & & -k_{off} * I & \\ & k_{on}[L] * I & & -k_{off} * I \\ -k_{on}[P] * I & & k_{off} * I & \\ & -k_{on}[L] * I & & k_{off} * I \end{bmatrix} \tag{2.8}$$

I represents the identity matrix of the same dimension as the corresponding block matrix in R.[37]

In equation 2.9, the relaxation matrix R, a block matrix, includes the sub-matrices R_L^f, R_P^f, R_L^b and R_P^b which describes the proton-proton relaxation in both the free and bound form of the ligand and the protein. The submatrices $R_{P,L}^b$ and $R_{L,P}^b$ represent the cross-relaxation pathway between protons in the protein-ligand complex.[35,37]

$$R = \begin{bmatrix} R_L^f & & & \\ & R_P^f & & \\ & & R_L^b & R_{L,P}^b \\ & & R_{P,L}^b & R_P^b \end{bmatrix} \tag{2.9}$$

The elements of the relaxation matrix are calculated using equations (2.10-2.11) as follows[37]:

$$R_{i,i} = \rho_i = \sum_{\substack{H_j \in A \\ j \neq i}} \frac{b^2}{d_{ij}^6} * (J(0) + 3J(\omega) + 6J(2\omega)) \tag{2.10}$$

$$R_{i,j} = \sigma_{ij} = \frac{b^2}{d_{ij}^6} * (6J(2\omega) - J(0)) \quad (2.11)$$

The spectral density function J(ω) is given by equation (2.12)[37]:

$$J(\omega) = \frac{2}{5}\left(\frac{\tau_c}{1+(\omega\tau_c)^2}\right) \quad (2.12)$$

Hereby, b is a constant defined in equation (2.13)[37]:

$$b = \frac{1}{2} * \frac{\mu_0}{4\pi} \hbar \gamma_H^2 \quad (2.13)$$

In these equations, ρ_i represents the longitudinal relaxation rate of proton H_i, $\sigma_{i,j}$ is the cross-relaxation rate between protons H_i and H_j, and $d_{i,j}$ is the distance between these protons. All protons H_j are considered for relaxation calculations ρ_i and $\sigma_{i,j}$, regardless of their distance $d_{i,j}$.[37]

2.1.2 Kirsten rat sarcoma virus protein (K-Ras)

K-Ras (Kirsten rat sarcoma virus) protein is a small membrane-bound GTPase involved in the signaling cascades regulating biological functions like cell proliferation and differentiation. K-Ras is a molecular switch turned from the active to the inactive state through the hydrolysis of Guanosine triphosphate (GTP) to Guanosine diphosphate (GDP).[39] K-Ras protein is encoded by the proto-oncogene K-Ras, the most frequently mutated RAS family member in human cancer cells. The most frequent mutations are substitutions of Glycine 12 to Valine or Aspartate and Glycine 13 to Aspartate.[40]

2.1.3 Objective

In this section, I present the simulation of a 2D-[^{1}H, ^{1}H]-NOESY spectrum for a K-Ras protein-ligand complex with (2E)-3-(1H-indol-2-yl) prop-2-enoic acid (Z5I) using a full matrix approach. Using a Python script, I generate a solution matrix that mirrors a measured NOESY spectrum, providing insight into the intensity evolution of diagonal peaks, intramolecular cross-peaks, and intermolecular cross-peaks as functions of mixing time τ_c, temperature T, and total protein and ligand concentration. Furthermore, the simulated data is compared to the temperature series of F_1-^{15}N, ^{13}C-filtered ^{1}H-^{1}H-NOESY spectra from a K-Ras protein-ligand complex with 7-fluoro-N, N-dimethyl-1-benzofuran-2-carboxamide.

2.2 Materials and Methods

2.2.1 Recoding of the NMR data

The protein was produced and provided by Dr. Nicolas Coudevylle.

A series of F_1-^{15}N, ^{13}C-filtered ^{1}H-^{1}H-NOESY spectra were recorded using 0.8 mM labeled ^{13}C-^{15}N K-Ras G13D bound to GMP-PNP in the presence of 4 mM 7-fluoro-N, N-dimethyl-1-benzofuran-2-carboxamide in 30 mM, deuterated NaPi buffer with 200 mM NaCl and 2 mM TCEP at a pD of 7.4, on a Bruker Avance III 700 MHz spectrometer (Figure S19). The pulse program noesygpphwgx1[41,42,43,44] was used, with an acquisition time of 0.098 s. Spectra acquisition included 2048 complex points in the direct dimension (t_2) and 256 in the indirect dimension (t_1). An interscan delay of 1 second was used, with each increment scanned 256 times, yielding an experiment duration of approximately 22 h 10 min. The same parameters were used for temperature measurements at 280 K, 289 K, and 298 K.
The spectra were processed using Topspin 4.1.4.

2.2.2 ^{1}H-^{1}H-NOESY spectrum simulation

A ^{1}H-^{1}H-NOESY spectrum of K-Ras G12V (GMPPNP bound) in complex with (2E)-3-(1H-indol-2-yl)prop-2-enoic acid (Figure 11, PDB: 8PIY) was successfully simulated using the full-matrix approach, following equation 2.2. The system model included all ligand's hydrogen atoms (9 in total) and protein's hydrogen atoms within a 7 Å radius of the ligand (135 atoms), resulting in a 288x288 quadratic solution matrix. All constants used in these calculations are listed in Table 1.

The equilibrium concentrations of the ligand, the protein, and the protein-ligand complex were determined by applying equations 2.14 and 2.15[38]:

$$\left[PL_{eq}\right] = [PL] = \frac{[L_t]+[P_t]+K_D-\sqrt{([L_t]+[P_t]+K_D)^2-4*[L_t][P_t]}}{2} \qquad (2.14)$$

$$\left[L_{eq}\right] = [L] = [L_t] - [PL] \; and \; \left[P_{eq}\right] = [P] = [P_t] - [PL] \qquad (2.15)$$

To model the rotational tumbling time (τ_c), I used Stoke's law (equation 2.16), which depends on the hydrodynamic radius (r_H), the temperature, the viscosity (η) of the solvent, and the Boltzmann constant (k). r_H was calculated via equation 2.17, considering the protein's molecular weight (MW), average density, and Avogadro's number.[45]

$$\tau_c = \frac{4*\pi*\eta*r_H^3}{3*k*T} \qquad (2.16)$$

$$r_H = \sqrt[3]{\frac{3*MW}{4*\pi*\rho*N_A}} + r_w \qquad (2.17)$$

Under fast exchanging conditions (k_{ex}>>Δω for free and bound states) the experimental NOESY spectrum was assumed to contain overlapping peaks from free and bound states. In the solution matrix of the simulation, those species have individual values that need to be summed up before simulated and experimental data can be compared. Consequently, the cross-peak (cp) intensities were calculated by summing interactions between both states (equation 2.18), and diagonal peaks (dp) were summed similarly (equation 2.19).[35,38]

$$CP(AX) = CP(A_f X_f) + CP(A_f X_b) + CP(A_b X_f) + CP(A_b X_b) \quad (2.18)$$

$$DP(A) = CP(A_f A_b) * 2 + DP(A_f) * DP(A_b) \quad (2.19)$$

To improve comparability with experimental spectra, the cp intensities were normalized using equation 2.20).

$$CP(AX)_N = \frac{CP(AX)}{\sqrt{DP(A)*DP(X)}} * \frac{\sqrt{N_1*N_2}}{N_1*N_2} \quad (2.20)$$

N1 and N2 are the number of equivalent spins within spin groups 1 and 2.

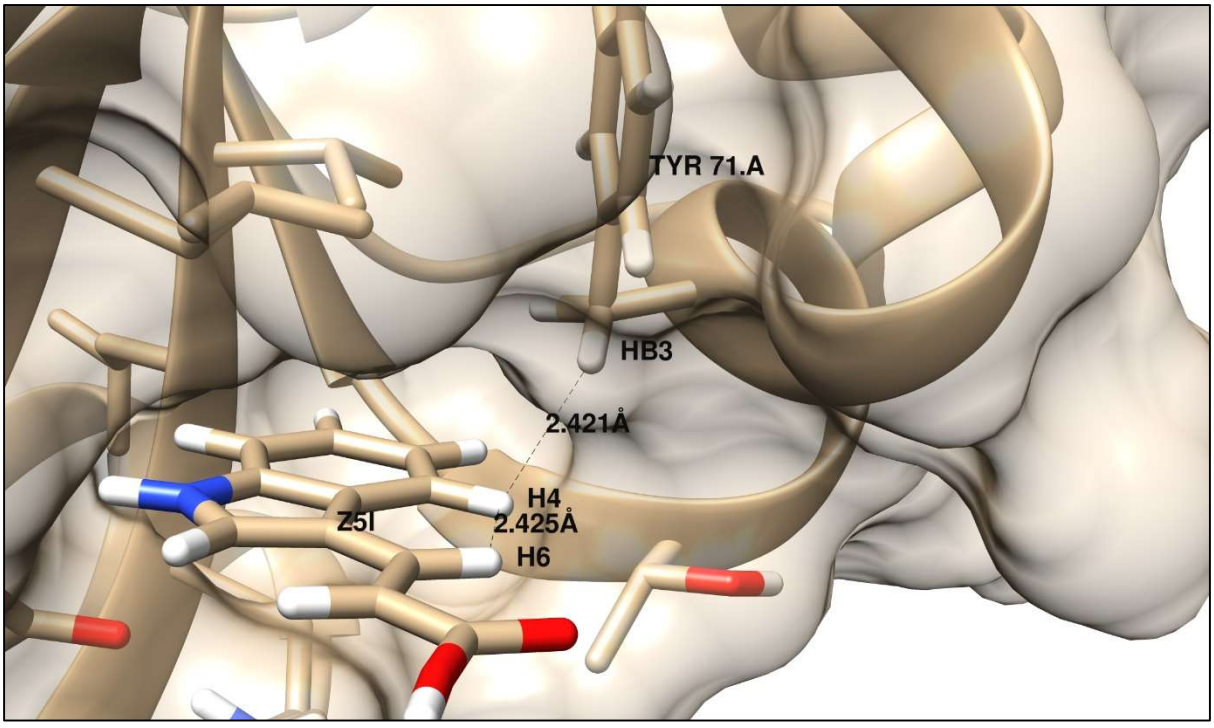

Figure 11: 3D structure of first conformer of the structure of K-Ras G12V (GMPPNP bound) in complex with (2E)-3-(1H-indol-2-yl)prop-2-enoic acid (Z5I)(PDB: 8PIY). H4 and H6 of Z5I are highlighted, as well as 71 Tyr HB3. The relevant distances are shown as dashed lines.

The simulated matrix was analyzed under varying conditions: first, across different mixing times (τ_{mix}), and second, with temperature as a variable. Temperature adjustments required recalculating temperature-dependent parameters, such as the solvent's viscosity η, the rotational tumbling time τ_c, and the dissociation constant K_D. The temperature-dependent behavior of K_D was modeled using the Van't Hoff enthalpy equation (equation 2.21)[46], while

the solvent's viscosity was described with a three-parameter exponential function (equation 2.22).[47]

$$\ln(K_D) = \frac{\Delta H}{RT} - \frac{\Delta S}{R} \Leftrightarrow K_D = e^{\frac{\Delta H}{RT} - \frac{\Delta S}{R}} \quad (2.21)$$

$$\mu = A * 10^{B/T - C} \quad (2.22)$$

Assuming the K_D is known for the protein-ligand interaction, ΔS can be calculated using the known K_D, the corresponding temperature, and a constant value of ΔH, where ΔH is positive for an endothermic reaction and negative for exothermic reactions. With ΔS calculated, K_D values for varying temperatures can be derived accordingly, assuming that ΔS and ΔH are constant with varying temperatures.[46]

Table 1: Constant variables and constants used in the simulation of the ^{1}H-^{1}H-NOESY spectrum

Variabel	Value	Unit
ω_0	4398229715.02571	Rad/s
ħ	$1.054571817 * 10^{-34}$	m^2*kg/s
μ_0	$4 * \pi * 10^{-7}$	H/m
γ_H	$26.752 * 10^7$	Rad/s*T
N_A	$6.02214086 * 10^{23}$	1/mol
k	$1.38064852 * 10^{-23}$	$m^{2*}kg/s^2$*K
ρ	1370000	g/m^3
B_0	16.4	T
R	8.31446261815324	J/K*mol
A	$2.414*10^{-5}$	Pa*s
B	247.8	K
C	140	K

2.3 Results

The Nuclear Overhauser effect (NOE) provides insights into the spatial proximity of proton spins within three-dimensional space, influenced by factors such as rotational tumbling time, mixing time, and inter-proton distance. In a two-dimensional ^{1}H-^{1}H-NOESY spectrum, the interaction between spatially proximity protons – even from two distinct molecules – can be detected if they are in close proximity long enough to permit cross-relaxation.[16] This property makes NOESY spectra valuable for analyzing protein-ligand complexes, particularly for assessing parameter effects like mixing time, temperature, and concentration.

The simulated NOESY spectrum assumed fast exchange conditions and considered cross-peaks between individual ligand protons and protein protons. Figure 12 illustrates the evolution of diagonal peaks for ligand proton H4 (Figure 12A) and protein proton Tyr 71 HB3 (Figure 12B), as well as the intramolecular cross-peak between ligand proton H4 and H6 (Figure 12C), the intermolecular cross-peak between ligand proton H4 and protein proton Tyr 71 HB3 (Figure 12D), and the normalized intra- and intermolecular cross-peaks (Figure 12 E-F) as a function of the mixing time τ_{mix} for the K_D values of 0.6 mM (Green curve), 1.6 mM (Black curve) and 2.6 mM (Red curve). Initial conditions included a protein concentration of 800 µM, ligand concentration of 4000 µM, ΔH of ±50 kJ/mol, K_D of 1600 µM, temperature of 298 K, and a k_{off} of 10^8 $M^{-1}s^{-1}$. k_{on} was calculated using equation 2.4. Cross-peaks were normalized per equation 2.20.

At τ_{mix} = 0 s, the diagonal peaks are at their maximum, with no relaxation. As the mixing time increases, the diagonal peaks start decaying, with the rate of decay of the protein diagonal peak being larger than that of the diagonal peak of the ligand (Figure 12A-B). Both the intra- and the intermolecular cross-peak are 0 at τ_{mix} = 0 s (Figure 12C-D).
As τ_{mix} increases, the cross-peak intensities rise until reaching their respective maxima, followed by a decay. The intramolecular cross-peak maximum at ~400 a.u. is reached after ~0.7 s (black curve, Figure 12C). In contrast, the maximum of the intermolecular peak at ~70 a.u. is reached after ~0.27 s (black curve, Figure 12D). The initial buildup rate of the intramolecular cross-peak is slightly larger than that of the intermolecular cross-peak (1670 a.u./s > 1440 a.u./s for K_D 0.6 mM, 1370 a.u./s > 1160 a.u./s for K_D 1.6 mM, 1170 a.u./s > 980 a.u./s for K_D 2.6 mM; initial (0 s-0.01 s) build-up rates of intra- vs intermolecular cross-peak intensities).
The normalized intra- and intermolecular cross-peaks increase, nearing 1, as the mixing time increases (Figure 12E-F). The normalized curves show contradictory behavior to the non-normalized curves, as the buildup of the normalized intermolecular cross-peak is faster than that of the intermolecular curve (0.42 s^{-1} < 0.92 s^{-1} for K_D 0.6 mM, 0.35 s^{-1} < 0.74 s^{-1} for K_D 1.6 mM, 0.29 s^{-1} < 0.63 s^{-1} for K_D 2.6 mM; initial (0 s-0.01 s) build-up rates of intra- vs intermolecular cross-peaks).
When changing the dissociation constant, the curves shift. The ligand diagonal peak relaxes more slowly for increasing K_D, if only slightly, while the effect on the protein diagonal peak is almost indistinguishable. The maxima of the cross-peak intensities are decreasing with increasing K_D. This behavior is confirmed by the normalized cross-peak curves, where the initial buildup rate of intra- and intermolecular cross-peaks is also decreasing with increasing K_D.

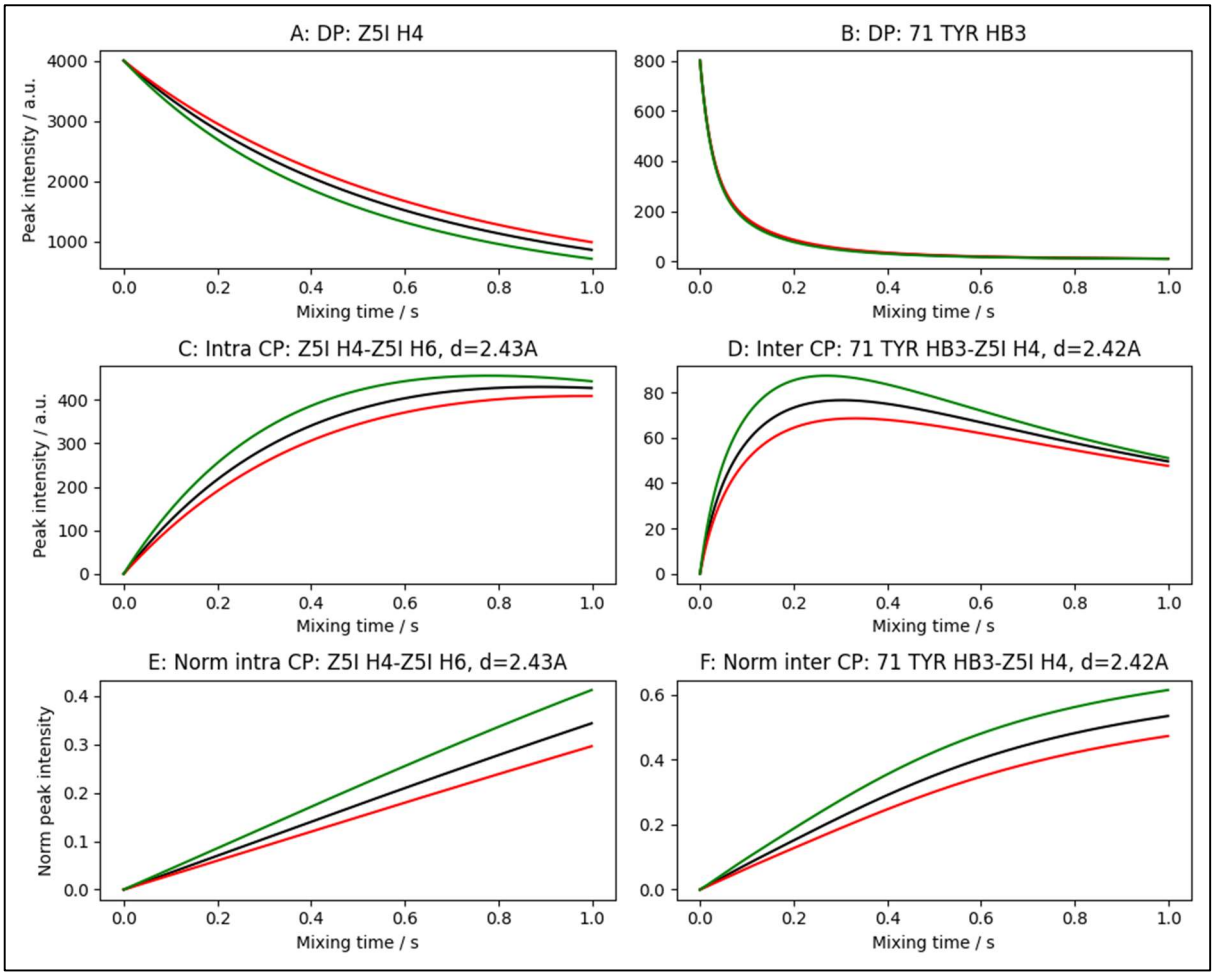

Figure 12: Simulated evolution of diagonal- and cross-peak intensities of the simulated ^{1}H-^{1}H NOESY spectrum of K-Ras G12V (GMPPNP bound) in complex with (2E)-3-(1H-indol-2-yl)prop-2-enoic acid (Z5I) as a function of the mixing time. Simulation parameters: [L]=4000 µM, [P]=800 µM, $k_{off}=10^{8}\ M^{-1}s^{-1}$, v_0=700 MHz, T=298 K. **Black** curves correspond to a K_D of 1.6 mM, **Green** curves correspond to a K_D of 0.6 mM, and **red** curves correspond to a K_D of 2.6 mM. (**A**) Decay of the diagonal peak for Z5I H4. (**B**) Decay of the diagonal peak for 71 Tyr HB3. (**C**) The buildup of the intramolecular cross-peak between Z5I H4 and Z5I H6 with a proton distance of 2.43Å. (**D**) The buildup of the intermolecular cross-peak between Z5I H4 and Tyr-71 HB3 with a proton distance of 2.42Å. (**E**) Normalized buildup of the intramolecular cross-peak Z5I H4-Z5I H6. (**F**) Normalized buildup of the intermolecular cross-peak between Z5I H4-71 Tyr HB3.

Figure 13 illustrates the evolution of temperature-dependent variables. It shows an increase in K_D for an exothermic reaction and a decrease for endothermic reactions with rising temperatures (Figure 13A-B). On the other hand, the concentration of the protein-ligand complex increases for an exothermic and decreases for an endothermic reaction (Figure 13C-D). The dynamic viscosity and the rotational tumbling time τ_c decrease with temperature at a proportional rate (Figure 13E-F). An increasing dissociation constant at 298 K shifts the curves. For the exothermic reaction, the shift is more pronounced for higher temperatures and shifts the K_D curve upward and the concentration curve downward. For the endothermic reaction, the shift is more pronounced at lower temperatures, shifting the K_D curve upward and the concentration curve downward.

An exothermic binding reaction was assumed for the temperature-dependent simulation.

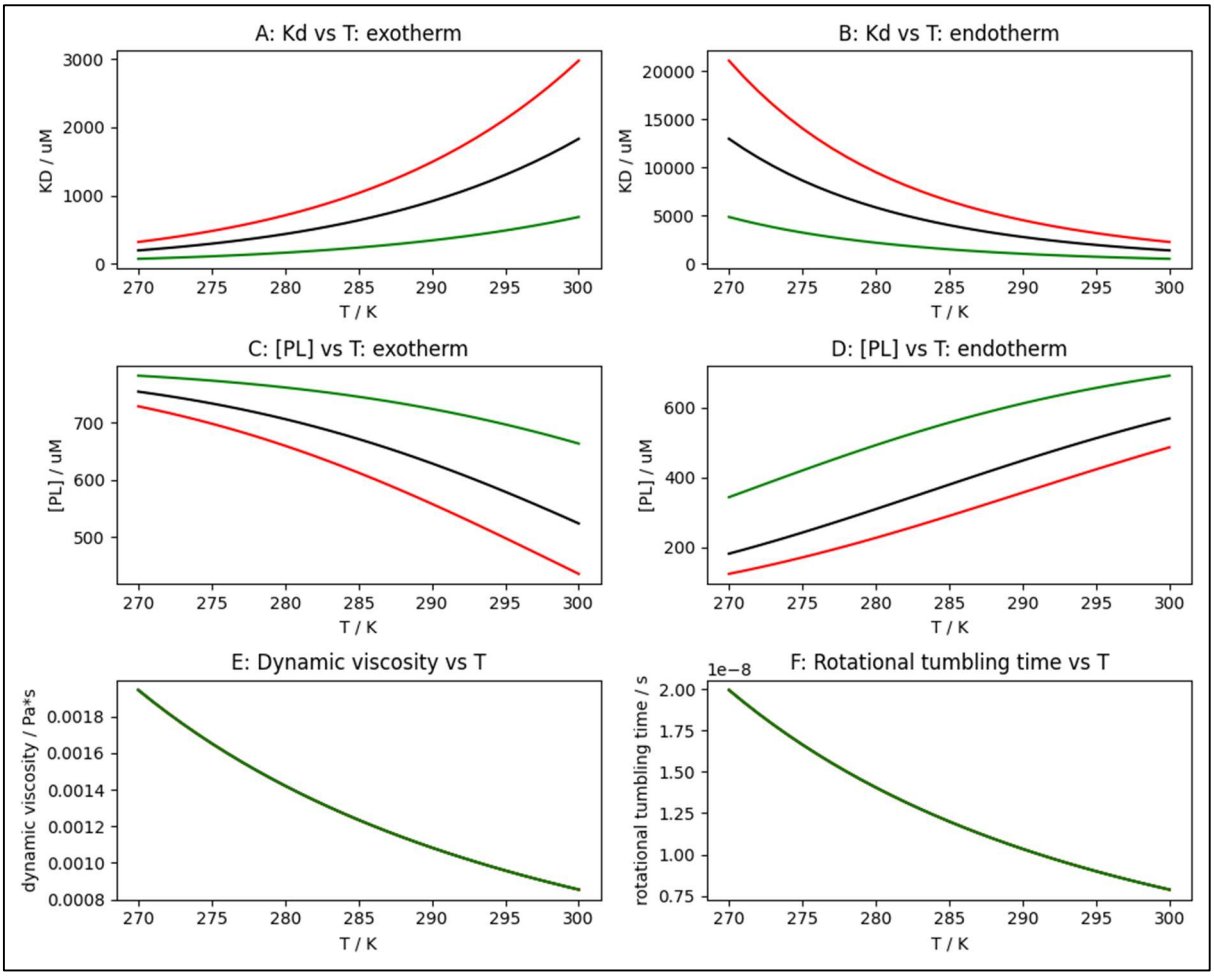

Figure 13: Temperature dependence of various thermodynamic and kinetic parameters. Simulation parameter: [L]=4000 µM, [P]=800 µM, k_{off}=10^8 $M^{-1}s^{-1}$, ν_0=700 MHz, τ_{mix}=0.1 s, ΔH=-50 kJ/mol. **Black** curves correspond to a K_D of 1.6 mM at 298 K, **Green** curves correspond to a K_D of 0.6 mM at 298 K, and **red** curves correspond to a K_D of 2.6 mM at 298 K. (**A**) K_D as a function of temperature for an exothermic binding reaction (ΔH=-50 kJ/mol). (**B**) K_D as a function of temperature for an endothermic binding reaction (ΔH=50 kJ/mol). (**C**) Complex concentration [PL] as a function of temperature for exothermic binding. (**D**) Complex concentration [PL] as a function of temperature for endothermic binding. (**E**) Dynamic viscosity of the solution as a function of temperature. (**F**) Rotational tumbling time τ_c as a function of temperature.

The experimental data showed a temperature-related trend. Figure 14 illustrates the evolution of all diagonal peaks corresponding to the protons of the ligand (Figure 14A) and the protein (Figure 14B), along with three intramolecular cross-peaks between ligand protons (Figure14C), five intermolecular cross-peaks (Figure 14D), and the normalized cross-peaks (Figure 14E-F)(Peaks listed in Tables S16-S18, full NOESY spectra in Figure S19, exemplary peak in Figure 15). The K_D for the binding interaction was determined to be 1.6 mM at 298 K, with measurements conducted at a mixing time of 0.1 s. It was assumed that the cross-peaks arise from the interaction between one ligand proton and a protein methyl group, resulting in three protein protons, where N_1 equals one, and N_2 equals three.

Interestingly, the diagonal- and cross-peaks of the simulated data (Figure 16) are predicted to exhibit different behaviors as temperature increases. Specifically, the diagonal peaks of both protein and ligand increase (Figure 16A-B), while the intramolecular and intermolecular cross-peaks decrease (Figure 16C-D) with rising temperature. The diagonal peaks of the experimental data also increase with temperature (Figure 14A-B). In contrast, the intramolecular cross-peaks and intermolecular cross-peaks 4 and 5 rise with increasing

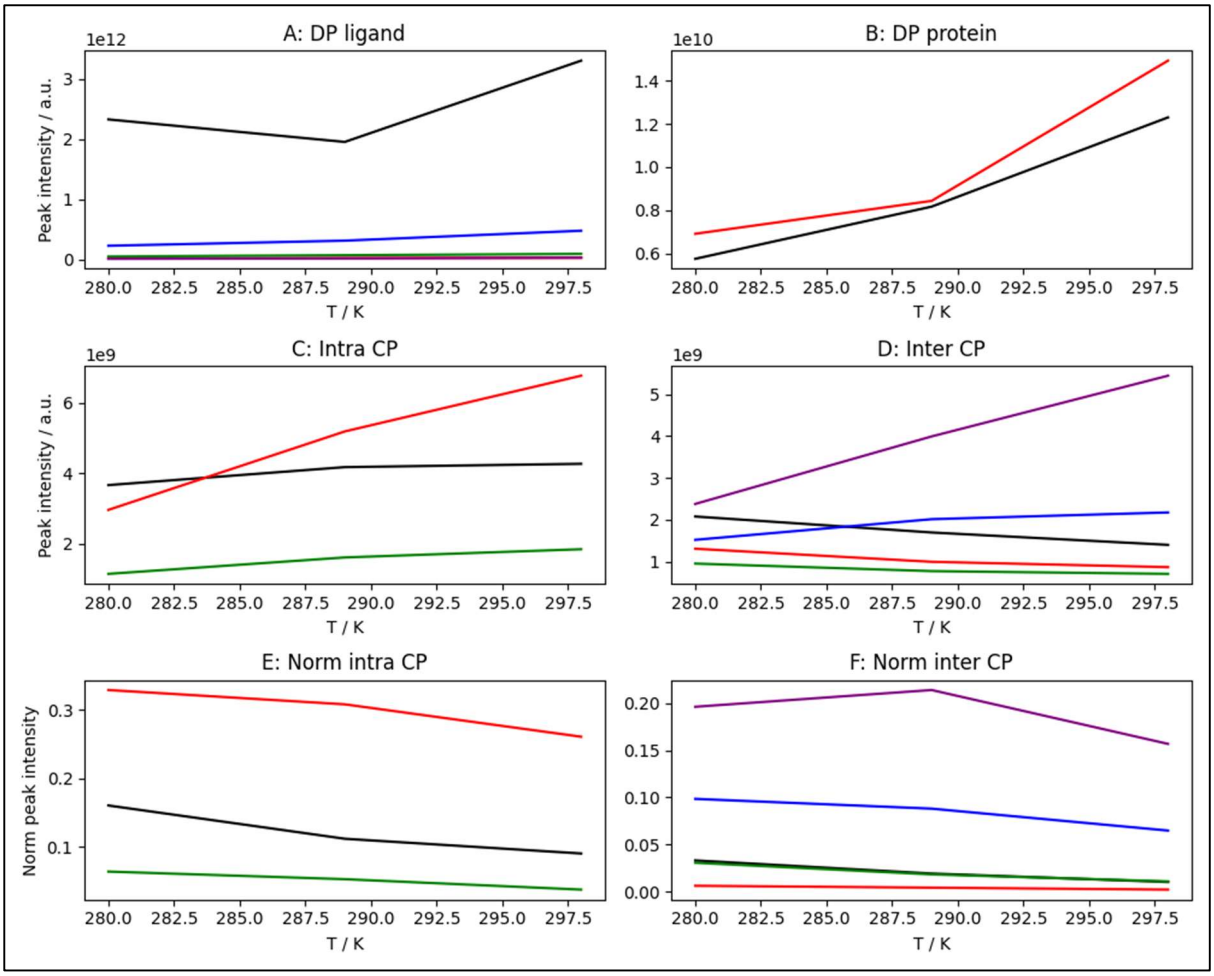

Figure 14: Temperature-dependent evolution of diagonal- and cross-peak intensities in F_1-^{15}N, ^{13}C-filtered ^{1}H-^{1}H-NOESY spectra of 0.8 mM K-Ras G13D (GMP-PNP bound) with 4 mM 7-fluoro-N, N-dimethyl-1-benzofuran-2-carboxamide at 280 K, 289 K and 298 K. Measurements were taken at 700 MHz using cryo platform in 30 mM deuterated NaPi, 200 mM NaCl and 2 mM TCEP at pD 7.4. **(A)** Diagonal peaks for the ligand: DP 3 (*black*), DP 4 (*red*), DP 5 (*green*), DP 6 (*blue*), DP 7 (*purple*). **(B)** Diagonal peaks for the protein: DP 1 (*black*), DP 2 (*red*). **(C)** Intramolecular cross-peaks: DP 1 (*black*), DP 2 (*red*), DP 3 (*green*). **(D)** Intermolecular cross-peaks: CP 1 (*black*), CP 2 (*red*), CP 3 (*green*), CP 4 (*blue*), CP 5 (*purple*) **(E)** Normalized intramolecular cross-peaks: CP 1 (*black*), CP 2 (*red*), CP 3 (*green*) **(F)** Normalized intermolecular cross-peaks: CP 1 (*black*), CP 2 (*red*), CP 3 (*green*), CP 4 (*blue*), CP 5 (*purple*). CP stands for cross-peak and DP for diagonal-peak.

temperature (Figure 14C-D), which deviates from the simulated data. However, intermolecular cross-peaks 1, 2, and 3 decreases with rising temperatures (Figure 14D), aligning with the simulated data. Following the normalization of the intra- and the intermolecular cross-peak intensities, the experimental curves (Figure 14E-F) resemble the simulated curves (Figure16E-F) and decrease as the temperature increases. This decrease is quantified by dividing the normalized intensity at 280 K by that at 298 K (fractions for K_D=1.6 mM, Figure 16E-F, black curves)(Table 2). The fractions of the normalized intra- and intermolecular cross-peaks were found to be 2.31 and 2.16, respectively. The experimental data exhibited some variation in these fractions, ranging from 1.26 to 1.77 for the intra-molecular and 1.25 to 3.16 for the intermolecular cross-peaks.

When changing the dissociation constant at 298 K in the simulation, the curves shift. The shift grows more pronounced with increasing temperature for diagonal- and cross-peaks alike. However, the diagonal-peak curves are shifted upward, and the cross-peak curves are shifted downward.

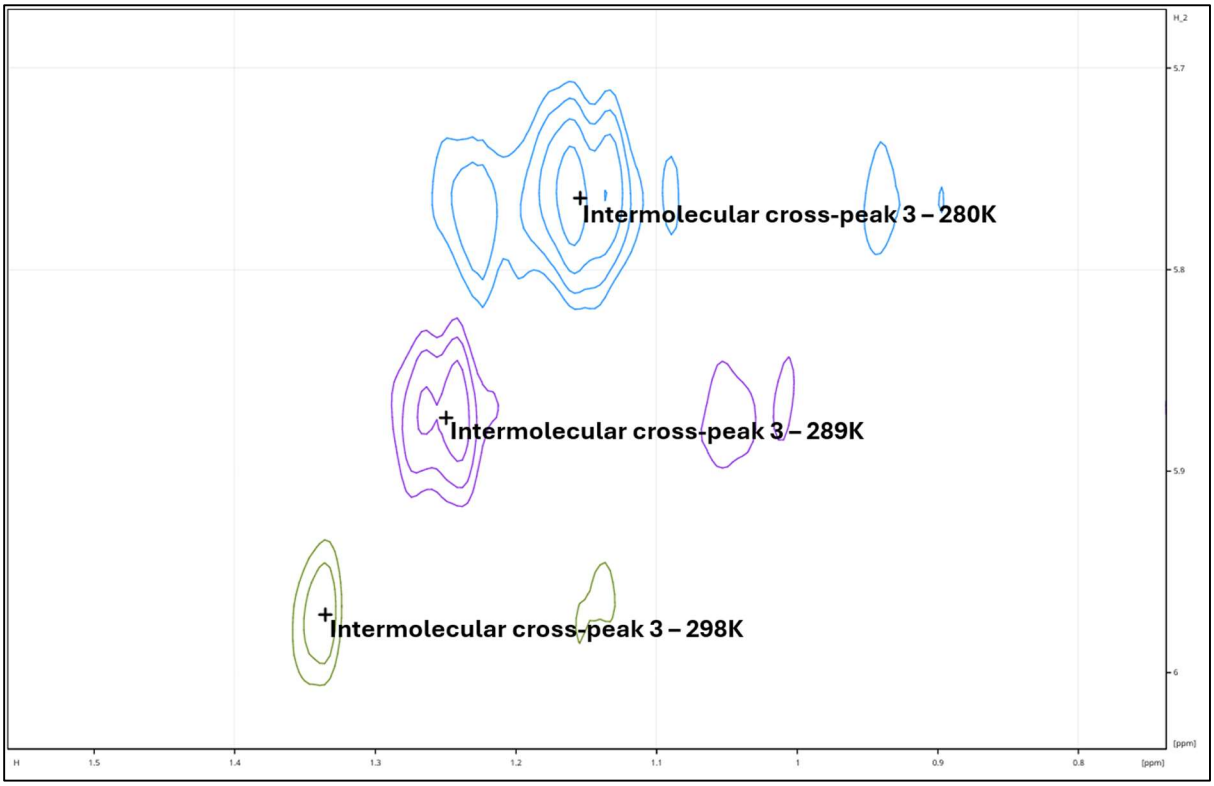

Figure 15: Closeup of the intermolecular cross-peak 3 in the F_1-^{15}N, ^{13}C-filtered ^{1}H-^{1}H-NOESY spectra of 0.8 mM K-Ras G13D (GMP-PNP bound) with 4 mM 7-fluoro-N, N-dimethyl-1-benzofuran-2-carboxamide at 280 K (**blue**), 289 K (**purple**) and 298 K (**green**). Measurements were taken at 700 MHz using cryo platform in 30 mM deuterated NaPi, 200 mM NaCl and 2 mM TCEP at pD 7.4. The temperature variation causes the shifts between the peaks.

Figure 17 illustrates the evolution of the intensities of the diagonal peak of ligand proton H4 (Figure 17A), the diagonal peak of protein proton Tyr 71 HB3 (Figure 17B), the cross-peak of ligand protons H4 and H6 (Figure 17C), the cross-peak of ligand proton H4 and protein proton Tyr 71 HB3 (Figure 17D), and the normalized cross-peaks described before (Figure 17E-F) as a function of the total protein concentration and the total ligand concentration. The normalization was done according to equation 2.20. The initial parameters were the same as in the previous simulations with a mixing time of 0.1 s. The figures are contour plots, where the x- and y-axis depict the total protein and ligand concentration in µM, and the z-axis represents the intensity in arbitrary units (a.u.) as contour levels. Both the diagonal peaks of the protein and ligand show a maximum if the total concentration of the interaction partner is minimal. The intramolecular and intermolecular cross-peak intensities behave similarly, showing a maximum for large total ligand and protein concentrations. The normalized cross-peaks are showing different behavior. While the normalized intramolecular cross-peak shows the maximum for large protein excess, the normalized intermolecular cross-peak increases with both total protein and ligand concentration.

Table 2: Tabulated reactions of the normalized cross-peak intensities of the filtered NOESY at 280 K and 298 K, as well as the fractions of the cross-peak intensities of the simulated data and their respective protons

Data origin	Interaction type	Cross-peak Nr.	Involved peaks/protons	Fraction I(280)/I(298)
Simulated data	**Intramolecular**	1	Z5I H4-Z5I H6	2.31
	Intermolecular	1	71 Y HB3-Z5I H4	2.16
Experimental data	**Intramolecular**	1	DP4-DP5	1.77
		2	DP4-DP7	1.26
		3	DP5-DP7	1.7
	Intermolecular	1	DP1-DP6	3.16
		2	DP1-DP3	2.63
		3	DP2-DP5	2.77
		4	DP2-DP4	1.52
		5	DP2-DP7	1.25

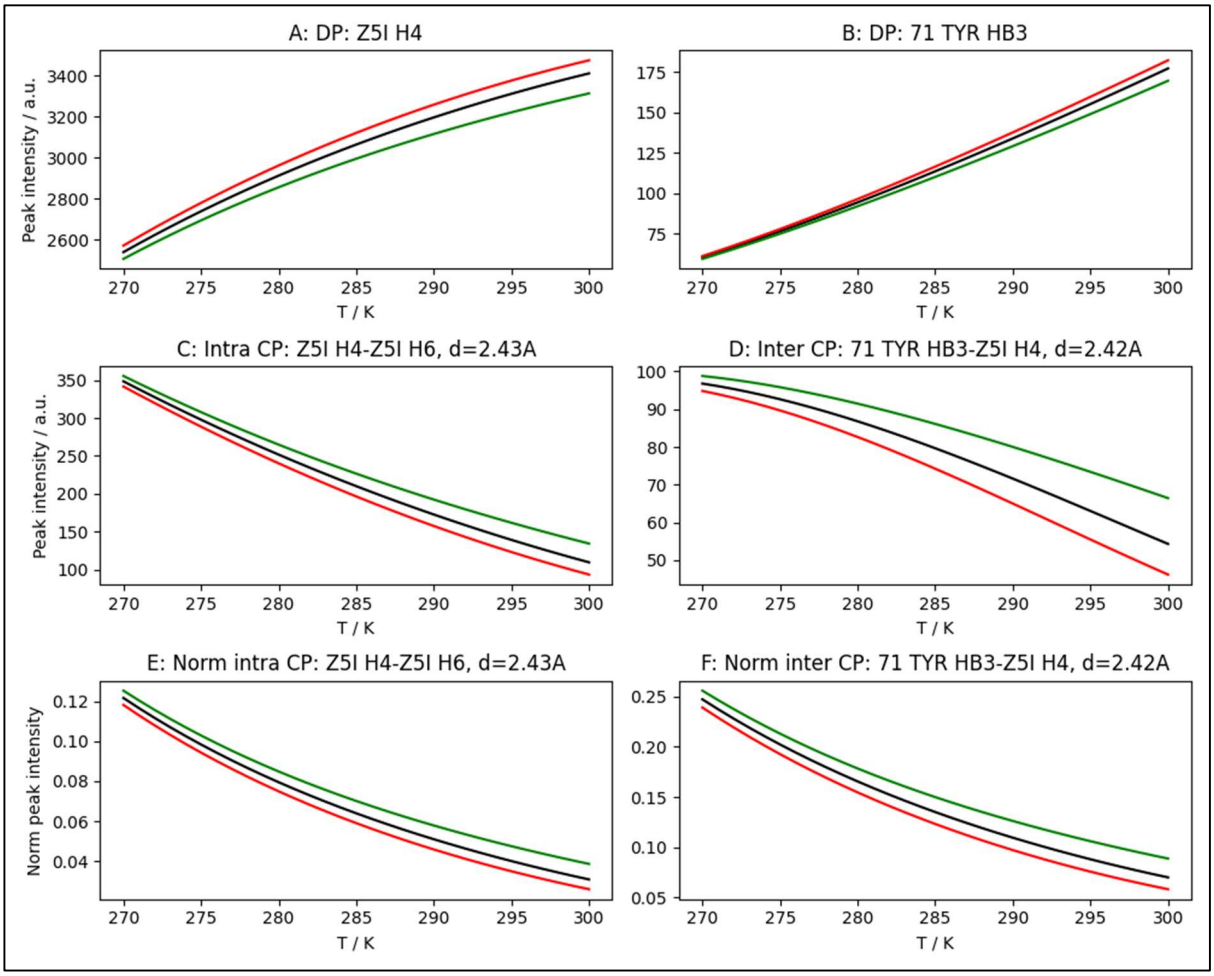

Figure 16: Temperature-dependent evolution of diagonal- and cross-peak intensities in the simulated ^{1}H-^{1}H NOESY spectrum of K-Ras G12V (GMPPNP bound) and (2E)-3-(1H-indol-2-yl)prop-2-enoic acid (Z5I). Simulation conditions: [L]=4000 µM, [P]=800 µM, $k_{off}=10^{8}\,M^{-1}s^{-1}$, ν_0=700 MHz, ΔH=-50 kJ/mol, and τ_{mix}=0.1 s. **Black** curves correspond to a K_D of 1.6 mM at 298 K, **Green** curves correspond to a K_D of 0.6 mM at 298 K, and **red** curves correspond to a K_D of 2.6 mM at 298 K. (**A**) Decay of the diagonal peak of Z5I H4. (**B**) Decay of the diagonal peak of 71 Tyr HB3. (**C**) Buildup of the intramolecular cross-peak between Z5I H4-Z5I H6 with an interproton distance of 2.43 Å. (**D**) The buildup of the intermolecular cross-peak between Z5I H4 and 71 Tyr HB3 (interproton distance of 2.42 Å). (**E**) Normalized buildup of the intramolecular cross-peak between Z5I H4 and Z5I H6 (interproton distance of 2.43 Å). (**F**) Normalized buildup of the intermolecular cross-peak between Z5I H4 and 71 Tyr HB3 (interproton distance 2.42 Å).

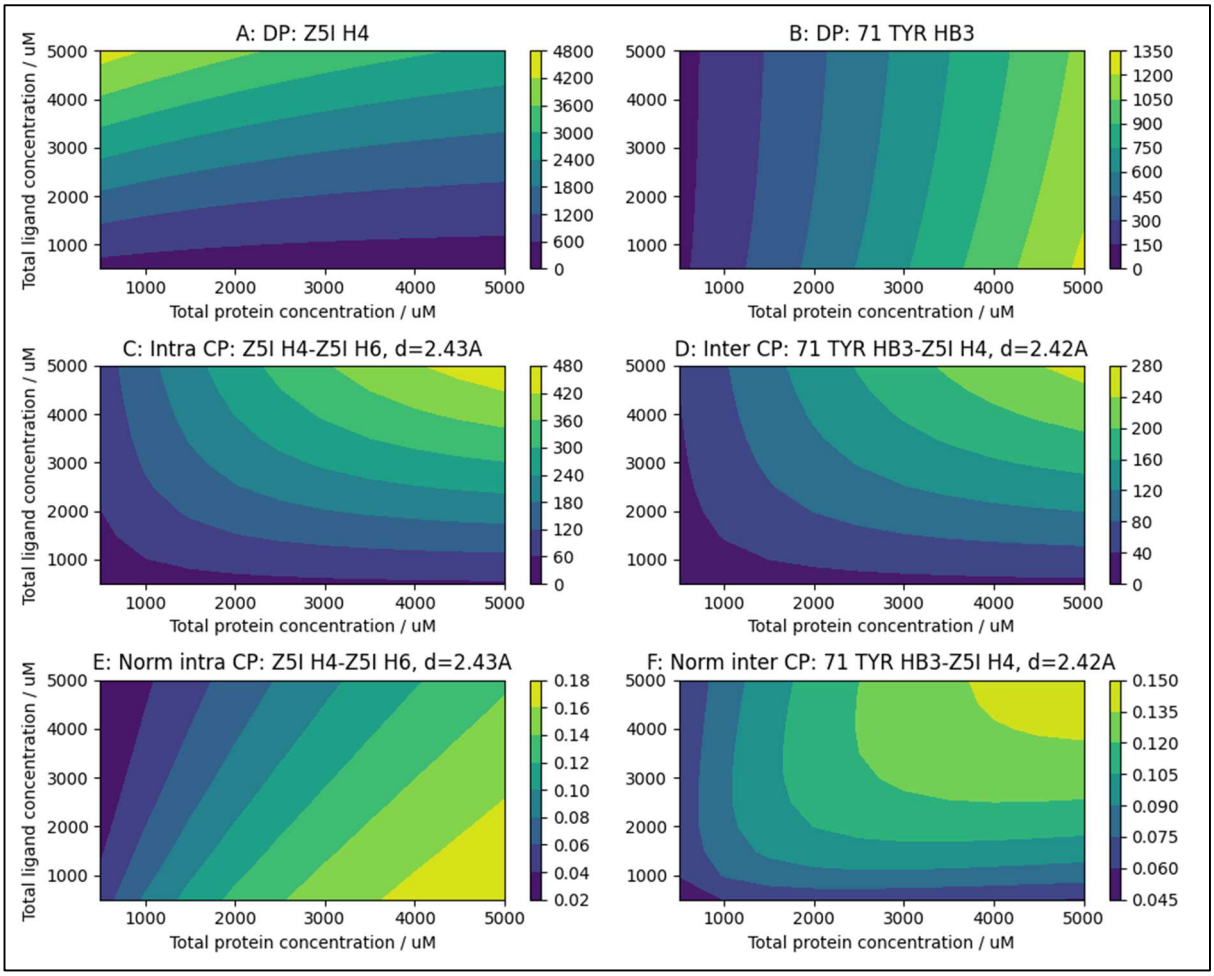

Figure 17: Concentration-dependent contour plots of diagonal- and cross-peak intensities in the simulated ^{1}H-^{1}H NOESY spectrum of K-Ras G12V (GMPPNP bound) and (2E)-3-(1H-indol-2-yl)prop-2-enoic acid (Z5I). Simulation conditions: T=298 K, $k_{off}=10^{8}\,M^{-1}s^{-1}$, ν_0=700 MHz, ΔH=-50 kJ/mol, K_D=1600 µM, and τ_{mix}=0.1 s. (**A**) Contour plot of the diagonal peak of Z5I H4. (**B**) Contour plot of the diagonal peak of 71 Tyr HB3. (**C**) Contour plot of the intramolecular cross-peak between Z5I H4-Z5I H6 with an interproton distance of 2.43 Å. (**D**) Contour plot of the intermolecular cross-peak between Z5I H4 and 71 Tyr HB3 (interproton distance of 2.42 Å). (**E**) Normalized contour plot of the intramolecular cross-peak between Z5I H4 and Z5I H6 (interproton distance of 2.43 Å). (**F**) Normalized contour plot of the intermolecular cross-peak between Z5I H4 and 71 Tyr HB3 (interproton distance 2.42 Å). The contour level depicts the intensities of the peaks in a.u. **Dark blue** corresponds to smaller numbers, and **light green** corresponds to larger numbers. A legend for the respective plot is depicted on the right of each spectrum.

2.4 Discussion

This project involved simulating a ^{1}H-^{1}H-NOESY spectrum for a protein-ligand complex (K-Ras and (2E)-3-(1H-indol-2-yl) prop-2-enoic acid (Z5I) using a full matrix approach in Python. The simulated solution matrix approximated the behavior of an experimental NOESY spectrum by plotting the evolution of diagonal- and cross-peak intensities under varying mixing times τ_{mix}, temperature T, and total protein-ligand concentrations. The influence of temperature-dependent dissociation constant K_D, protein-ligand complex concentration [PL_{eq}], dynamic viscosity μ, and rotational tumbling time τ_c was also factored into the simulation, giving insights into how each variable might affect the NOESY spectrum. The simulations were further performed with three different dissociation constants to show the effects on the evolution of peak intensities depending on the different conditions.

Peak intensities were examined with and without normalization, which helped to correct signal variations due to line shape, incomplete relaxation, and coherent transfer effects. Normalized curves, independent of diagonal peaks, helped demonstrate cross-peak buildup in relation to proton distance and cross-relaxation rates.[48] As expected, diagonal peaks decayed with increasing mixing time, while the cross-peaks initially increased before eventually decaying.[48] The normalized intermolecular cross-peak showed a faster buildup rate than the intramolecular cross-peak, suggesting a larger effective cross-relaxation rate. The cross-peak intensity's initial buildup rate seems inversely proportional to the K_D (Figure 12 D-F).

The K_D for exothermic binding reactions decreased with rising temperature, while endothermic reactions showed an opposite trend, consistent with the established principles of binding thermodynamics.[46] Temperature simulations for these reactions revealed predictable patterns; as temperature increased, dynamic viscosity and τ_c decreased proportionally due to higher kinetic energy, which lowers intermolecular friction.[45,47] Additionally, the effects of increasing K_D values on the protein-ligand complex concentration were as expected. Higher K_D values, indicating weaker binding, corresponded to a decrease in the concentration of protein-ligand complex. For the assumed exothermic binding, protein-ligand interaction likely arises from non-covalent forces, releasing energy and displacing water molecules upon binding.[49,50]

To validate these simulations, experimental NOESY spectra were obtained for the K-Ras complex with 7-fluoro-N, N-dimethyl-1-benzofuran-2-carboxamide at 280 K, 289 K, and 298 K. The comparison highlighted variations in normalized cross-peak fractions between simulated and experimental data, which could result from unknown proton-proton distances, overlapping signals, or inaccurate normalization factors. Expanding the sample size to include multiple model systems with known distance constraints could improve the accuracy of normalization and interpretation.

The observed differences between simulated and experimental fractions may stem from the abovementioned reasons. For intermolecular cross-peaks, the fraction from simulated data aligns reasonably well with the range observed in experimental data. For intramolecular cross-peaks, the fraction from simulated data is slightly larger than the range observed in

experimental data. However, this deviation remains within a plausible range when compared across the intermolecular cross-peak fractions.

Both the simulated and experimental data showed larger cross-peak intensities and smaller diagonal-peak intensities at lower temperatures. This corresponds to an increase in dynamic viscosity and the rotational tumbling time as temperature decreases. Since both proton-proton relaxation and cross-relaxation are proportional to τ_c, the relaxation rates increase with reduced temperature (Figure 13).
Also, the temperature change affected the K_D for the assumed exothermic binding reaction. The smaller K_D indicated stronger protein and ligand binding at lower temperatures, supported by the higher concentration of protein-ligand complex observed for lower temperatures (Figure 13).
These findings suggest that the increased cross-peak intensities at lower temperatures arise from enhanced cross-relaxation rates and decreased K_D, resulting in a higher concentration of the protein-ligand complex. This is further supported by the behavior of cross-peak curves for K_D values of 0.6 mM and 2.6 mM (Figure 16, D-F). The cross-peak intensities shift upward for the lower K_D (0.6 mM, green curve) and downward for the higher K_D (2.6 mM, red curve).

Experimental data is needed to validate the influence of total protein and ligand concentration on cross-peak intensity. The simulated data suggest that high protein and ligand concentrations are necessary to maximize cross-peak intensity.

This study did not investigate the relationship between proton distance and intensity variation with rising temperature. Further experiments are necessary to validate the simulated data conclusively, but they were not included in this work due to a limited timeframe.

Additional experiments to enhance analysis and validation could include conducting a concentration series, modifying viscosity independently of temperature, using different magnetic field strengths, and testing with alternative protein-ligand complexes. The script could be adapted upon validation for broader applications, such as optimizing NOESY spectrum parameters and maximizing cross peak intensities. Maximizing cross-peak intensities is especially important for the study of protein-ligand interactions and can result in more accurate extrapolation of distance restraints or the appearance of more intermolecular cross-peaks in the experimental spectrum.

It should be noted that the script requires precise coordinates for the protein-ligand system to calculate distances accurately; ideally, these coordinates are from the measured system, though a similar system can be used for approximation if needed.

2.4.1 Conclusion

This chapter highlights the efforts to simulate a 2D-[^{1}H, ^{1}H]-NOESY spectrum for a K-Ras protein-ligand complex with (2E)-3-(1H-indol-2-yl) prop-2-enoic acid (Z5I) using a full matrix approach. The simulations explored the effects of varying mixing time, temperature, and total protein and ligand concentrations and were compared to experimental data.

The mixing time simulations demonstrated the expected behavior. Diagonal peaks decrease from their maximum at τ_{mix}=0 s, while cross peaks increase to a maximum before decaying. These cross-peak maxima were inversely proportional to the K_D, confirming that the simulation aligns with the experimental expectations.

Temperature simulations showed diagonal-peak intensities increasing and cross-peak intensities decreasing with rising temperature. This behavior was attributed to changes in K_D, dynamic viscosity, and rotational correlation time, which together influenced cross-relaxation rates. The trends observed in simulated temperature-dependent spectra were consistent with experimental results for a similar protein-ligand complex.

Simulation of varying protein and ligand concentrations indicates that cross-peak intensities are maximized at high concentrations of both components, providing insights into optimal experimental conditions for future studies.

Experimental data from a temperature series of F_1-^{15}N, ^{13}C-filtered ^{1}H-^{1}H-NOESY spectra, recorded using 0.8 mM labeled ^{13}C-^{15}N K-Ras G13D bound to GMP-PNP in the presence of 4 mM 7-fluoro-N, N-dimethyl-1-benzofuran-2-carboxamide, supported the simulations. The ratio of cross-peak intensities at 280 K and 298 K for both intra- and intermolecular interactions showed comparable trends between simulated and experimental data, with minor variations potentially due to overlapping signals or inaccuracies in normalization.

In summary, the simulation successfully reproduced expected spectral behaviors and provided insights into the relationship between experimental conditions and NOESY peak intensities. Future work should further validate these simulations by conducting concentration series, isolating the effect of viscosity, and expanding the dataset with alternative protein-ligand complexes. These efforts will enhance the understanding of NOESY spectra and support the development of more accurate models for studying molecular interactions and for maximizing cross peak intensities.

Literature

1. Erdogan, A. & Rao, S. S. C. Small Intestinal Fungal Overgrowth. *Curr Gastroenterol Rep* **17**, 16 (2015).
2. Hube, B. From commensal to pathogen: stage- and tissue-specific gene expression of Candida albicans. *Current Opinion in Microbiology* **7**, 336–341 (2004).
3. Moran, C., Grussemeyer, C. A., Spalding, J. R., Benjamin, D. K. & Reed, S. D. CANDIDA ALBICANS AND NON-ALBICANS BLOODSTREAM INFECTIONS IN ADULT AND PEDIATRIC PATIENTS: COMPARISON OF MORTALITY AND COSTS. *Pediatric Infectious Disease Journal* **28**, 433–435 (2009).
4. Arendrup, M. C. Epidemiology of invasive candidiasis: *Current Opinion in Critical Care* **16**, 445–452 (2010).
5. Donlan, R. M. Biofilm Formation: A Clinically Relevant Microbiological Process. *CLIN INFECT DIS* **33**, 1387–1392 (2001).
6. Kumamoto, C. A. Candida biofilms. *Current Opinion in Microbiology* **5**, 608–611 (2002).
7. Wapinski, I., Pfeffer, A., Friedman, N. & Regev, A. Natural history and evolutionary principles of gene duplication in fungi. *Nature* **449**, 54–61 (2007).
8. Ladurner, A. G., Inouye, C., Jain, R. & Tjian, R. Bromodomains Mediate an Acetyl-Histone Encoded Antisilencing Function at Heterochromatin Boundaries. *Molecular Cell* **11**, 365–376 (2003).
9. Mietton, F. *et al.* Selective BET bromodomain inhibition as an antifungal therapeutic strategy. *Nat Commun* **8**, 15482 (2017).
10. Ferri, E., Petosa, C. & McKenna, C. E. Bromodomains: Structure, function and pharmacology of inhibition. *Biochemical Pharmacology* **106**, 1–18 (2016).

11. Dhalluin, C. *et al.* Structure and ligand of a histone acetyltransferase bromodomain. *Nature* **399**, 491–496 (1999).
12. Filippakopoulos, P. & Knapp, S. The bromodomain interaction module. *FEBS Letters* **586**, 2692–2704 (2012).
13. Filippakopoulos, P. & Knapp, S. Targeting bromodomains: epigenetic readers of lysine acetylation. *Nat Rev Drug Discov* **13**, 337–356 (2014).
14. Rees, D. C., Congreve, M., Murray, C. W. & Carr, R. Fragment-based lead discovery. *Nat Rev Drug Discov* **3**, 660–672 (2004).
15. Sawyer, T. K. New screening tools for lead compound identification. *Nat Chem Biol* **1**, 125–125 (2005).
16. Keeler, J. J. *Understanding NMR Spectroscopy*. (Wiley, Chichester, West Sussex, 2011).
17. Meiboom, S. & Gill, D. Modified Spin-Echo Method for Measuring Nuclear Relaxation Times. *Review of Scientific Instruments* **29**, 688–691 (1958).
18. Wang, L. *et al.* Recent progress in fragment-based drug discovery facilitated by NMR spectroscopy. *Magnetic Resonance Letters* **2**, 107–118 (2022).
19. Davis, B. J. & Erlanson, D. A. Learning from our mistakes: The 'unknown knowns' in fragment screening. *Bioorganic & Medicinal Chemistry Letters* **23**, 2844–2852 (2013).
20. Mayer, M. & Meyer, B. Characterization of Ligand Binding by Saturation Transfer Difference NMR Spectroscopy. *Angew. Chem. Int. Ed.* **38**, 1784–1788 (1999).
21. Meyer, B. & Peters, T. NMR Spectroscopy Techniques for Screening and Identifying Ligand Binding to Protein Receptors. *Angew Chem Int Ed* **42**, 864–890 (2003).
22. Aubin, Y., Gingras, G. & Sauvé, S. Assessment of the Three-Dimensional Structure of Recombinant Protein Therapeutics by NMR Fingerprinting: Demonstration on

Recombinant Human Granulocyte Macrophage-Colony Stimulation Factor. *Anal. Chem.* **80**, 2623–2627 (2008).

23. Williamson, M. P. Using chemical shift perturbation to characterise ligand binding. *Progress in Nuclear Magnetic Resonance Spectroscopy* **73**, 1–16 (2013).
24. Pellecchia, M., Sem, D. S. & Wüthrich, K. Nmr in drug discovery. *Nat Rev Drug Discov* **1**, 211–219 (2002).
25. Tiwari, A. & Singh, S. Computational approaches in drug designing. in *Bioinformatics* 207–217 (Elsevier, 2022). doi:10.1016/B978-0-323-89775-4.00010-9.
26. Sanders, M. P. A. *et al.* From the protein's perspective: the benefits and challenges of protein structure-based pharmacophore modeling. *Med. Chem. Commun.* **3**, 28–38 (2012).
27. Kumar, B. V. S. S. *et al.* Ligand-based and structure-based approaches in identifying ideal pharmacophore against c-Jun N-terminal kinase-3. *J Mol Model* **17**, 151–163 (2011).
28. Cox, O. B. *et al.* A poised fragment library enables rapid synthetic expansion yielding the first reported inhibitors of PHIP(2), an atypical bromodomain. *Chem. Sci.* **7**, 2322–2330 (2016).
29. Liu, M. *et al.* Improved WATERGATE Pulse Sequences for Solvent Suppression in NMR Spectroscopy. *Journal of Magnetic Resonance* **132**, 125–129 (1998).
30. Piotto, M., Saudek, V. & Sklenář, V. Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J Biomol NMR* **2**, 661–665 (1992).
31. Mayer, M. & Meyer, B. Group Epitope Mapping by Saturation Transfer Difference NMR To Identify Segments of a Ligand in Direct Contact with a Protein Receptor. *J. Am. Chem. Soc.* **123**, 6108–6117 (2001).

32. Aguilar, J. A., Nilsson, M., Bodenhausen, G. & Morris, G. A. Spin echo NMR spectra without J modulation. *Chem. Commun.* **48**, 811–813 (2012).
33. Avramovic, M. Biochemistry and biological NMR investigation of diseases relevant proteins. (2023) doi:10.25365/THESIS.72990.
34. Vom, A. *et al.* Detection and Prevention of Aggregation-based False Positives in STD-NMR-based Fragment Screening. *Aust. J. Chem.* **66**, 1518 (2013).
35. Ni, F. Recent developments in transferred NOE methods. *Progress in Nuclear Magnetic Resonance Spectroscopy* **26**, 517–606 (1994).
36. Orts, J. *et al.* NMR-Based Determination of the 3D Structure of the Ligand–Protein Interaction Site without Protein Resonance Assignment. *J. Am. Chem. Soc.* **138**, 4393–4400 (2016).
37. Orts, J., Griesinger, C. & Carlomagno, T. The INPHARMA technique for pharmacophore mapping: A theoretical guide to the method. *Journal of Magnetic Resonance* **200**, 64–73 (2009).
38. Gossert, A. D. & Jahnke, W. NMR in drug discovery: A practical guide to identification and validation of ligands interacting with biological macromolecules. *Progress in Nuclear Magnetic Resonance Spectroscopy* **97**, 82–125 (2016).
39. Van Dam, T. J. P., Bos, J. & Snel, B. Evolution of the Ras-like small GTPases and their regulators. *Small GTPases* **2**, 4–16 (2011).
40. Hofmann, M. H., Gerlach, D., Misale, S., Petronczki, M. & Kraut, N. Expanding the Reach of Precision Oncology by Drugging All *KRAS* Mutants. *Cancer Discovery* **12**, 924–937 (2022).
41. Breeze, A. L. Isotope-filtered NMR methods for the study of biomolecular structure and interactions. *Progress in Nuclear Magnetic Resonance Spectroscopy* **36**, 323–372 (2000).

42. Zwahlen, C. *et al.* Methods for Measurement of Intermolecular NOEs by Multinuclear NMR Spectroscopy: Application to a Bacteriophage λ N-Peptide/ *boxB* RNA Complex. *J. Am. Chem. Soc.* **119**, 6711–6721 (1997).
43. Ogura, K., Terasawa, H. & Inagaki, F. An improved double-tuned and isotope-filtered pulse scheme based on a pulsed field gradient and a wide-band inversion shaped pulse. *J Biomol NMR* **8**, (1996).
44. Iwahara, J., Wojciak, J. M. & Clubb, R. T. Improved NMR spectra of a protein-DNA complex through rational mutagenesis and the application of a sensitivity optimized isotope-filtered NOESY experiment. *Journal of Biomolecular NMR* **19**, 231–241 (2001).
45. Cavanagh, J. *Protein NMR Spectroscopy: Principles and Practice*. (Academic Press, Amsterdam Boston, 2007).
46. Sakamoto, S. *et al.* Contributions of Interfacial Residues of Human Interleukin15 to the Specificity and Affinity for Its Private α-Receptor. *Journal of Molecular Biology* **389**, 880–894 (2009).
47. Kestin, J., Sokolov, M. & Wakeham, W. A. Viscosity of liquid water in the range −8 °C to 150 °C. *Journal of Physical and Chemical Reference Data* **7**, 941–948 (1978).
48. Nichols, P. *et al.* The Exact Nuclear Overhauser Enhancement: Recent Advances. *Molecules* **22**, 1176 (2017).
49. Olsson, T. S. G., Williams, M. A., Pitt, W. R. & Ladbury, J. E. The Thermodynamics of Protein–Ligand Interaction and Solvation: Insights for Ligand Design. *Journal of Molecular Biology* **384**, 1002–1017 (2008).
50. Reynolds, C. H. & Holloway, M. K. Thermodynamics of Ligand Binding and Efficiency. *ACS Med. Chem. Lett.* **2**, 433–437 (2011).

Christophe Dhalluin, J. E.-M.-M. (1999). Structure and ligand of a histone acetyltransferase bromodomain. *Nature 399, 491–496*, https://doi.org/10.1038/20974.

Supplementary Information

Buffers used for expression and purification

Table S3: M9 expression medium

M9 expression medium – 1.5L	
Compound	**Concentration**
10x M9 salts	150 mL
$^{15}NH_4Cl$	1.5 g
Glucose	9 g
1M $MgSO_4$	1.5 mL
1M $CaCl_2$	0.45 mL
Antibiotics	1.5 mL
100x Trace elements	15 mL
ddH_2O	To 1.5 L

Table S4: 10x M9 salts

10x M9 salts – 1L	
Compound	**Concentration**
KH_2PO_4	30 g
Na_2HPO_4	68 g
NaCl	5 g
ddH_2O	To 1 L

Table S5: 100x Trace elements

100x Trace elements – 1L	
Compound	**Concentration**
EDTA	5 g
$FeCl_3 – 6H_2O$	830 mg
$ZnCl_2$	84 mg
0.1M $CuSO_4 – 5H_2O$	765 µL
0.2M $CoCl_2 – 6H_2O$	210 µL
0.1M H_3BO_3	1.6 mL
1M $MnCl_2$	8.1 µL
ddH_2O	To 1 L

Table S6: Buffer A used for the purification of CaBdf1 BD2 (adapted from Mietton *et al.* (2017)[9])

Buffer A pH 7.4	
Compound	**Concentration**
Tris	50 mM
NaCl	500 mM
β-Mercaptoethanol	5 mM

Table S7: Buffer B used for the purification of CaBdf1 BD2 (adapted from Mietton *et al.* (2017)[9])

Buffer B pH 7.4	
Compound	**Concentration**
Tris	50 mM
NaCl	500 mM
β-Mercaptoethanol	5 mM
Imidazole	300 mM

Table S8: Buffer used for the size exclusion chromatography of CaBdf1 BD2 (adapted from Avramovic (2023)[33])

SEC buffer pH 7.4	
Compound	**Concentration**
Na_2HPO_4/NaH_2PO_4	50 mM
NaCl	50 mM

Table S9: Lysis buffer for CaBdf1 BD2 (adapted from Mietton *et al.* (2017)[9])

Lysis buffer pH 7.4	
Compound	**Concentration**
Tris	50 mM
NaCl	500 mM
β-Mercaptoethanol	5 mM
Glycerol	10%
Protease inhibitor	One tablet/50 mL buffer
Benzonase	1 µL/50 mL buffer

Compositions of screening boxes

Table S10: Compositions of Screening Box 1

Box 1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
A	1A2 2A2 3A2			1A5 4A5		2A7 3A7 5A7	2A8 5A8	2A9 4A9	1A10 2A10 3A10	1A11 2A11 3A11
B	4B2 5B2	1B3 2B3 3B3		1B5 2B5 4B5	1B6 2B6 3B6	2B7 5B7	2B8 3B8 4B8	2B9 3B9 5B9	3B10 4B10 5B10	1B11 2B11 3B11
C	2C2 3C2 4C2	1C3 3C3	4C4 5C4		1C6 2C6 3C6	1C7 3C7 4C7	1C8 2C8 3C8	2C9 5C9	1C10 2C10 3C10	1C11 3C11
D	2D2 5D2	1D3 3D3	2D4 3D4 4D4	1D5 2D5 (3D5)	2D6 4D6		1D8 3D8 4D8	2D9 3D9 4D9	3D10 (4D10) 5D10	2D11 4D11
E		2E3 4E3 5E3	1E4 2E4 3E4	1E5 2E5 3E5	3E6 4E6	1E7 2E7 5E7	1E8 2E8 3E8	1E9 2E9 4E9	1E10 2E10 5E10	1E11 3E11

F	2F2 3F2 4F2		4F4 5F4	2F5 3F5 4F5	4F6 5F6	1F7 2F7	2F8 3F8 4F8	2F9 3F9 4F9		1F11 2F11 3F11
G	2G2 3G2 4G2	1G3 5G3	1G4 2G4 4G4		3G6 4G6	1G7 2G7 3G7		1G9 2G9	3G10 5G10	1G11 3G11
H	1H2 3H2 4H2		2H4 4H4	1H5 2H5 5H5	2H6 5H6		1H8 2H8 3H8	1H9 3H9 4H9	1H10 2H10 3H10	1H11 3H11 5H11

Table S11: Composition of screening Box 2

Box 2	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
A	4A2 5A2 8A2	3A3 6A3 7A3	2A4 7A4 10A4	10A5 11A5		9A7 10A7 11A7	6A8 11A8	6A9 7A9 8A9	5A10 6A10 10A10	5A11 7A11 8A11
B	6B2 8B2 11B2	4B3 5B3 6B3	6B4 7B4 8B4	7B5 8B5 10B5	8B6 10B6		5B8 8B8 11B8	6B9 7B9 8B9	6B10 7B10 8B10	7B11 8B11 11B11
C	6C2 9C2 12C2	6C3 7C3	6C4 7C4 8C4	3C5 7C5 9C5	6C6 10C6 11C6	7C7 8C7 9C7	4C8 5C8 6C8	6C9 7C9 9C9	6C10 7C10 9C10	7C11 8C11 11C11
D	7D2 8D2	6D3 10D3 11D3	7D4 9D4 10D4	2A6 6D5 6G6	6D6 9D6 10D6	1D7 9D7	5D8 6D8 7D8	5D9 6D9 7D9	7D10 8D10 9D10	7D11 9D11 10D11
E	2E2 7E2 8E2	6E3 7E3 8E3	5E4 6E4 8E4	5E5 7E5 10E5	6E6 7E6 8E6	6E7 7E7 9E7	4E8 5E8 6E8	5E9 6E9 7E9	7E10 9E10 10E10	7E11 8E11 9E11
F	7F2 8F2 12F2	2F3 3F3 4F3	7F4 10F4 11F4	6F5 7F5 10F5	6F6 8F6 11F6	6F7 7F7 8F7	5F8 6F8 7F8	6F9 7F9 10F9	6F10 8F10 9F10	7F11 9F11 10F11
G	7G2 9G2 12G2	6G3 7G3 8G3	6G4 10G4	1G5 6G5 10G5		5G7 10G7 11G7	2G8 6G8 7G8	6G9 7G9 8G9	7G10 10G10	6G11 9G11 10G11
H	6H2 10H2 12H2	2H3 6H3 7H3	6H4 7H4 10H4	8H5 9H5 10H5	6H6 9H6 10H6	1H7 6H7 7H7	5H8 6H8 7H8	5H9 6H9 10H9	5H10 7H10 10H10	8H11 9H11 10H11

Table S12: Composition of screening Box 3

Box 3	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
A	9A2 11A2 12A2	10A3 11A3 12A3								
B		7B3 8B3 10B3	9B4 10B4 11B4					9B10 10B10 11B10		

C			10C4 11C4 11A4	10C5 11C5 11F5		10B7 10C7 11C7	10B9 10C8 11C8			
D		11B3 12D3 12F3					9D8 10D8 10G8			
E	10E2 11D6 12E2	9E6 11E3 12E3	9E4 10E4 11D4				7E8 8E8 10E8		10D10 10E11 11D11	
F		5F3 7F3 10F3				9F7 10F7 10E7	9F8 10F8 11F8	10F10 11F10 11H10		
G		10G3 11G3 12G3								
H		10H3 11H3 12H3				8H7 9H7 10H7	8H8 9H8 11H8			

Summary of Screening analysis

Table S13: Tubes/Ligands showing large CSPs

Ligands with significant CSP	Number of tubes	Number of ligands
Total	58	166
W/o STD and CPMG signals	14	39
Only CPMG signals	13	37
Only STD signals	12	34
STD and CPMG signals	19	56

Table S14: Tubes/Ligands showing small CSPs

Ligands with insignificant CSP	Number of tubes	Number of ligands
Total	84	234
W/o STD and CPMG signals	38	105
Only CPMG signals	16	44
Only STD signals	12	33
STD and CPMG signals	18	52

Table 15: Tubes/Ligands showing no CSPs

Ligands w/o CSP	Number of tubes	Number of ligands
Total	30	81
W/o STD and CPMG signals	14	37
Only CPMG signals	7	19
Only STD signals	4	11

STD and CPMG signals	5	14

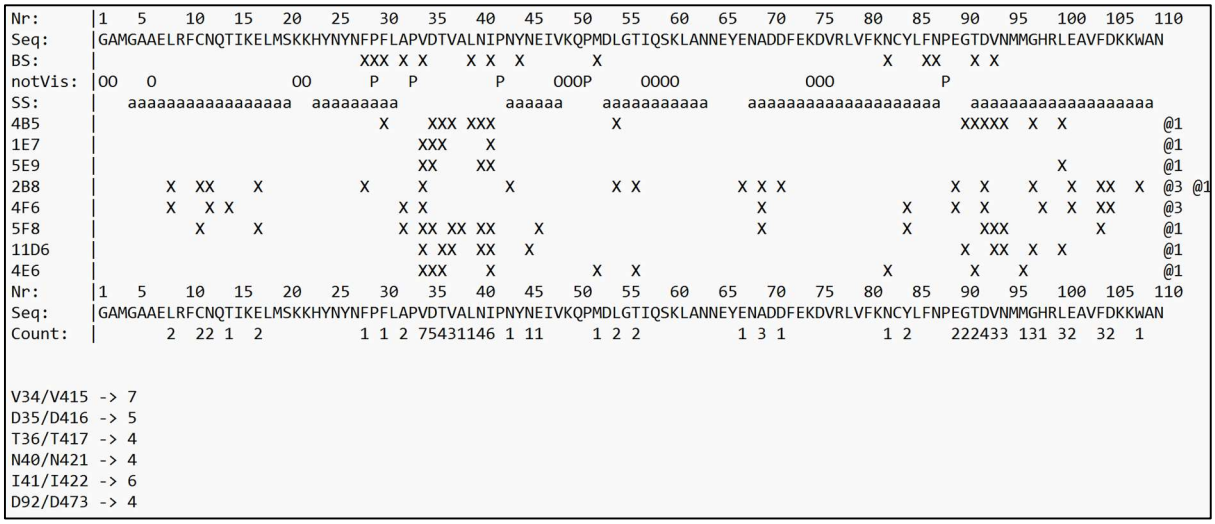

Figure S18: CSP distribution of ligands 4B5, 1E7, 5E9, 2B8, 4F6, 5F8, 11D8 and 45E, @X means the CSP of an unassigned residue. Nr: Number of primary sequence; Seq: Primary sequence of CaBdf1 BD2; BS: Residues involved in the binding of the ligand from the published structure (PDB code 5N18); noVis: 0 marks the position of not assigned residues in the HSQC spectrum, P marks the positions of prolines; SS: Secondary structure elements, a marks alpha helices; Count: Counts the times the specific residue is shifted across all depicted ligands.

Raw data of the NOESY temperature series

Table S16: Raw data of NOESY spectrum measured at 280 K

#(280)	Pos F1	Pos F2	LW F1 (Hz)	LW F2 (Hz)	Height	Comment
inter CP1	0.455655	6.973648	35.25362	85.98359	2.08E+09	1->6
inter CP2	0.330367	3.069735	35.11373	64.4477	1.31E+09	1->3
inter CP3	1.164818	5.76535	37.25682	71.80179	9.5E+08	2->5
DP5	5.768753	5.764146	34.38593	69.37033	4.56E+10	Ligand
DP6	6.975222	6.978028	0	68.92317	2.25E+11	Ligand
DP3	3.072157	3.071593	17.3875	62.04168	2.33E+12	Ligand
intra CP1	4.262595	5.767783	41.48703	69.91967	3.67E+09	4->5
intra CP2	4.26385	7.709275	38.65427	67.24956	2.97E+09	4->7
DP7	7.704097	7.710281	29.29165	76.67186	7.09E+09	Ligand
intra CP3	5.765928	7.709729	32.50036	68.5025	1.15E+09	5->7
DP4	4.263781	4.263839	None	None	1.15E+10	Ligand
DP2	1.164896	1.165225	None	None	6.91E+09	Protein

DP1	0.329385	0.330047	None	None	5.76E+09	Protein
inter CP4	1.155279	4.250494	None	None	1.52E+09	2->4
inter CP5	1.152183	7.712891	None	None	2.38E+09	2->7

Table S17: Raw data of NOESY spectrum measured at 289 K

#(289)	Pos F1	Pos F2	LW F1 (Hz)	LW F2 (Hz)	Height	Comment
inter CP1	0.566108555	7.092671575	52.10397	65.711106	1.69E+09	1->6
inter CP2	0.420776354	3.186216185	29.14681	44.5088329	9.93E+08	1->3
inter CP3	1.249033289	5.870457523	0	79.042088	7.7E+08	2->5
DP5	5.869905202	5.870050362	30.47115	65.222585	6.77E+10	Ligand
DP6	7.094895789	7.097893816	0	70.4409523	3.1E+11	Ligand
DP3	3.182382122	3.180759116	21.59625	60.7051737	1.96E+12	Ligand
intra CP1	4.348040023	5.869474399	48.74865	69.0948187	4.18E+09	4->5
intra CP2	4.34946812	7.811556802	40.43988	65.3277229	5.19E+09	4->7
DP7	7.809428034	7.81222508	26.43184	65.8640478	1.38E+10	Ligand
intra CP3	5.870795881	7.812137807	27.06629	63.8307636	1.62E+09	5->7
DP4	4.34948293	4.349497204	None	None	2.06E+10	Ligand
DP2	1.248915056	1.248750495	None	None	8.44E+09	Protein
DP1	0.420447813	0.420518316	None	None	8.17E+09	Protein
inter CP4	1.245475191	4.349041018	None	None	2.01E+09	2->4
inter CP5	1.244273141	7.809349328	None	None	3.99E+09	2->7

Table S18: Raw data of NOESY spectrum measured at 298 K

#(298)	Pos F1	Pos F2	LW F1 (Hz)	LW F2 (Hz)	Height	Comment
inter CP1	0.662409	7.19367899	32.8249395	72.87798884	1.4E+09	1->6
inter CP2	0.530376	3.278330828	365.8399607	49.70980129	8.64E+08	1->3
inter CP3	1.336905	5.970875251	21.98904015	62.38124367	7.04E+08	2->5
DP5	5.971773	5.972567197	20.15660721	62.4672428	8.93E+10	Ligand
DP6	7.196437	7.196497379	0	64.73796703	4.75E+11	Ligand
DP3	3.283277	3.281163612	15.07675794	63.44138104	3.31E+12	Ligand
intra CP1	4.447465	5.972455616	25.10159538	63.60505996	4.27E+09	4->5
intra CP2	4.446271	7.907178891	26.95361521	62.99727426	6.77E+09	4->7
DP7	7.906241	7.909524667	19.05414683	63.72513193	2.69E+10	Ligand

intra CP3	5.971358	7.907407318	18.14255908	62.40547673	1.85E+09	5->7
DP4	4.446234	4.446257173	None	None	2.5E+10	Ligand
DP2	1.336727	1.336862313	None	None	1.49E+10	Protein
DP1	0.530082	0.530323716	None	None	1.23E+10	Protein
inter CP4	1.336547	4.441791105	None	None	2.17E+09	2->4
inter CP5	1.337529	7.907737908	None	None	5.44E+09	2->7

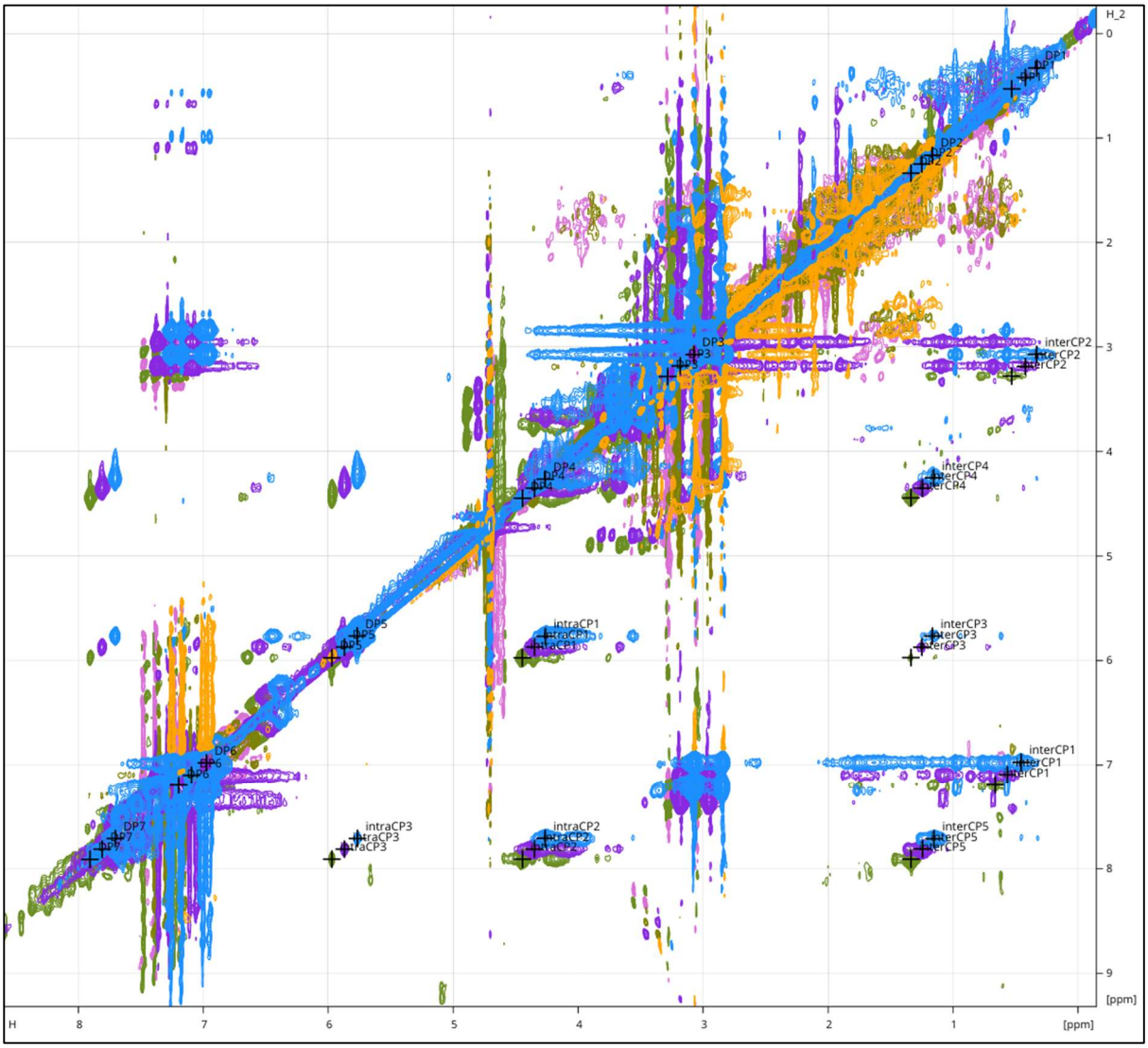

Figure S19: Overlay of F_1-^{15}N, ^{13}C-filtered ^{1}H-^{1}H-NOESY spectra at three temperatures: 280 K (**blue**), 289 K (**purple**), and 298 K (**green**), for 0.8 mM ^{13}C, ^{15}N labeled K-Ras G13D (GMP-PNP bound) in the presence of 4 mM 7-fluoro-N, N-dimethyl-1-benzofuran-2-carboxamide. Spectra were acquired on a 700 MHz cryo platform in 30 mM deuterated NaPi buffer with 200 mM NaCl and 2 mM TCEP at pD 7.4. Abbrev: IntraCP: Intramolecular cross-peak; InterCP: Intermolecular cross-peak; DP: Diagonal-peak.