



UNIVERSITAT DE
BARCELONA

Biophysical Techniques for the Investigation of Protein-Ligand Complexes

Salvatore Scaffidi

ADVERTIMENT. La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX (www.tdx.cat) i a través del Dipòsit Digital de la UB (diposit.ub.edu) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX ni al Dipòsit Digital de la UB. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX o al Dipòsit Digital de la UB (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

ADVERTENCIA. La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR (www.tdx.cat) y a través del Repositorio Digital de la UB (diposit.ub.edu) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR o al Repositorio Digital de la UB. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR o al Repositorio Digital de la UB (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

WARNING. On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX (www.tdx.cat) service and by the UB Digital Repository (diposit.ub.edu) has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized nor its spreading and availability from a site foreign to the TDX service or to the UB Digital Repository. Introducing its content in a window or frame foreign to the TDX service or to the UB Digital Repository is not authorized (framing). Those rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author.

Universitat de Barcelona
Facultat de Farmàcia i Ciències de l'Alimentació

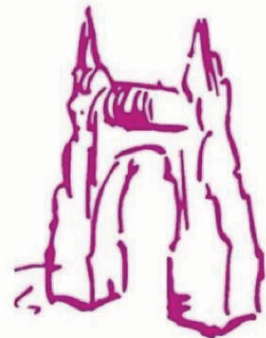
Biophysical Techniques for the Investigation of Protein-Ligand Complexes

Salvatore Scaffidi

2021



UNIVERSITAT DE
BARCELONA



FACULTAT DE FARMÀCIA I
CIÈNCIES DE L'ALIMENTACIÓ

Programa de Doctorat en Biomedicina
Universitat de Barcelona

Directors de Tesi: Dr. Carles Galdeano Cantador i Prof. Xavier
Barril Alonso

UNIVERSITAT DE BARCELONA

FACULTAT DE FARMÀCIA I CIÈNCIES DE L'ALIMENTACIÓ

DEPARTAMENT DE FARMÀCIA I TECNOLOGIA FARMACÈUTICA
I FÍSICOQUÍMICA

PROGRAMA DE DOCTORAT EN BIOMEDICINA

BIOPHYSICAL TECHNIQUES FOR THE INVESTIGATION OF
PROTEIN-LIGAND COMPLEXES

Aquesta tesi ha estat realitzada per Salvatore Scaffidi sota la direcció del Dr. Carles Galdeano Cantador, Professor Lector Serra Hunter i el Prof. Xavier Barril Alonso, Professor d'Investigació ICREA, tots dos del Departament de Farmàcia i Tecnologia Farmacèutica i Físicoquímica de la Facultat de Farmàcia i Ciències de l'Alimentació de la Universitat de Barcelona. Es presenta aquesta memòria per optar al títol de doctor per la Universitat de Barcelona en el Programa de Doctorat en Biomedicina.

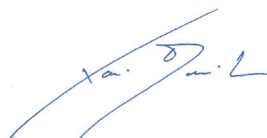
Carles Galdeano Cantador

Director de tesi




Xavier Barril Alonso

Director de tesi



Salvatore Scaffidi

Doctorand



Salvatore Scaffidi

January 2021

*Alla mia Famiglia che ha sempre creduto in me,
Grazie di cuore.*

Acknowledgements

I'd like to thank, first of all, Carles, who has always guided me in the most important decisions of my career from the first moment I met him in Dundee. He has inspired me and has transmitted his passion for science, teaching me how to fight for my dreams and never give up, but most important to be resilient! Then it comes Xavi, that is always there for you, helping you in no matter what! He is not only a great and brilliant scientist to be proud to meet, but mostly, a very good person to rely on. I appreciate the time that he has spent to answer my questions whenever I have needed and I'm really grateful for the wise advices he has given to me during these 3 years.

I'd also like to thank my partners in crime, Miriam and Andrea, that support me every day and with whom I have spent incredible moments in the lab that I'll never forget. I also thank Moira, Serena and Maciej for the amazing time we had in Barcelona together; crazy and funny indelible memories. With that I'd like to thank Alex, Alvaro, Dylan, Juan, Marina, Guillermo, Roger and my students Miriam C. and Marta for the good times together! I'd like also to acknowledge the people from the Medicinal Chemistry Lab at the Faculty, in particular Dr Carmen Escolano and Sergio, who have helped me during the time spent with them. I specially thank Marta Taulés for being a great and wonderful support for the SPR experiments and for becoming a friend after this time. A big thanks also to Marga and Teresa for being so patient and teaching me how to understand the NMR experiments. A special thanks goes to Dani that has helped me, especially in the most stressful moments, and motivated me every single time I wanted to give up. Thanks also to my flatmates, Arianna, Monica and Cristina for listening my every day complains and for giving me good advices, and most of all for all the drinks and parties we had during the quarantine. Finally, from the bottom of my heart I thank my sister, my aunt Pina and my grandparents, that always

supported me and that gave me the possibility to follow my dreams. Specially I thank you, Nonno, that you were always smiling just knowing that I was happy. From there, I know, together with my parents, you are always by my side!

The experimental work of this Thesis was performed from April 2017 to July 2020 and was supported with funds from the *Ministerio de Economía y Competitividad* (SAF2015-68749-R, RTI2018-096429-B-I00) and *Agència de Gestió d'Ajuts Universitaris i de Recerca* (AGAUR, 2017SGR1746). For this PhD thesis, Salvatore Scaffidi received an APIF PhD fellowship from *Universitat de Barcelona*.

The experimental work was carried out at the Barril & Galdeano's Lab at the *Departament de Farmàcia i Tecnologia Farmacèutica i Fisicoquímica de la Facultat de Farmàcia i Ciències de l'Alimentació de la Universitat de Barcelona*.

Abbreviations

ABBREVIATION LIST

ADME: Absorption Distribution Metabolism and Excretion	K_d: dissociation constant
AIDS: Acquired ImmunoDeficiency Syndrome	MALDI: Matrix-assisted Laser Desorption/Ionization
ATR: Attenuated Total Reflectance	mCRPC: metastatic Castration Resistant Prostate Cancer
BCPs: Bromodomain-Containing Proteins	MD: Molecular Dynamic
BSA : Bovine Serum Albumin	MoA: Mechanism of Action
CETSA: Cellular Thermal Shift Assay	MS: Mass Spectrometry
COSMIC: Catalogue Of Somatic Mutation in Cancer	MST: MicroScale Thermophoresis
cryo-EM: cryo-Electron Microscopy	MW: Molecular Weight
CSP: Chemical Shift Perturbation	Ni-NTA: Nickel-Nitrilotriacetic acidic
DSF: Differential Scanning Fluorimetry	NMR: Nuclear Magnetic Resonance
e.g.: exempli gratia	PAINS: Pan-Assay Interference compoundS
FBDD: Fragment-Based Drug Discovery	PBS: Phosphate Buffer Saline
FDA: Food and Drug Administration	PCR: Polymerase Chain reaction
FP: Fluorescence Polarization	POI: Protein Of Interest
FRET: Förster Resonance Energy Transfer	PPI: Protein-Protein Interaction
GPCR: G Protein-Coupled Receptor	PROTAC: PROteolysis TArgeting Chimera
GST: Glutathione-S-Transferase	RTL: Retinol
HIV: Human Immunodeficiency Virus	RU: Resonance Unit
HRMS: High-Resolution Mass-Spectrometry	SBDD: Structure-Based Drug Design
HSQC: Heteronuclear Single-Quantum Correlation	SFX: Serial Femtosecond Crystallography
HTS: High-Throughput Screening	SPR: Surface Plasmon Resonance
i.e.: id est	STD: Saturation Transfer Difference
IC₅₀: inhibitory concentration	TR-FRET: Time Resolved- FRET
IMiDs: ImmunoModulatory imide Drugs	UPS: Ubiquitin Proteasome System
IR: InfraRed	XFELs: X-ray Free Electron Lasers
ITC: Isothermal Titration Calorimetry	

Table of contents

Table of Contents

Chapter 1	1
Introduction	1
1.1 Preface	3
1.2 A historical view of the biophysical techniques:fromtheresolutionofthefirstcrystal structure to2020.....	5
1.3 Biophysical techniques in the drug discovery process	8
1.3.1 Differential Scanning Fluorimetry (DSF) assay	9
1.3.2 Surface Plasmon Resonance (SPR).....	12
1.3.3 Fluorescence Polarization (FP) assay.....	15
1.3.4 Nuclear Magnetic Resonance (NMR) spectroscopy.....	17
1.3.4.1 Protein-Observed NMR experiments: chemical shift perturbations	18
1.3.4.2 Ligand-Observed NMR experiments: Saturation Transfer Difference (STD)	20
1.3.5 Förster Resonance Energy Transfer (FRET)	22
1.3.5.1 Time-Resolved FRET (TR-FRET)	23
1.4 Introduction to fragment screening.....	25
1.4.1 The fragment library	27
1.4.2 Fragment screening	28
1.4.3 Fragment optimization.....	31
Chapter 2	35
Objectives	35
2.1 General Objective.....	37
2.2 Specific objectives.....	37
Chapter 3	39
Methods andMaterials	39
3.1 General Materials.....	41
3.1.1 Reagents	41
3.1.2 Bacterial Strains and Growth Media	41
3.2 Molecular Cloning	41
3.2.1 Protein Constructs and Expression Plasmids	41
3.2.2 Kits	42
3.2.2.1 Site-directed mutagenesis (CRBP project).....	42

Table of contents

3.3	Preparation of Competent E.coli cells.....	43
3.4	Transformation of E.coli cells	43
3.5	Determination of DNA concentration, protein concentration and cell growth	44
3.6	Preparation of M9 medium broth (Fbw7-Skp1 project)	44
3.7	Protein Expression.....	45
3.7.1	Fbw7-Skp1 complex.....	45
3.7.2	Fbw7-Skp1 complex ¹⁵ N labelled-M9 Marley expression	45
3.7.3	Rat-CRBP proteins.....	46
3.7.4	Brd4(BD1) protein.....	46
3.8	Cell lysis.....	47
3.9	Purification	48
3.9.1	Fbw7-Skp1 complex and Fbw7-Skp1 ¹⁵ N labelled.....	48
3.9.2	Rat-CRBPs proteins	49
3.9.3	Brd4(BD1)	49
3.10	Fragment Library.....	50
3.11	Biophysical Techniques	51
3.11.1	DSF experiments	51
3.11.2	SPR experiments	52
3.11.2.1	Fbw7-Spk1 SPR screening	53
3.11.2.2	Fragment hit characterization	55
3.11.3	Retinol Kinetic affinity characterization	56
3.11.4	NMR experiments.....	56
3.11.4.1	1D-NMR of Fbw7-Skp1 complex	57
3.11.4.2	Saturation transfer difference (STD).....	57
3.11.5	Fluorescence Polarization (FP) assay.....	58
3.11.6	TR-FRET assay.....	59
3.12	Synthesis of Brd4(BD1) compounds	61
3.12.1	2-Chloro-N'-(4-phenylthiazol-2-yl)acetohydrazide (Pre-SSR1)	62
3.12.2	3-(Chloromethyl)-5-phenylthiazolo[2,3-c] [1,2,4]triazole (SSR1).....	62
3.12.3	3-(Hydroxymethyl)-5-phenylthiazolo[2,3-c] [1,2,4]triazole (SSR-2)	63
3.12.4	N'-(4-phenylthiazol-2-yl)acetohydrazide (Pre-SSR3)	64
3.12.5	3-methyl-5-phenylthiazolo[2,3-c][1,2,4]triazole (SSR-3)	65
3.12.6	N'-(4-phenylthiazol-2-yl)propionohydrazide (Pre-SSR4).....	66
3.12.7	3-Ethyl-5-phenylthiazolo[2,3-c][1,2,4]triazole (SSR-4).....	67
Chapter 4.....		69
Results.....		69

Table of contents

4.1	Identify and confirm the first hits binding to Fbw7-Skp1 complex.....	71
4.1.1	Background: The Ubiquitin Proteasome System (UPS).....	71
4.1.2	Hijacking the UPS: PROTAC technology	73
4.1.3	Fbw7: a challenging attractive target in drug discovery	76
4.1.3.1	Fbw7-Skp1 structure	77
4.1.3.2	Fbw7-Skp1 substrate recognition and mutations.....	78
4.2	Objective of the Chapter 4.....	81
4.3	Cloning, expression, purification and characterization of Fbw7-Skp1 complex	83
4.3.1	Cloning and expression of Fbw7-Skp1 complex in E.coli.....	83
4.3	Purification and characterization of Fbw7-Skp1 complex	84
4.4	Fragment Screening.....	86
4.4.1	DSF	86
4.4.2	SPR experiments	88
4.4.2.1	Dose response curve assessed by SPR: Fbw7 linked by amine coupling reaction	92
4.4.2.2	Dose response curve assessed by SPR: Fbw7 immobilization with a GST tag.....	96
4.4.3	STD-NMR to confirm binding	102
4.4.4	Elucidation of the binding site of the fragments.....	107
4.4.4.1	X-ray crystallography	107
4.4.4.2	MD simulations	108
4.4.4.3	2D-NMR (CSP)	109
4.4.4.3.1	Cloning, expression and purification of Fbw7-Skp1 complex ¹⁵ N labelled in E.coli.	110
4.5	Mechanism of action elucidation	111
4.5.1	Fluorescence polarization Assay	111
4.6	Folding and stability of Fbw7-Skp1 complex.....	115
4.7	Summary and future perspective	116
Chapter 5.....		123
Results.....		123
5.1	Examine the thermodynamics and kinetics of the retinol binding to Cellular Retinol Binding proteins (CRBPs).....	125
5.1.1	Background: Cellular retinol binding protein	125
5.2	Objective of the Chapter 5.....	127
5.3	Cloning, expression, purification and characterization of CRBPs	129
5.3.1	Site-directed mutagenesis	129
5.4	Purification and characterization of CRBPs	130
5.5	SPR: titration of retinol.....	130
5.6	Summary	132

Table of contents

Chapter 6	135
Results	135
6.1 Automated fragment evolution platform: Validating small-molecule binding to Brd4(BD1).....	137
6.1.1 Background	137
6.2 Objective of the Chapter 6.....	141
6.3 Synthesis of SSR3, SSR4, SSR1 and SSR2	143
6.4 Cloning, expression and purification of Brd4(BD1) in E.coli	145
6.5 Time-resolved Fluorescence Resonance Energy Transfer (TR-FRET) experiment	146
6.6 Summary	153
Chapter 7	157
Discussions	157
7.1 General discussion	159
Chapter 8	165
Conclusions	165
8.1 Fbw7-Skp1 project (Chapter 4)	167
8.2 CRBPs project (Chapter 5).....	168
8.3 Brd4(BD1) project (Chapter 6)	168
Annex	171
Bibliography	181

List of figures

<u>Chapter 1</u>	1
Figure 1.1 Flowchart of a possible strategy for compound screening, validation, and characterization using biophysical techniques.....	9
Figure 1.2 Examination of differences in the temperature dependent fluorescence profile in DSF.....	10
Figure 1.3 Determination of the melting Temperature in DSF.....	11
Figure 1.4 SPR principles.....	13
Figure 1.5 Example of on/off kinetic rate in SPR and correlation with K_d	14
Figure 1.6 FP principles.....	16
Figure 1.7 HSQC-NMR spectra.....	19
Figure 1.8 STD-NMR principles.....	21
Figure 1.9 FRET principles.....	23
Figure 1.10 TR-FRET principles.....	24
Figure 1.11 Example of a FBDD screening cascade.....	30
Figure 1.12 FBDD optimization strategies.....	31
Figure 1.13 Venetoclax development	32
<u>Chapter 3</u>	39
Figure 3.1 Principle moment of inertia of the fragment library.....	51
Figure 3.2 Formula to calculate the amount of protein to immobilize in a sensor chip.....	53
<u>Chapter 4</u>	69
Figure 4.1 The ubiquitin proteasome system.....	73
Figure 4.2 PROTAC molecules hijack the ubiquitin proteasome system.....	76
Figure 4.3 Overall structure of the ternary Fbw7-Skp1-C-terminal Cyclin E peptide complex.....	78
Figure 4.4 Conserved CPD consensus motif found in all known Fbw7 substrates.....	79
Figure 4.5 The distribution of Fbw7 mutations in the FBXW7 genome.....	79
Figure 4.6 FBXW7 gene mutation frequency for different human cancer types.....	80
Figure 4.7 Akta chromatograms of Fbw7-Skp1 purification.....	85

List of figures and tables

Figure 4.8 Electrophoresis gels and MALDI spectrum of the Fbw7-Skp1 complex.....	86
Figure 4.9 DSF Plots of Fbw7-Skp1 in complex with a peptide of the natural substrate Cyclin and F1A1, F1A2, F1A3 F1A4, F1A5.....	87
Figure 4.10 SPR plot screening of a 96-wells plate at a single concentration.....	88
Figure 4.11 Dose-response curves of the most potent fragments.....	94
Figure 4.12 Sensogram of the two best fragments, F3A8 and F6H3.....	96
Figure 4.13 Dose-response curves of the most potent fragments with a GST immobilization protocol.....	99
Figure 4.14 Sensograms of the two best fragments F3A8 and F6H3 with a GST-immobilization protocol.....	102
Figure 4.15 F3A8 (250 μ M) in complex with Fbw7-Skp1 (10 μ M).....	103
Figure 4.16 F4G7 (125 μ M) in complex with Fbw7-Skp1 (2.5 μ M).....	103
Figure 4.17 F6D5 (250 μ M) in complex with Fbw7-Skp1 (2.5 μ M).....	103
Figure 4.18 F6H3 (100 μ M) in complex with Fbw7-Skp1 (1.25 μ M).....	104
Figure 4.19 F7E2 (100 μ M) in complex with Fbw7-Skp1 (1.25 μ M).....	104
Figure 4.20 F1G4 (100 μ M) in complex with Fbw7-Skp1 (1.25 μ M).....	104
Figure 4.21 F1D5 (100 μ M) in complex with Fbw7-Skp1 (1.25 μ M).....	105
Figure 4.22 F6F12 (250 μ M) in complex with Fbw7-Skp1 (1.25 μ M) 0.05% tween-20.....	105
Figure 4.23 F4E10 (100 μ M) in PBS.....	106
Figure 4.24 F4E10 (100 μ M) in complex with 0.05% Tween-20.....	106
Figure 4.25 F4E10 (100 μ M) in complex with Fbw7-Skp1 (1.25 μ M) 0.05% tween-20.....	106
Figure 4.26 Density map of Fbw7-Skp1 soaked with F3A8.....	107
Figure 4.27 Most probable binding site of a fragment in the degron site of Fbw7.....	109
Figure 4.28 Electrophoresis gels and MALDI spectrum of the ¹⁵ N Fbw7-Skp1 complex.....	111
Figure 4.29 Curve replicates showing binding of the labelled peptide FICT-DISC1 to Fbw7-Skp1.....	113
Figure 4.30 FP competition assay between DISC-1 and FICT-DISC1.....	113
Figure 4.31 IC ₅₀ of FICT-DISC1 in the presence of different fragments.....	114
Figure 4.32 IC ₅₀ of FICT-DISC1 in the presence of different concentration of DISC1 once using a high percentage of DMSO (8%).....	115

List of figures and tables

Figure 4.33 ¹ H-Fbw7-Skp1 complex in PBS pH= 8.0 and ¹ H-Fbw7-Skp1 complex in PBS 5% DMSO pH= 8.0.....	116
Figure 4.34 ¹ H-Fbw7-Skp1 complex in PBS 5% DMSO pH= 8.0 after 24h and ¹ H-Fbw7-Skp1 complex in PBS 5% DMSO pH= 8.0 after 48h.....	116
<u>Chapter 5</u>	123
Figure 5.1 Structure of CRBP proteins.....	126
Figure 5.2 Sequencing of CRBP L78→I mutant plasmid.....	129
Figure 5.3 Purification of CRBP proteins.....	130
Figure 5.4 Kinetic fitting of retinol binding to CRBP-II wild-type and to CRBP-II L78→I.....	131
Figure 5.5 Sensogram of retinol binding to CRBP-I.....	132
<u>Chapter 6</u>	135
Figure 6.1 Structure and molecular recognition of BET bromodomains.....	139
Figure 6.2 X-ray structure of 1XA in complex with Brd4(BD1) (PDB 4LR6).....	141
Figure 6.3 Synthetic route of SSR3, SSR4, SSR1 and SSR2.....	143
Figure 6.4 Possible side reactions during Pre-SSR1 synthesis.....	144
Figure 6.5 Side product of pre-SSR3.....	145
Figure 6.6 Brd4(BD1) his-tag purification.....	146
Figure 6.7 Normalized TR-FRET assay results.....	149
Figure 6.8 X-ray structure of Brd4(BD1) in complex with SSR4, SSR2, SSR3.....	154
 List of tables	
Table 1.1 Fragment-based drugs in clinical trials.....	25
Table 3.1 General buffers for purification.....	47
Table 3.2 Main parameters for protein immobilization.....	54
Table 3.3 SPR-fragment screening buffers.....	54
Table 4.1 SPR results.....	92
Table 4.1 GST-SPR results.....	98
Table 6.1 TR-FRET results.....	148

List of figures and tables

Chapter 1

Introduction

1.1 Preface

Proteins are the most important pillars of our body and the hardest workers of our cells. Indeed, they specifically bind small molecules, nucleic acids and multiple molecular interactions like protein-protein, protein-nucleic acid and protein-ligand interactions control and enable life. Understanding all of these interactions is of crucial importance to restore the normal physiology of pathological-altered cell pathways. The entirety of approved small-molecule drugs act through approximately 200 human proteins, a small number compared to the more than 20000 protein-coding human genes [1]. Therefore, a huge effort has to be done to better identify and modulate “undruggable” proteins involved in pathways with strong disease phenotypes.

Most pharmaceutical companies research is based on target-focused drug discovery, where the goal is to modulate the biological activity of a particular target to provide a desired response to cure or treat a specific disease. In the last years, our increasing knowledge of molecular biology and the development of new technologies and approaches have helped scientists in their understanding of the molecular mechanisms that underlie disease processes. Potent, selective, and cell-penetrant small-molecules, often referred as “chemical probes,” also provide powerful tools to aid elucidation of protein function inside the cell [2].

Since the late 1990s, within the pharmaceutical industry, significant investment and efforts have focused in both phenotypic and target based high-throughput screenings (HTS) of large library collections (>100,000 compounds) [3]. Typically, HTS requires robotized bioassay to screen thousands of compounds in a whole cell set-up or using purified protein. The use of this approach turned out to be really fruitful with the approval in the

market of drugs for different diseases/targets: e.g. gefinitib, lapatinib, erlotinib to treat cancer; tipranavir, maraviroc for HIV, sitagliptin for diabetes and many others [3].

Although these approaches succeeded at identifying biologically active compounds, more recently, significant advances in biophysical and structural techniques for monitoring weak to moderate binding affinities of protein–ligand interactions have facilitated the development and success of other approaches, i.e. fragment-based drug discovery (FBDD). In fragment screening, compared to HTS, smaller libraries (usually around 1,000) of compounds of relatively small size (fragments, MW usually <300 Da) are screened at high concentration (usually >0.5 mM) for direct, non-covalent binding to a given target [4]. Fragment screening is now firmly established as an early-stage lead discovery approach, very often performed in parallel with HTS against any target of interest. The key for the success of fragment approaches relies in biophysical and structural methods, which were previously only used for quality controls or during the late stages of lead optimization and now are being increasingly used for screening and validation during the early stages of the discovery process [4].

HTS and FBDD, together with computational approaches, have speeded up the drug discovery pipeline. In parallel, these approaches will continue providing to us strong information to better understand our cell machinery. Indeed, progresses made from the drug discovery side have been so far really helpful to better understand how difficult targets work, how they interact with other proteins and finally, how they can be activated or inhibited. However, a lot of work still remain to be done in order to bring to light the “dark proteome”, to modulate proteins that until now are considered undruggable and to better understand the secrets hidden by our genetic code, the DNA.

1.2 A historical view of the biophysical techniques: from the resolution of the first crystal structure to 2020

Since the first three-dimensional protein structures published during the 50s and 60s, a considerable number of structures of relevant targets have been elucidated, allowing scientist all over the world to get more insights in the structural determinants of these targets. The pioneering work of Drs John Kendrew and Max Perutz in solving the crystal structures of myoglobin and haemoglobin explained the oxygen-carrying/storing properties of these proteins and also shed light on the molecular basis of sickle cell anaemia, and at the same time on potential treatments for this disease [5]. Similarly, thanks to the determination of the amino acid sequence of insulin by Dr Fred Sanger and its three-dimensional structure by Dr Dorothy Hodgkin, it was possible to engineer synthetic insulins for the treatment of diabetes. These ground-breaking studies demonstrated how important is the identification of structure-function relationships and furthermore showed the importance of having in hands the three-dimensional structures of therapeutically relevant proteins. Moreover, insights on protein fold and structures (Drs Linus Pauling and Robert Corey proposed the α -helix and β -sheet structures in 1951) [6] and the availability of the three-dimensional structure allowed the identification and characterization of potential binding sites on targets and formed the foundation of what we call, structure-based drug design (SBDD) [5].

During the 80s, structures-guided programs started aiming to find inhibitors for different targets, leading to the discover of successful drug candidates like inhibitors of the thymidylate synthase for the treatment of cancer and inhibitors of the viral neuraminidase to combat influenza [5]. Another great

achievement of the SBDD approach was the use of the structure of HIV protease in the design of the four FDA-approved antiviral protease inhibitors (saquinavir, nelfinavir, indinavir and ritonavir) for the treatment of HIV/AIDS [7-8]. A major change in the field was during the 90s, when equipment for X-ray structure determination and computing and graphics equipment required for the corresponding molecular modelling was available in most institutions in both academia and industry [9]. Indeed, for X-ray crystallography the major developments were as a consequence of an increase in the speed of the structure determination. Currently, synchrotron radiation, coupled to new faster instrumentation is capable of rapid data collection.

In parallel to X-ray crystallography, Nuclear Magnetic Resonance (NMR) gained ground during the 90's. An Abbott-based group, led by Dr Stephen Fesik, developed the SAR by NMR approach [10]. NMR has been applied to design potent, novel leads against a number of targets, including the first inhibitor of an important target in cancer, the Bcl-2 family (ABT-737/navitoclax) [11]. At that time, it became quite clear that structural insights could not only speed up but, above all, improve the success of drug discovery efforts. In parallel to X-ray and NMR, other biophysical techniques such as Differential Scanning Fluorimetry (DSF), Surface Plasmon Resonance (SPR, as an example, the first SPR instrument was brought to the market by Pharmacia Biosensor AB in 1990), Isothermal Titration Calorimetry (ITC) or more recently Microscale Thermophoresis (MST) and others, have matured during these years to become key components of drug discovery platforms, enabling drug discovery for more challenging targets [12-13].

In the last years, two new promising technologies are taking their place in structural biology: serial femtosecond crystallography (SFX) using ultra-short pulses of coherent and extremely intense radiation generated by X-ray free electron lasers (XFELs) [14] and high-resolution single-particle cryo-electron microscopy (cryo-EM) [15]. XFELs can be used for targets where X-ray crystallography fails. Indeed, X-ray suffers from limitations such as the need for large, well-diffracting crystals and radiation damage that can hamper native structural determination [16]. Instead, XFELs works on small crystals that are not commonly suitable for diffraction on regular synchrotron beamlines and for macromolecules where radiation damage represents a limitation [17]. Furthermore, it helps to obtain information on molecular dynamics, since no cryogenic temperature are needed [12]. Cryo-EM, instead, is a well know technique, that until very recent advances in electron detection and image processing, was limited to large complexes (200–4,000 kDa) due to its low resolution. In the last years, cryo-EM has rapidly emerged as a powerful technique for protein structure determination at high resolution and has come to rival X-ray crystallography [18]. Due to these improvements, there are a lot of targets that have been structurally disentangled by cryo-EM, including TRPV1 [19], γ -secretase [20] and β -galactosidase [21] for which, for example, a 2.2 Å-resolution structure was obtained. Furthermore, for some systems, cryo-EM has been able to achieve the resolution needed to elucidate details of protein–ligand interactions, and for that reason is considered to be of utility in drug discovery. Noteworthy, scientists at Astex Pharmaceuticals and Isohelio demonstrated, using as test systems β -galactosidase (Bgal) and the oncology target pyruvate kinase 2 (PKM2), that cryo-EM can become a powerful tool for FBDD since it gives sufficient quality data and throughput to determine unambiguous binding mode for ligands [22]. Today, more than

in the past, it became clear that SBDD programs will benefit of combining data obtained from different biophysical techniques and other approaches, i.e. HTS and computational approaches.

1.3 Biophysical techniques in the drug discovery process

Drug discovery is a long, complex and expensive process. On average it takes up to 15 years and 1.5 billion dollars to bring a new drug to market [23]. The path to obtain a successful drug, indeed, is hard and needs a multidisciplinary approach and team from both basic and clinical researchers. From an SBDD perspective, after identifying an unmet medical need, the next step is to identify the molecular biology behind it, and therefore the target that causes the disease. Once the target is identified and characterized, the long pathway to develop a drug starts. In this process, biophysical methods also have an important part to play and nowadays are firmly combined at several stages. In order to affect the biology of a specific system, it is of crucial importance to find a hit compound as a starting point to probe the *druggability* of the target. Here, biophysical techniques help to elucidate the binding of a ligand to a given target, directly (K_d) or indirectly (IC_{50}); providing the basis to decipher structure-activity relationships. Also, in this stage biophysical techniques contribute to the elucidation of binding site and binding mode of ligands, allowing to obtain structural information, even for weak binding molecules. As mentioned, a plethora of biophysical techniques are available for this hit identification, and lead development and several parameters should be considered, including the throughput of the technique, the consumption of material (recombinant protein and ligand), the requirement of covalent immobilization, the detectable range of K_d s, and obviously the accessibility of the instrument [4]. DSF, fluorescence polarization (FP) assays and Förster

resonance energy transfer (FRET), NMR, SPR, ITC, X-ray crystallography, MST and Mass Spectrometry (MS) are the most applied techniques in drug discovery. A possible strategy of compound screening is illustrated in Figure 1.1. In the following subchapters, exclusively the techniques employed in this thesis will be deeply described.

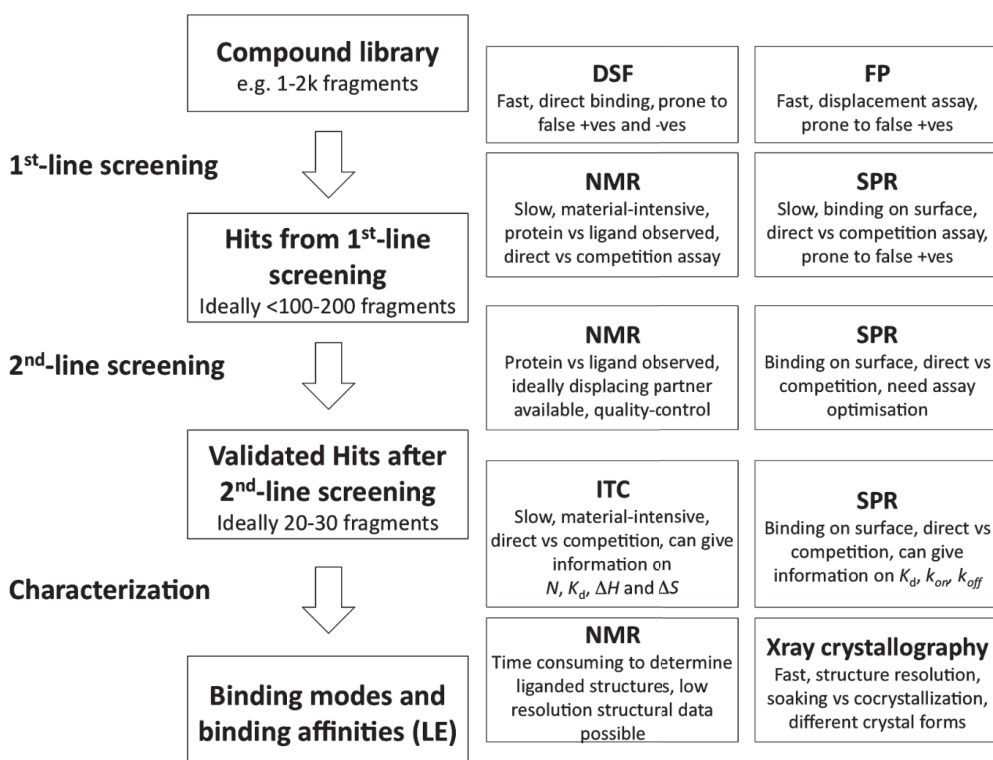


Figure 1.1. Flowchart of a possible strategy for compound screening, validation, and characterization using biophysical techniques. Adapted from [4].

1.3.1 Differential Scanning Fluorimetry (DSF) assay

DSF (also named thermal shift and ThermoFluor) is a high throughput screening, inexpensive, simple and quick method that measures the progressive denaturation of a protein through heating. Proteins exist in thermodynamic equilibrium between multiple conformational states (folded/native state) that upon slow increase of the temperature undergo denaturation (unfolded state) since this heating induces changes in the water

shells and the hydrogen bonds, thus leading to associated adjustments in the conformation of the protein and exposing to solvent hydrophobic residues normally buried inside the core of the protein [4]. The assay consists of combining a protein solution with a fluorescent hydrophobic dye (SYPRO is the most common) that fluoresces brightly when bound to the hydrophobic region of a protein. Therefore, as the protein unfolds during heating and reveals binding sites for the dye (hydrophobic regions), fluorescence increases proportionally to unfolded protein abundance [24]. Melting temperature (T_m) is then calculated as the midpoint of the resulting fluorescence versus temperature (Fig. 1.2).

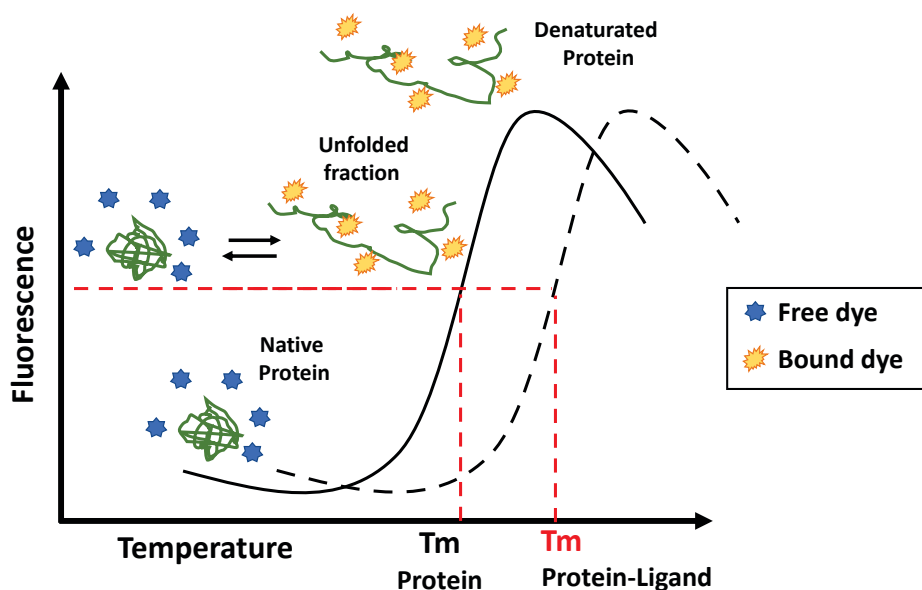


Figure 1.2. Examination of differences in the temperature dependent fluorescence profile of protein plus dye in the presence and absence of a potential ligand (small molecule) may reveal a change in melting temperature (T_m), the temperature at which there is 50 % denaturation, to a higher value indicative of binding.

Since the stability of a given protein may be affected by many factors like buffer condition, protein-protein interaction, mutations, and ligands binding, DSF is widely used in academia and in pharmaceutical companies at the very early stage of the drug discovery process. Observation of a change in the melting temperature upon addition of a ligand in the protein solution to a higher value may be indicative of binding. Indeed, a specific

small ligand can enhance the stability of the protein through improved folding thanks to binding or through a more favourable environment of ionic strength, counterions and others [25]. Therefore, DSF is largely applied to assess the binding of small molecules or fragments, in HTS or FBDD, or to validate and characterize hits. Usually, the melting temperature can be obtained by plotting the derivative of the fluorescence signal against temperature (Fig. 1.3).

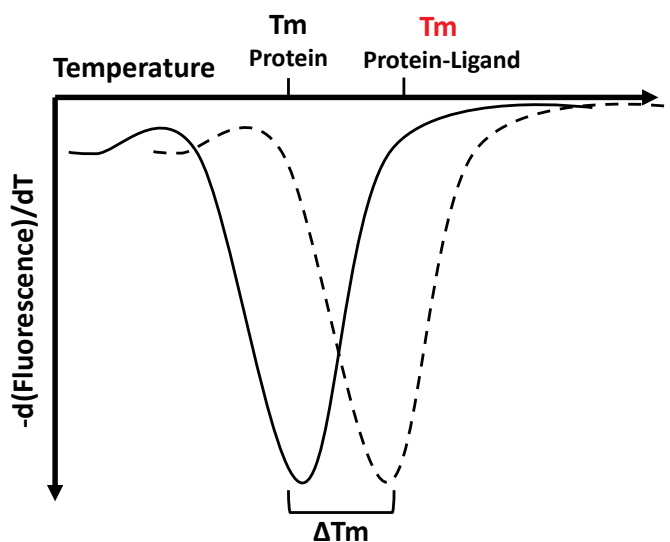


Figure 1.3. The melting temperature is determined by the point of inflection of the curve showed in Figure 1.2. This can most easily be assessed by plotting the derivative of the fluorescent signal against temperature.

Hundreds of runs can be performed in one day using a quantitative polymerase chain reaction (qPCR) instrument, making this technique ideal for high throughput screening in a low-cost manner. Another advantage of this technique is the versatility of its application and that it does not require high amounts of materials. DSF can be applied for most target as purified proteins, however other derivative applications based in the same physicochemical principle have been developed, i.e. to probe target engagement (Cellular Thermal Shift Assay (CETSA)) [26]. Of course, the technique has its own pitfall and can give rise to false positives and false negatives. Nevertheless, DSF represents a good starting point for a primary screening in the drug discovery pipeline if then combined with other orthogonal techniques.

1.3.2 Surface Plasmon Resonance (SPR)

SPR is an optical technique that allows real-time, label-free detection of biomolecular interactions. SPR phenomenon occurs when a beam of polarized light is shone, under conditions of total internal reflection (TIR), through a prism at the interface of a gold layer between two media with different refractive indices, e.g. the glass of a sensor surface (high refractive index) and a buffer (low refractive index). The light penetrates some distance into and beyond the gold (in a phenomenon called evanescent wave formation) at a particular angle of incidence, then the electrons in the gold absorb the energy coming from the light and excite charged density waves, called “surface plasmons” that will start propagating along the surface of the metal. At the maximum absorption (resonance condition) the reflected light decreases drastically and sharply and can be detected. The evanescent wave extends in the solution for approximately 100–300 nm and therefore, only a change in the refraction index close to the gold-aqueous solution interface, due for example to something binding at the surface of the sensor, will cause a shift in the resonance angle. Any change in the refractive index of the solution is recorded and can be monitored in real time. The gold layer is functionalized with a dextran matrix that allows immobilization of a protein and subsequently ligand can be flown across it. The outcome of the analysis is a sensogram with responses measured in resonance units (RU) vs time in seconds. The SPR sensogram gives mainly two kinds of information: (a) the rate of interaction (association, dissociation, or both), which provides information on kinetic rate constants and analyte concentration, and (b) the binding level, which supplies data on the affinity constants and can be used for qualitative or semi-quantitative applications (Fig. 1.4). Kinetic constants provide valuable data about the dynamics of a biological system, which could be further applied for the selection and rational design of new molecules of

therapeutic interest [27]. Above all, the use of SPR analysis is taking ground in order to improve the drug efficacy during the lead optimization process.

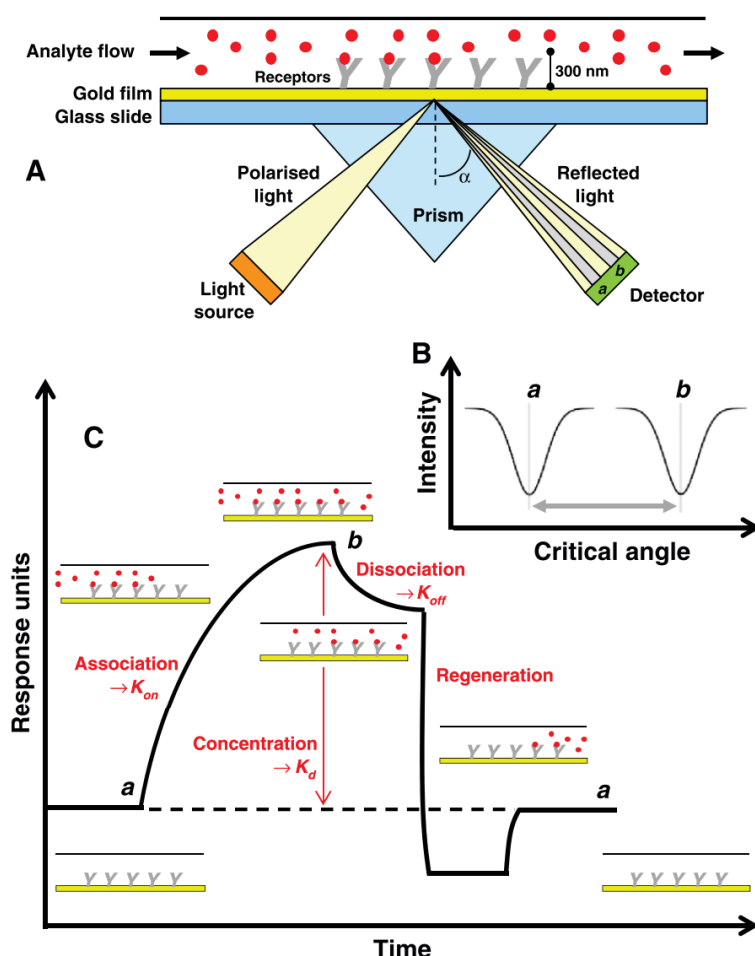


Figure 1.4. **A)** The polarized light strikes the gold layer through the prism and generates the evanescent wave that extends up to 300 nm in the solution. The light will be reflected and recorded at a particular angle of incidence. Once an analyte (red) binds to the receptor immobilized on the surface (grey) of the gold layer, there will be a shift in the resonance angle that will be recorded. **B)** Change in the critical angle of incident light from the angle *a* to angle *b* on binding of an analyte molecule to a receptor molecule. **C)** Sensogram that illustrates the association of the analyte once binding to the receptor (k_{on}), the steady state when all receptors are occupied thus achieving the maximum response (R_{max}) that can be used to extrapolate the binding affinity (K_d) and the dissociation rate when the binding site becomes unoccupied (k_{off}). The surface can then be regenerated to start the experiment again. Adapted from [28].

Drug residence time ($1/k_{off}$) emerged as an important factor to predict in vivo behaviour of a drug and relates to the life-time of drug-target complex. Therefore, improving kinetic properties can reduce off-target effect and enhance drug efficacy [29]. Figure 1.5 highlights how drug candidates can

have different properties (different k_{off}) although the K_d remains unchanged [30].

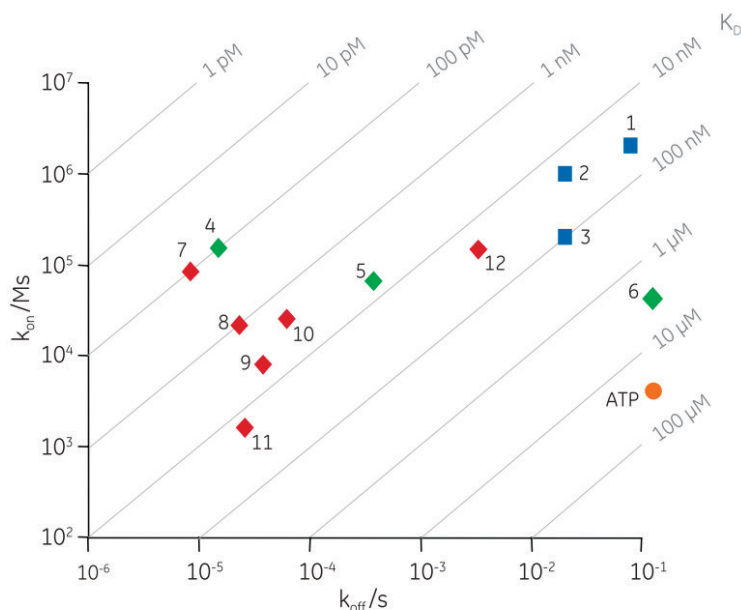


Figure 1.5. On/off rate plots from SPR kinetic analyses provide an informative way to demonstrate the resolved scale of affinities and how these affinities can be distributed over a range of on- and off-rates. The ability to characterize slow off-rate ligands is particularly useful during the optimization of the pharmacokinetic properties of lead compounds and drugs. Adapted from [30].

The main strengths of SPR rely on the wide range of binding detection (from pM to mM) and on the low amount of protein (picograms) and ligand needed, that makes this technique suitable for both HTS and FBDD. These advantages make SPR a powerful technique that stands across other popular techniques such as NMR or ITC. In contrast, the covalent immobilization of the target to the surface, if needed, represent one of the major of weakness of this technique.

However, it is important to highlight that depending on the type of experiments to perform, to overcome the limitation of a covalent immobilization, different protocols can be followed with several commercial sensor chips containing different moieties such as biotin-avidin based, antibody based (e.g. antibody anti-Glutathione-S-Transferase (GST)), affinity immobilization (e.g. Nickel-Nitrilotriacetic acidic (Ni-NTA) for recombinant proteins containing a poly-

histidine) or disulphide based (mobile link depending on reducing agent such as glutathione). In this way, choosing a specific protocol, it is possible to constrain the protein on the chip as desired.

1.3.3 Fluorescence Polarization (FP) assay

FP is a fluorescence-based technique in which a fluorescent labelled molecule is excited with polarized light. The related anisotropy reveals information on the molecular mobility, which is dependent on size and shape. The principle of FP derives from the fact that the degree of polarization of a fluorophore is inversely related to its molecular rotation, itself being largely driven by Brownian motion. Quantitatively, FP/FA is defined as the difference of the emission light intensity parallel (F_{\parallel}) and perpendicular (F_{\perp}) to the excitation light plane normalized by the total fluorescence emission intensity and can be expressed in terms of FP (P) or anisotropy (r):

$$P = (F_{\parallel} - F_{\perp}) / (F_{\parallel} + F_{\perp})$$

$$r = (F_{\parallel} - F_{\perp}) / (F_{\parallel} + 2F_{\perp})$$

The smaller the rotation of a fluorescent molecule, the higher the anisotropy. Indeed, FP can be detected once a fluorescent labelled molecule binds to a target protein, limiting the freedom of rotation of the fluorophore. Hence, FP can provide a direct readout of the binding between a fluorescent ligand and a protein. In order to perform FP, usually a fluorescence dye is used and attached to a fast-rotating molecule like, for example, a known peptide, substrate of the target protein (in the case of protein-protein interactions). Therefore, the labelled peptide alone in solution gives a very low FP signal, since the emission of the fluorescence is randomized, which increases drastically when the peptide binds to the

protein (Fig. 1.6) [31-32].

Since obviously not all small molecules are fluorescent, FP assays for screening compound libraries are carried out in a competitive inhibition mode by titrating different concentrations of small molecules against a sample containing selected concentrations of the protein and the fluorescent-version of a known ligand in order to generate a dose–response curve, which can be used to determine IC_{50} s (and through back-calculation K_d s) [4]. The displacement of the fluorescent peptide by direct binding of the small molecule or by allosteric modulation will be reflected in a decrease of the FP signal.

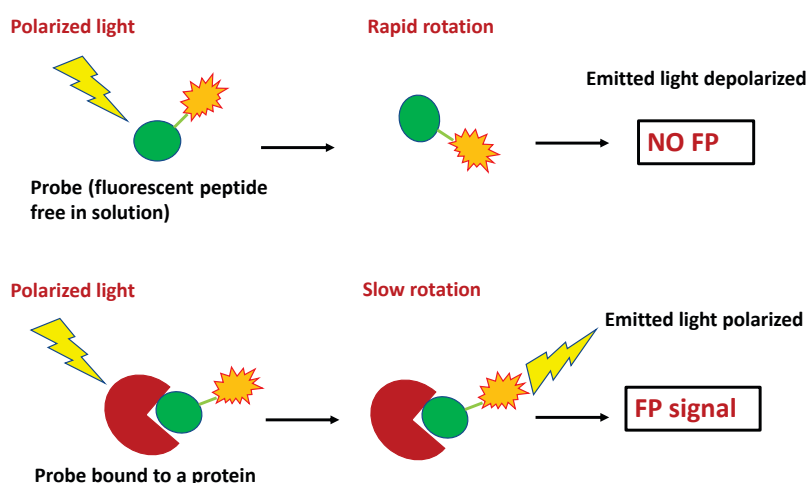


Figure 1.6. When the fluorescent probes are free in solution, rapidly rotating in the absence of their target protein, the FP will result in low signal. Conversely, binding of the fluorescent probe to a large, slowly rotating protein will result in high FP signal.

Intensive optimization of the experimental conditions is required when employing FP and several points should be considered. As a ratiometric method, FP is relatively insensitive to absorptive interferences or inner filter effects, but can suffer from autofluorescence and light scattering as these effects can confound sample FP calculation. Additionally, the labelled tag should not affect the binding with the target protein. To check that, a titration

of different concentrations of the protein with a known concentration of the labelled peptide has to be done to extrapolate the K_d of this labelled peptide. Once the experiment works, a competition experiment with the unlabelled peptide (in this case without the fluorescent tag) has to be performed. This second titration yields an “apparent” dissociation constant (K_{dapp}) for the fluorescent complex that should be larger than that measured in the absence of the unlabelled peptide. From this K_{dapp} , it can be calculated the K_d of the unlabelled peptide, allowing to confirm that the introduction of the fluorophore has not affected the binding to the protein. An important parameter that should be checked and can be derived by these preliminary, but necessary experiments, is the Z factor. The Z factor quantifies the quality of the assay and can be calculated as following:

$$Z' = 1 - \frac{3(\sigma_{\text{free}} + \sigma_{\text{bound}})}{mP_{\text{bound}} - mP_{\text{free}}}$$

where σ_{free} , σ_{bound} , mP_{free} , and mP_{bound} are the standard deviations and the means of the negative (free) and positive (bound) controls giving rise to a polarized light, respectively. A value greater than 0.7 is considered to yield a good assay format. Once established the right assay conditions, small-molecules can be screened at an intermediate throughput speed and the IC_{50} and K_d can be calculated with a moderate accuracy, with a detectable range “only” down to the K_d of the fluorescent probe [4].

1.3.4 Nuclear Magnetic Resonance (NMR) spectroscopy

NMR spectroscopy represents a powerful tool for characterizing protein–ligand interactions in solution under near physiological conditions [33]. It has a wide range of applications and is routinely applied in almost all drug discovery pipelines. The advantages that NMR offer are remarkable and

indeed it allows to assess the affinity and specificity of interactions, to solve structures of protein–ligand complexes and therefore identify binding epitopes on proteins and ligands, to characterize structural rearrangements induced by binding and to detect even very weak interactions (mM range). This last characteristic fits perfectly with FBDD (to note that FBDD was first made possible 20 years ago by the development of the so called structure-activity relationship by NMR) [10]. Moreover, NMR is less prone to artefacts compared to other biophysical techniques and therefore low incidence of false positives and false negatives are observed. In contrast, one of the main limitations of NMR is its low-throughput. However, in the last years, the NMR field has experienced a radical change in automatization, rendering faster screenings of compound libraries. Furthermore, the development of high sensitivity probes (e.g., cryoprobes), new pulse sequences, efficient isotopic labelling techniques, and more powerful magnets, have all contributed significantly to minimize the main limitations [33]. NMR experiments for protein–ligand interactions can be divided in two main categories: either studying them from the perspective of the protein (Protein-Observed NMR) or from the perspective of the ligand (Ligand-Observed NMR).

1.3.4.1 Protein-Observed NMR experiments: chemical shift perturbations

Protein-observed experiments are much more informative than ligand-observed experiments because the former can yield structural information. But isotopic labelling and assignment of the NMR chemical shifts are a prerequisite. Upon protein-ligand binding, changes in the physicochemical properties of both protein and ligand can be observed. Once a protein and a ligand interact there is a rearrangement of the electron density of the atoms involved in the interaction, that can be due, for example, to differences in the

hydrophobicity at the interaction surface. These differences have a strong influence on the chemical shift of magnetic nuclei that can be recorded with NMR [34]. However, conventional ^1H -NMR experiments (1D) are not generally suitable to resolve the protons signal of the protein and therefore, 2D experiments have to be performed studying active nuclei rather than protons, such as ^{15}N and ^{13}C . The abundance of these nuclei is very low (0.37 % for ^{15}N and 1.1% for ^{13}C), for this reason, isotopic labelling of the protein is necessary. A widely used experiment is the two-dimensional proton-nitrogen correlation ($^1\text{H},^{15}\text{N}$ HSQC), which measures $^1\text{H},^{15}\text{N}$ spin pairs. Heteronuclear $^1\text{H}-^{15}\text{N}$ correlation NMR experiments generate a spectrum containing at least one signal per residue (named chemical shift) except for proline and two extra ones per glutamine and asparagine side-chains. $^1\text{H},^{15}\text{N}$ pairs are very sensitive to the chemical environment and their chemical shifts can be monitored upon addition of ligands. The perturbation of each peak can be measured (chemical shift perturbations or CSPs) to determine ligand binding sites (Fig. 1.7).

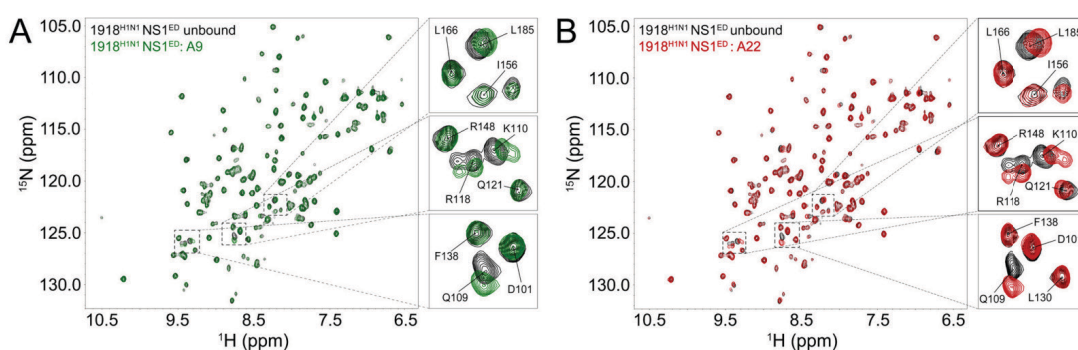


Figure 1.7. Overlay of $^1\text{H}-^{15}\text{N}$ HSQC spectra of the 1918H1N1NS1ED protein alone (black spectra) after adding (A) a ligand A9 (green spectra) and (B) a ligand A22 (red spectra). Adapted from [35].

One of the limitations of this approach lays in the size of the protein (usually <40 kDa), since in proteins with high molecular weight the population of signals is too crowded to be assigned. To overcome this drawback, selective amino acid labelling could be performed, and with that

there will be less noise but also less information and binding event can be missed. Another important drawback is the quantity of material; huge amount of protein (>0.1 mM per 200 μ l) is often required limiting the use for proteins expressed with a low yield. Finally, as mentioned, the signals from the two-dimensional protein NMR spectra have to be assigned (depending on the complexity of the protein structure it can be a laborious puzzle to solve and 3D experiments are needed).

CSP can be related to direct ligand binding (short distance effect) or to conformational changes of the protein structure after allosteric binding (long distance effect). This can make the interpretation of the chemical-shift changes ambiguous, even when assignments are available. The differential CSP method, in which the spectra of the target protein is compared when it is bound to two slightly different ligands, is an efficient solution to solve this problem [33, 34, 35, 36]. In spite of these limitations, protein-based NMR represents a powerful technique in the elucidation of binding of ligands for “NMR-friendly” targets, with low molecular weight and that can be expressed in abundance.

1.3.4.2 Ligand-Observed NMR experiments: Saturation Transfer Difference (STD)

Another important approach in NMR is the ligand-observed NMR in which, instead of focussing the attention on proteins NMR-signals outcome, the signals of the ligands are monitored. Crucially, there is no need for isotopic-labelling of the protein or previous assignment of the protein. Moreover, in comparison with protein-based NMR, the ligand approach requires less amount of material and experiments are easier to carry-out and could be applied to any protein of interest with a molecular weight >10 KDa. Compounds with different chemical structures can be screened in mixture, reducing the time of the experiments and speeding-up the

ligand discovery process. Ligand-observed experiments exploit the differences in the relaxation properties of a free versus a bound ligand. One commonly applied experiment is Saturation Transfer Difference (STD) that is based in the Nuclear Overhauser Effect (NOE) and relies in transfer of magnetization (signal) directly from the protein to the ligand, which is only possible in the bound ligand state, e.g. by exciting the aliphatic methyl group region of the protein (Ala, Val, Leu, Ile). (Fig. 1.8).

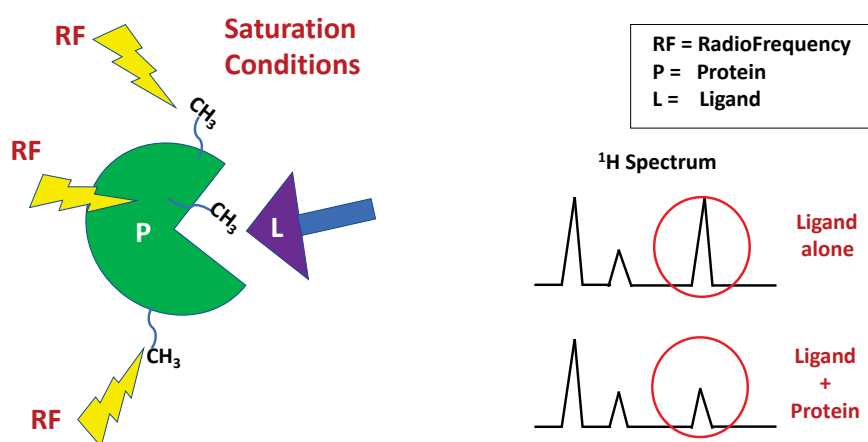


Figure 1.8. In STD, the protein methyl groups are saturated and magnetization is transferred to the bound ligand. A decrease in signal intensity is observed. The ¹H spectra before and after protein addition are used to calculate a difference spectrum (STD) that will reveal only protons involved in the binding.

An STD spectrum is generated by the difference between an on-resonance spectrum and an off-resonance spectrum. In the on-resonance spectrum a frequency-selective pulse hits the protein in the range in which only protein protons are present (e.g., at -1 ppm) to saturate the protein. If there is a ligand binding to the protein, the magnetization from the protein is transferred to the ligand and this will result in the nearly complete disappearance of protein signals and partial attenuation of protons ligand signals (Fig. 1.8). Interestingly, the degree of saturation received by each ligand proton is not equal, but rather depends on their proximity to the protein. Therefore, this property can be used to determine the binding epitope of the ligand [33]. In

the off-resonance spectrum, no change in signal intensities is observed since the frequency-pulse is irradiated far from both protein-ligand protons range (e.g. 40 ppm). Therefore, in the STD spectrum, only the signals of the binding molecule are visible and STD % can help in elucidating structure-activity relationship.

1.3.5 Förster Resonance Energy Transfer (FRET)

FRET is a powerful technique capable to inform about dynamic interactions between proteins and a plethora of biochemical signalling events based on the development of specific biosensors [37]. This technique relies on the transfer of energy from a fluorophore-donor excited at a specific wavelength to a fluorophore-acceptor, recording the intensity of fluorescence of the emitting acceptor. For transfer to occur, the two fluorophores must be in close distance (≤ 10 nm), because the rate of FRET transfer is inversely proportional to the sixth power of the distance between the fluorophore and acceptor. Also, the orientation of the two fluorophores has to allow dipole-dipole pair matching. A parallel orientation affords 100% transfer efficiency, while there is no transfer if the orientation is perpendicular (Fig. 1.9). The efficiency of the FRET transfers also relies on other factors like: i) quantum yield of the donor, ii) the refractive index of the solution and iii) the spectral overlap between donor and emitter. FRET represents the simplest way to access intra- and intermolecular distances on the nanometre length scale. For this reason, FRET is considered a ‘molecular ruler’ for probing macromolecular structures [38]. Indeed, although there are other techniques with higher resolution such as X-ray crystallography, FRET is more accessible, less technically complex, and do not require restrictive sample preparation (e.g. crystallization for X-ray) [39].

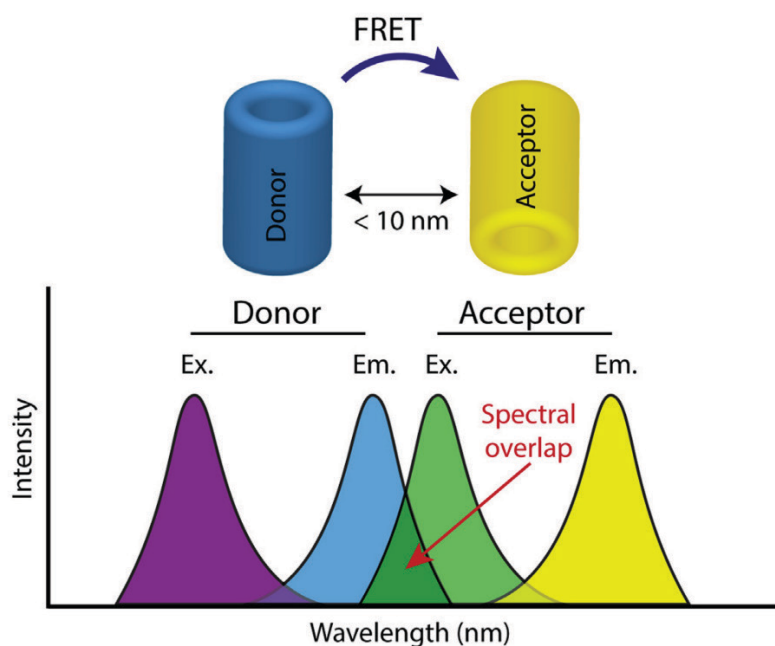


Figure 1.9. FRET occurs when the emission spectrum of the donor overlaps the excitation spectrum of the acceptor. FRET efficiency is maximum when the distance between donor and acceptor molecules is within 10 nm and the dipoles are with parallel orientation. Adapted from [37].

FRET has a wide range of applications, including studying protein-protein interactions, measuring protein activation upon binding of a positive modulator (e.g. in GPCRs), assessing post-translational modifications, detecting enzyme cleavage activity and characterising ligand binding. In the last case, a common approach is to extrapolate binding indirectly by competition when the ligand of interest is not fluorescent. Furthermore, if the binding site of the putative ligand is known and a tryptophan is buried inside the cavity, intrinsic FRET (iFRET) can be assessed [40]. The advantage of iFRET relies on the possibility of avoiding laborious labelling of the protein.

1.3.5.1 Time-Resolved FRET (TR-FRET)

Another useful application of FRET is Time-Resolved FRET (TR-FRET) that allows to eliminate background fluorescence caused by sample components, but requires more specialised donor fluorophores and

detection equipment. This background is extremely transient (with a lifetime in the nanosecond scale) and can be avoided by measuring FRET of fluorophores with a very long fluorescence lifetime. The peculiarity of the technique relies in that the donor species used have a fluorescent lifetime many orders of magnitude longer (from microseconds to milliseconds) compared to those used in standard FRET, which have a decay of nanoseconds, where excitation of the sample and measurement of emission occurs simultaneously. The fluorophores of choice are usually chelates of lanthanides (most commonly europium, terbium and samarium). Its fluorescence lifetime is in the range of microseconds (μs), while in most fluorophores the lifetime of its fluorescence is in the range of nanoseconds (ns). These feature in TR-FRET allows to measure the resulting emission signal after any interfering signal has completely decayed, greatly reducing background and thus increasing the signal-to-noise ratio [41] (Fig. 1.10). Although TR-FRET is not the first technique of choice in a drug discovery process, it represents a versatile, medium throughput technique capable of giving important information to better elucidate the binding of a ligand.

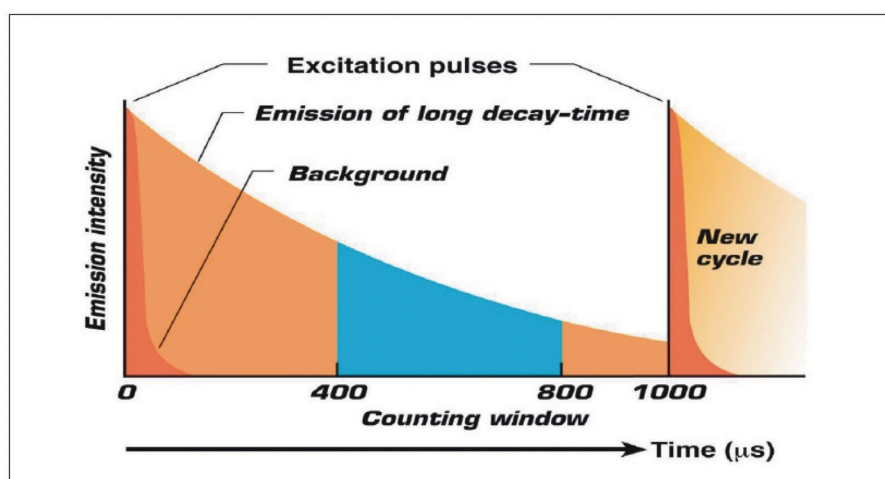


Figure 1.10. Time-resolved FRET measurements. The fluorescent lifetime of most common fluorophore is in the nanoscale range (100 ns or less). TR-FRET fluorophores exhibit a long fluorescent lifetime (200 to 1500 μs) allowing to avoid the measure background noise and after a delay time (e.g. 400 μs) FRET can be measured (blue window). Adapted from [41].

1.4 Introduction to fragment screening

Fragment-based screening, developed within the pharmaceutical industry in the late 90s, has rapidly established itself as a robust and powerful approach to identify good quality hits against any protein target. Moreover, fragment-based screening is also now an established approach for interrogating target *druggability* in vitro and in whole cells [42, 43]. Forty-seven fragment-derived drugs reached clinical trials and 4 candidates are already approved by the FDA: the BRAF inhibitor vemurafenib, the BCL-2 inhibitor venetoclax, the FGFR1-4 inhibitor erdafitinib and the last approved, in August 2019, the CSF1R-KIT inhibitor pexinartinib (Table 1).

Table 1.1. Fragment-based drugs in clinical trials (Nov 2020)

Drug	Company	Target
Approved		
Erdafitinib	Astex/J&J	FGFR1-4
Pexidartinib	Plexxikon	CSF1R, KIT
Vemurafenib	Plexxikon	B-RAF ^{V600E}
Venetoclax	AbbVie/Genentech	Selective BCL-2
Phase 3		
Asciminib	Novartis	BCR-ABL
Lanabecestat	Astex/AstraZeneca/Lilly	BACE1
Verubecestat	Merck	BACE1
Phase 2		
AMG 510	Amgen	KRAS ^{G12C}
ASTX660	Astex	XIAP/cIAP1
AT7519	Astex	CDK1,2,4,5,9
AT9283	Astex	Aurora, JAK2
AUY-922	Vernalis/Novartis	HSP90
AZD5363	AstraZeneca/Astex/CR-UK	AKT
AZD5991	AstraZeneca	MCL1
CPI-0610	Constellation	BET
DG-051	deCODE	LTA4H
eFT508	eFFECTOR	MNK1/2
Indeglitazar	Plexxikon	pan-PPAR agonist
LY2886721	Lilly	BACE1
LY3202626	Lilly	BACE1
LY517717	Lilly/Protherics	FXa
MAK683	Novartis	PRC2 EED
Navitoclax (ABT-263)	Abbott	BCL-2/BCL _{XL}

Introduction

Onalespib	Astex	HSP90
PF-06650833	Pfizer	IRAK4
PF-06835919	Pfizer	KHK
Phase 1		
ABBV-744	Abbott	BD2-selective BET
ABT-518	Abbott	MMP-2 & 9
ABT-737	Abbott	BCL-2/BCL _{XL}
ASTX029	Astex	ERK1,2
AT13148	Astex	AKT, p70S6K, ROCK
AZD3839	AstraZeneca	BACE1
AZD5099	AstraZeneca	Bacterial topoisomerase II
BI 691751	Boehringer Ingelheim	LTA4H
ETC-206	D3	MNK1/2
GDC-0994	Genentech/Array	ERK2
HTL0014242	Sosei Heptares	mGlu5 NAM
IC-776	Lilly/ICOS	LFA-1
LP-261	Locus	Tubulin
LY2811376	Lilly	BACE1
Mivebresib	AbbVie	BRD2-4
Navoximod	New Link/Genentech	IDO1
PLX5568	Plexxikon	RAF
S64315	Vernalis/Servier/Novartis	MCL1
SGX-393	SGX	BCR-ABL
SGX-523	SGX	MET
SNS-314	Sunesis	Aurora

Indeed, FBDD is a first in line approach to assess protein-protein interaction, multicomplex targets, intrinsically disordered proteins and other unconventional targets in which classical screening of larger libraries (HTS) failed. One of the advantages of FBDD is represented by the possibility to explore much more chemical space compared to HTS. It was calculated that the possible number of small drug-like molecules can be around 10^{63} , more than the number of stars in the universe [44]. The current HTS libraries of about 10^6 small-molecules represent a very tiny fraction of the possible chemical space and it appears clear that to increase the chances of finding small molecules it will be necessary to screen libraries that contain billions of molecules with drug-size (30 non-hydrogen (heavy) atoms) [45]. A challenge quite difficult to undertake. Instead, in FBDD, since fragments contain less than 20 heavy atoms with a molecular weight less than 300 Da,

there are just over 166 billion possibilities for molecules having up to 17 heavy atoms [44]. While this is still a huge number, current FBDD screening collections (generally in the order of 10^3 compounds) provide much better sampling than the equivalent HTS collections.. Noteworthy, the chances of binding a fragment into a target are much higher compared to drug-like molecules since they do not have to fulfil as many interaction constraints. Indeed, the likelihood of compound binding goes up as its complexity goes down. Ligand Efficiency (LE) represents a good parameter to measure binding affinity per heavy atoms in order to classify the quality of binding [46, 47]:

$$LE = \frac{\Delta G_{binding}}{\#HA} = \frac{-RT \ln K_d}{\#HA}$$

where #HA is the number of heavy atoms, ΔG binding (Gibbs binding free energy), R the gas constant, T the absolute temperature and K_d the dissociation constant. It has been demonstrated that fragments can bind to all possible targets, even to those that have proven to be very difficult with traditional HTS approaches [47].

1.4.1 The fragment library

In contrast to HTS, the design of a target-direct library is not required and a small library (700-2000 fragments) can be screened for different targets. The most important criteria in the design of the library are the purity, the solubility and the stability of compounds. Avoiding artefacts from the library is the first rule to succeed in fragment-screening but is not straightforward. As demonstrated in a recent paper published in *J. Med. Chem.* by Dr Alessio Ciulli, Dr Helen Walden and co-workers at the University of Dundee [48], special attention has to be taken on metal contaminations that can mislead hit identification. The researchers conducted a screen against a ubiquitin-

conjugating enzyme, Ubet2T, combining different biophysical techniques: DSF, biolayer interferometry, two-dimensional (HSQC) protein-detected NMR experiments and ITC. These orthogonal techniques led to a promising hit with a K_d of 17 μM . After a SAR by catalogue process to improve potency, they realized there was something wrong since the new fragments were not giving the desired and expected results. Only after solving the crystal-structure the mystery was resolved. They found out that a metal, zinc, was binding to the protein. After addition of EDTA, a metal chelator, the activity of the fragment was lost and titration of zinc in ITC gave the same outcome previously observed. Therefore, before starting a fragment screening, the purity of the library must be confirmed.

Another important property to carefully consider is the solubility. Since fragments have weak binding affinity for the target (μM - mM range) high concentrations of fragments are required ($>500 \mu\text{M}$) and they must be soluble in a wide range of buffers at different conditions. Furthermore, it is crucial to avoid pan-assay interference compounds (PAINS) and known aggregators. PAINS are compounds containing substructures that give rise to apparent but artefactual activity in assays [49]. There is not a specific mechanism but in general it includes forming covalent adducts with the protein or producing hydrogen peroxide [45].

In summary, the use of a good library represents a good starting point to find hits that can be then optimized to create compounds with higher affinity to probe the biology of the target or to have the perfect lead compound to be then evolved in a drug-like molecule.

1.4.2 Fragment screening

Fragments generally bind to target protein with a weak affinity in the range of 0.1-10 mM. For that reason, it is of fundamental importance the choice of

a technique that is sensitive enough to detect such weak binding, robust to avoid false detection of fragment hits and with the throughput capacity. The key for the success of FBDD resides then in the careful selection of a biophysical screening cascade.

A typical cascade of a fragment library can be divided into a primary and secondary screening and finally in a fragment characterization step. In the primary screening, fragments with low-affinities can be identified (K_d in the range of mM or three digits micromolar). DSF, SPR or even ligand-based NMR (cocktails) can be selected as biophysical techniques for this primary screening. The primary screening usually reduces drastically the number of fragments of the library. Once a subset of fragments candidates has been identified to bind to the target, a second screening can be performed. The aim of this secondary screening is to confirm hits for the selected target. Usually ligand-based NMR (e.g. STD, CPMG, water-LOGSY) or SPR are applied. The confirmed hits can be then characterised by ITC to get high content information such as enthalpy (ΔH), entropy ΔS and stoichiometry (n). In parallel fragments can be progressed with X-ray crystallography or CSP-NMR. Structural information, if obtained, can be really useful for the optimization of the fragments. However, is not always possible to elucidate the binding mode and binding site of the fragments and therefore computational analysis can be implemented in the screening.

The hit rate of the screening could correlate with the library size and composition, but this hit rate strongly depends on how challenging (druggable) the target is. Hit rates of fragment-based programs could fall in the range of 3 to 30 % [50–53]. Each company or academic institution has implemented its own cascade, such as the case of Astex Pharmaceutical that is applying the Pyramid™ platform that integrates a range of high-

throughput biophysical techniques for screening such as X-ray crystallography, NMR, and calorimetry with fragment library design and computational methodologies. A standard screening cascade strategy is illustrated in Fig. 1.11.

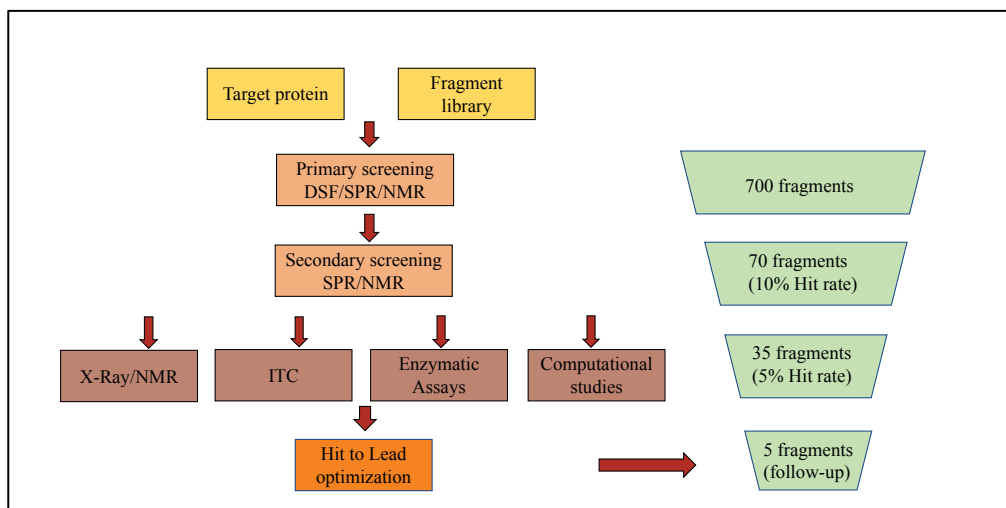


Figure 1.11. Example of a FBDD screening cascade and hit rates for each step.

As mentioned, an important point during the screening process is to detect false positives or false negatives. Each technique has its own drawback and can lead to false positive and false negatives. To limit the chances of bringing forward the wrong hit fragments it is of fundamental importance to identify the optimal combination of techniques in the cascade. One approach is to combine two or more screening techniques and follow-up only common hits. In this manner, the number of false positives will be reduced, but of course the number of false negatives will increase. An alternative strategy is to select a primary method that has been shown to work well for a particular target of interest and rely on this method until the compounds reach an affinity that allows them to be detected unequivocally with the same or other techniques [54]. There are not “ready- to-use” guidelines in the choice of the best screening cascade in FBDD, therefore a correct balance between cost and benefit, throughput and sensitivity, research goals and facilities access,

timeline and resources have to be taken.

1.4.3 Fragment optimization

Approaches employed for fragment optimization to achieve more potent compounds fall into three groups: fragment merging, linking and growing. Figure 1.12 illustrates the different fragment optimization strategies.

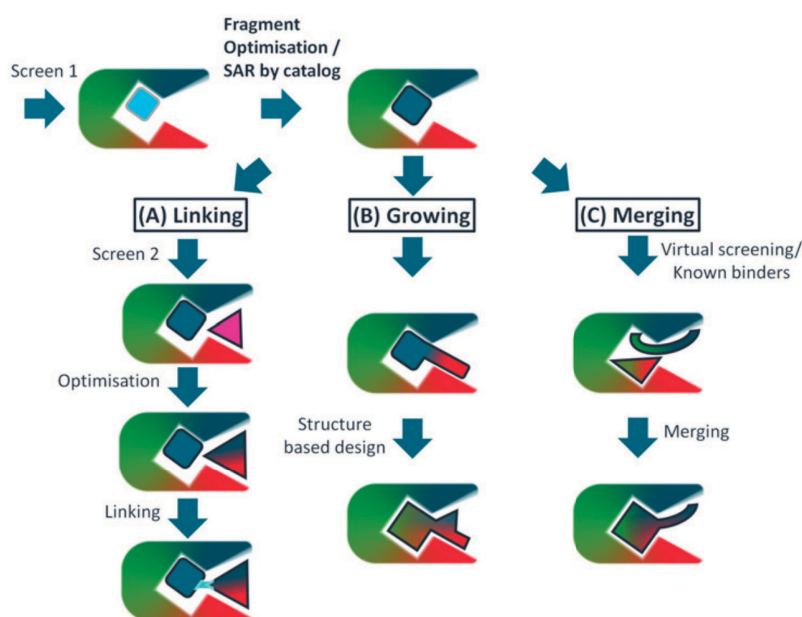


Figure 1.12. FBDD optimization strategies. Adapted from [13].

Fragment merging involves the incorporation of different regions of overlapping molecules into a single compound. In fragment linking, the separate molecules binding at non-overlapping sites in a proximity distance in the same cavity of the protein, are joined together to create a new bigger compound. A famous example of this strategy is the discovery of venetoclax a selective Bcl-2 inhibitor. Venetoclax is the culmination of a discovery program that was running for more than two decades. Dr Steve Fesik and his colleagues at Abbott published the X-ray and NMR structure of the protein Bcl-XL back in 1996. The exploratory “SAR by NMR” work was done on this protein, leading to ABT-263, which targets both proteins Bcl-XL and Bcl-2. Subsequent work revealed that a selective Bcl-2 inhibitor might be

preferable in some cases, and further medicinal chemistry led to venetoclax. This drug illustrates the power of fragments to tackle a difficult target by accessing unusual chemical space (Fig 1.13) [55].

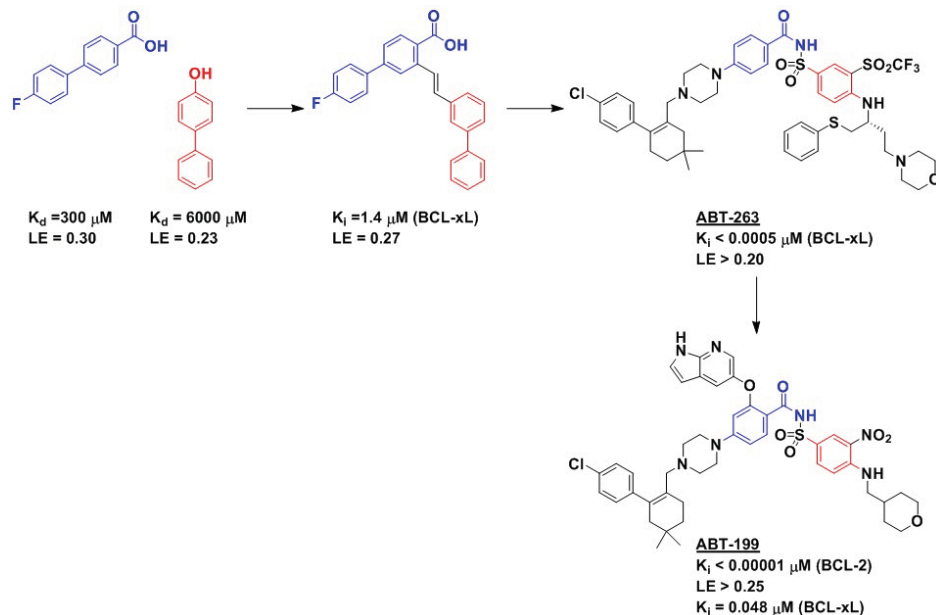


Figure 1.13. Venetoclax development: from the fragment identification to the final drug.

Fragment growing is the most applied method in FBDD, where molecules are enlarged using chemical synthesis to include additional binding contacts to fit the cavity of the target and to improve drug-like properties [13]. The discovery of vemurafenib, a selective inhibitor of the B-Raf V600E mutant kinase, is a good example of this approach [56].

Computational methods can also aid to find less obvious drug fragments candidates and evolve them in drug candidates. There are two primary strategies in computational drug design. The first (structure-based) is to assess the fit of a molecule to the binding site through target-based scoring functions. The second (ligand-based) is to determine how well a ligand can recapitulate features of known active binders, evaluated by ligand-based scoring functions. Structure-based methods can maximize the possibility of finding binders by exploiting experimentally obtained or modelled structural information. However, when a structure is not

available, ligand-based methods have also shown to be successful. Programs like PINGUI [57], ALTA [58] or DOTS [59] can be used for a fragment to lead optimization.

Our group is also developing a fragment evolution platform able to generate new chemical structures. This platform allows to explore the chemical space, given the binding mode of a fragment, and is able to produce non-obvious fragments that maintain or improve affinity for the target, but raising the possibility to obtain new entities as starting point for a lead optimization or with better physicochemical properties compared to the identified fragment. Validation of the platform will be discussed in chapter 6.

Chapter 2

Objectives

Objectives

2.1 General Objective

Drug discovery has seen huge changes over the past 20 years, with the advent of new technologies constantly changing the approach to drug discovery and design and positioning biophysical techniques as the fundamental tools to apply for a successful program. In this scenario, using biophysical techniques, the main objective of this work is the identification and characterization of small molecules with the ability to bind specific proteins.

2.2 Specific objectives

- 1) Fbw7 is an E3 ligase with an important role in cancer. However, until now no Fbw7 small molecules ligands have been identified. In this thesis we aim to **perform a FBDD program in order to identify fragments able to bind to this E3 ligase**. These fragments could be employed as starting points to elucidate the best strategy to target Fbw7 and to build novel PROTAC molecules.
- 2) Due to the poor aqueous solubility of retinoids, evolution has tuned their binding to cellular proteins to address specialized physiological roles by modulating uptake, storage, and delivery to specific targets. In this thesis, **we aim to disentangle the structure–function relationships of these protein class and disclose clues for engineering selective carriers**. The binding mechanism of one of most abundant retinol-binding isoforms and its mutant form will be elucidated.
- 3) Given the binding mode of fragments to a target of interest, optimization of the fragments can be laborious, difficult and time consuming. In this thesis, **small molecules binding to Brd4(BD1), identified by the automated fragment evolution platform developed by our group, will be assayed and the platform validated**.

Objectives

Chapter 3

Methods and Materials

3.1 General Materials

3.1.1 Reagents

General reagents were purchased from suppliers Sigma-Aldrich and Fisher Scientific. Mini-PROTEAN TGX precast gels as well as premixed electrophoresis running buffers were purchased from Bio-Rad. Protein ladder was Precision Plus Protein™ Dual Colour Standards from Bio-Rad. Coomassie Brilliant Blue was used to stain the protein gels. Tobacco etch virus (TEV) protease and thrombin protease were purchased from Invitrogen and Sigma-Aldrich, respectively.

3.1.2 Bacterial Strains and Growth Media

XL1 blue competent cells (Agilent) and MC062 cells were used for DNA preparation. *E. coli* BL21 and Rosetta (DE3) were used for the preparation of competent cells and were a kind donation of Dr Raimon Sabaté Lagunas. Luria-Bertani broth, also named lysogeny broth (LB) and agar were purchased from Sudelab supplier. M9 medium for ¹⁵N labelling was prepared in house following recipe describe in paragraph 3.6.

3.2 Molecular Cloning

3.2.1 Protein Constructs and Expression Plasmids

The proteins used for this work (unless mentioned otherwise) are given in table 3.1. The gene encoding Fbw7-Skp1 complex (Fbw7²⁶³⁻⁷⁰⁷-Skp1) in a pABLOmut vector is a kind donation of Dr Bing Hao (UConnHealth, USA). In this plasmid, Skp1 contains two internal deletions [60]. Plasmid is designed to express the protein of interest with a GST-tag and thrombin protease cleavage site and encodes gene conferring ampicillin resistance.

The genes encoding rat CRBP-I and CRBP-II constructs in a PV614375 and PV614379 plasmid, respectively, were provided by Genomics-Online for expression with a N-terminal His6-GST tag and a TEV protease cleavage site. Plasmid PV614379 containing rat CRBP-II L78→I construct was generated in house using QuickChange II site-directed mutagenesis kit (Agilent). Both plasmids encode gene conferring resistance to kanamycin. Finally, the gene encoding Brd4(BD1) with a His6-GST tag and kanamycin resistance site was a kind donation of Dr Alessio Ciulli (Dundee University, UK).

3.2.2 Kits

QuickChange II site-directed mutagenesis kit for the site-directed mutagenesis was purchased from Agilent. GeneJet plasmid miniprep kit for DNA amplification was purchased from ThermoFisher Scientific.

3.2.2.1 Site-directed mutagenesis (CRBP project)

Site-directed mutagenesis was performed using QuickChange II site-directed mutagenesis kit (Agilent). Primers were designed following the QuickChange manual and oligonucleotides were synthesized, desalted, purified and lyophilized. The polymerase chain reactions (PCR) were performed on 9700 GeneAmp® PCR System (Applied Biosystems®). Upon digestion of the parental DNA strands by Dpn1 restriction enzyme, the PCR product was transformed into XL1-Blue competent cells and grown on LB agar plates containing kanamycin (50 µg/ml) at 37 °C for 24-36 h. Single colonies were then picked from agar plates and grown for 12h-18 h in 10 ml of LB medium containing kanamycin (50 µg/ml). DNA was subsequently extracted and purified using GeneJet plasmid miniprep kit (ThermoFisher Scientific). Purified DNA was sequenced to confirm the presence of the

mutation by the Genomics Service (CCTiUB) of the University of Barcelona.

3.3 Preparation of Competent *E.coli* cells

A single colony of competent *E. coli*, grown in an agar plate (without antibiotic) the night before, was picked and added into 15 ml LB-broth (without antibiotic) in a 50 ml sterile falcon tube and shaken at 37 °C for 4 h at 170 rpm. When the bacterial culture reached an OD600 value of approx. 0.6, the falcon was removed from the shaker and cooled on ice. After 10 min, the bacteria culture was transferred to a pre-chilled centrifuge tube and the cells were harvested by centrifugation at 6000 rpm, 4 °C for 3 min. The cell pellet was re-suspended in 10 ml ice cold 0.1 M CaCl₂ and incubated for 20 min. The cells were once again harvested by centrifugation at 6000 rpm, 4 °C for 3 min. That new cell pellet was resuspended in 5 ml ice cold 0.1 M CaCl₂ containing 15% (v/v) glycerol and the cells were dispensed as 200 µl aliquots into micro-centrifugation tubes and stored at -80 °C until further use.

3.4 Transformation of *E.coli* cells

Competent *E. coli* cells were thawed in ice for 30 minutes. Once thawed, 50 µl were mixed with 5 ng of vector-DNA. The mixture was kept on ice for 30 min, followed by a heat shock for 1 min at 42 °C and then an immediately 5 min incubation on ice. Transformed *E. coli* cells were incubated for 1 h at 37 °C with 0.5 ml LB containing no antibiotics. After this, the *E. coli* cells were spread on LB-agar plates containing the appropriate antibiotic depending on the plasmid resistance (ampicillin for Fbw7-Skp1 construct, kanamycin for Brd4(BD1) and rat-CRBP constructs). Subsequent to over-night incubation at 37 °C, *E. coli*

transformants were visible as colonies.

3.5 Determination of DNA concentration, protein concentration and cell growth

DNA has an UV-light absorption maximum at 260 nm due to the aromatic rings of its bases. Proteins have an UV-light absorption maximum at 280 nm due to tryptophan. Therefore, DNA concentration and protein concentration were determined by measuring UV-light absorption at 260 nm and 280 nm, respectively, using a NanoDrop spectrophotometer (Thermo-Scientific). Cell growth was determined by measuring the amount of light scattering by microbial cells (Optical Density) at 600 nm (OD600), using a NanoDrop spectrophotometer (Thermo-Scientific).

3.6 Preparation of M9 medium broth (Fbw7-Skp1 project)

For 1 L of M9 medium, 6 g of Na_2HPO_4 , 3 g of KH_2PO_4 , 0.5 g of NaCl and 1 g of $^{15}\text{NH}_4\text{Cl}$ were mixed, the pH adjusted to 7.4 and the solution obtained autoclaved. Right after the inoculations of cells, 20 ml of 20% filter-sterile glucose, 1 ml 0.1 M filter-sterile CaCl_2 , 1 ml 1 M filter-sterile MgSO_4 , 1 ml 10 mg/ml filter-sterile biotin, 1 ml 10 mg/ml filter-sterile thiamine and 1 ml 100 mg/ml ampicillin were added. Furthermore, in order to boost the expression of the protein 1 ml of a micronutrient solution was added: Q solution (40 mM HCl, 25 mM $\text{FeCl}_2 \cdot 4\text{H}_2\text{O}$, 1.25 mM $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$, 1 mM H_3BO_3 , 0,076 mM $\text{CoCl}_2 \cdot 2\text{H}_2\text{O}$, 2.5 mM ZnCl_2 , 0.023 mM $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$, 3mM $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$, 0.2 mM $\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$) kind donation of Dr Xavier Salvatella Lab, IRB Barcelona

3.7 Protein Expression

3.7.1 Fbw7-Skp1 complex

For a typical 4 L *E. coli* protein expression experiment, 4 flasks, each containing 1 L of LB-media were autoclaved the day before of the starting of the experiment. In parallel, 10 ml LB ampicillin-media was inoculated with a single colony from a LB-ampicillin-agar plate and incubated over-night at 37 °C and 150 rpm. The next day, the autoclaved LB-media was supplemented with ampicillin and was inoculated with a ratio of 1 to 100 of the over-night starting culture. The flasks were shaken at 37 °C and 180 rpm until the *E. coli* cell culture density reached an OD600 of 0.6 to 0.8. At this point protein expression was induced with 1 mM IPTG. After induction, the flasks were shaken at 18 °C and 180 rpm for 12 to 18 h. The *E. coli* cell cultures were then centrifugated at 8000 rpm for 30 min. The cell pellets were stored at –20 °C or used as source of protein purification.

3.7.2 Fbw7-Skp1 complex ¹⁵N labelled-M9 Marley expression

For a typical 4 L *E. coli* protein expression experiment, 4 flasks, 3 containing 1 L of LB-media each and 1 containing M9 broth were autoclaved the day before starting the experiment. In parallel, 10 ml LB ampicillin-media was inoculated with a single colony from a LB-ampicillin-agar plate and incubated over-night at 37 °C by 150 rpm. The next day, the autoclaved LB-media was supplemented with ampicillin and was inoculated with a ratio of 1 to 100 of the over-night starting culture. The flasks were shaken at 37 °C and 180 rpm until the *E. coli* cell culture density reached an OD600 of 0.6 to 0.8. The *E. coli* cell cultures were then centrifugated at 8000 rpm for 30 min. The cell pellets were added to the 1 L M9 medium broth and the flask

was shaken at RT at 180 rpm for 1 h. Protein expression was induced with 1 mM IPTG and after induction, the flask was shaken at 18 °C for 12 to 18 h. The *E. coli* cell culture was then centrifugated at 8000 rpm for 30 min. The cell pellets were stored at –20 °C or used as source of protein purification.

3.7.3 Rat-CRBP proteins

For a typical 4 L *E. coli* protein expression experiment, 4 flasks, each containing 1L of LB-media were autoclaved the day before starting the experiment. In parallel, 10 ml LB kanamycin-media was inoculated with a single colony from a LB-kanamycin-agar plate and incubated over-night at 37 °C by 150 rpm. The next day, the autoclaved LB-media was supplemented with kanamycin and was inoculated with a ratio of 1 to 100 of the over-night starting culture. The flasks were shaken at 37 °C and 180 rpm until the *E. coli* cell culture density reached an OD600 of 0.6 to 0.8. Protein expression was induced with 1 mM IPTG. After induction, the flasks were shaken at 18 °C at 170 rpm for 12 to 18 h. The *E. coli* cell cultures were then centrifugated at 8000 rpm for 30 min. The cell pellets were stored at –20 °C or used as source of protein purification.

3.7.4 Brd4(BD1) protein

For a typical 4 L *E. coli* protein expression experiment, 4 flasks, each containing 1L of LB-media were autoclaved the day before starting the experiment. In addition, 10 ml LB kanamycin-media was inoculated with a single colony from a LB-kanamycin-agar plate and incubated over-night at 37 °C by 150 rpm. The next day, the autoclaved LB-media was supplemented kanamycin and was inoculated with a ratio of 1 to 100 of the over-night starting culture. The flasks were shaken at 37 °C and 180 rpm until the *E. coli* cell culture density reached an OD600 of 2.5. Protein expression was

induced with 0.4 mM IPTG. After induction, the flasks were shaken at 18 °C at 170 rpm for 12 to 18 h. The E. coli cell cultures were then centrifugated at 8000 rpm for 30 min. The cell pellets were stored at –20 °C or used as source of protein purification.

3.8 Cell lysis

Cell lysis is a method that disrupts cells and leads to the release of their proteins into the lysis buffer. Recombinant cell pellets were re-solubilized in 50 ml of Buffer A for Fbw7-Skp1, Buffer E for Brd4(BD1) and Buffer G for rat CRBPs (table 3.1) implemented with Pierce protease inhibitor cocktail (ThermoFisher Scientific) and lysed by sonication at 19 °C for 2 min with a break of 20 seconds every 10 seconds of lysis, followed by double centrifugation (8600 rpm, 4 °C, 30 min) to remove insolubilities. The supernatant was then collected for protein purification.

Table 3.1. General Buffers for purification

LOADING	BUFFER	PROTEIN BUFFER
Buffer A	Fbw7-Skp1	50 mM Hepes pH= 8.0, 200 mM NaCl, 5 mM DTT
Buffer B	Fbw7-Skp1	50 mM Hepes pH= 8.0, 200 mM NaCl, 5 mM DTT, 20 mM Gluthatione
Buffer C	Fbw7-Skp1	50 mM Hepes pH= 8.0, 50 mM NaCl
Buffer D	Fbw7-Skp1	50 mM Hepes pH= 8.0, 1000 mM NaCl
Buffer E	Brd4(BD1)	50 mM Hepes pH= 7.4, 150 mM NaCl, 30 mM Imidazole, 2 mM B- mercaptoethanol
Buffer F	Brd4 (BD1)	50 mM Hepes pH= 7.4, 150 mM NaCl, 250 mM Imidazole, 2 mM B-mercaptoethanol
Buffer G	CRBPs	50 mM Tris/HCl pH=8.0, 250 mM NaCl, 5 mM imidazole
Buffer H	CRBPs	50 mM Tris/HCl pH=8.0, 250 mM NaCl, 250 mM imidazole

Methods and Materials

Buffer I	CRBPs	50 mM Tris/HCl pH= 8.0, 50 mM NaCl
Buffer L	CRBPs	50 mM Tris/HCl pH= 8.0, 1000 mM NaCl

3.9 Purification

Chromatographic methods were performed using a ÄKTA start system (GE Healthcare, Uppsala, Sweden).

3.9.1 Fbw7-Skp1 complex and Fbw7-Skp1¹⁵N labelled

After cell lysis, the supernatant was filtered using Millex-AA Syringe Filter of 0.8 µm provided by Sigma-Aldrich and was applied to a 5 ml GSTrap HP column (GE Healthcare) and washed with Buffer A (table 3.1). Bound protein was eluted with 20 mM glutathione (Buffer B table 3.1) and collected. Proteins were identified by SDS-PAGE. Thrombin protease was added to the fraction and dialyzed overnight at 4 °C against buffer A using a 3.5 kDa dialysis membrane (Spectra/Por). To remove contaminations (uncleaved protein), a second GST purification was performed as described above. The flow-through was collected and dialyzed against Buffer C (table 3.1). Dialyzed solution was then loaded onto 5 ml anion exchange Hitrap Heparin (GE Healthcare). The protein was collected from the flow-through of the purification (Buffer C table 3.1) and identified by SDS-PAGE. The mass and purity were subsequently verified by mass spectrometry at the Molecular Characterization and Mass Spectrometry service (CCTiUB) of the University of Barcelona.

3.9.2 Rat-CRBPs proteins

After cell lysis, the supernatant was filtered using Millex-AA Syringe Filter of 0.8 μm provided by Sigma-Aldrich and was applied to a 5 ml HisTrap HP column (GE Healthcare) and washed with Buffer G (table 3.1). Bound protein was eluted using Buffer H (table 3.1) with a linear gradient from 5% to 100%. Fractions containing proteins were identified by SDS-PAGE and pooled. His-tagged TEV protease was added to the pooled fractions and dialyzed overnight at 4 °C against Buffer G (table 3.1) using a 3.5 kDa dialysis membrane (Spectra/Por). To remove contaminations (uncleaved protein) and the His-tagged TEV protease, a second HisTrap purification was performed as described above. The flow-through was collected and dialyzed against Buffer I (table 3.1). The protein sample was then loaded onto 5 mL anion exchange Hitrap Q (GE Healthcare) equilibrated and the proteins were eluted with Buffer L (table 3.1) with a NaCl-gradient (50 mM NaCl to 1 M NaCl). Fractions containing proteins were identified by SDS-PAGE and pooled. The mass and purity were subsequently verified by mass spectrometry at the Molecular Characterization and Mass Spectrometry service (CCTiUB) of the University of Barcelona.

3.9.3 Brd4(BD1)

After cell lysis, the supernatant was applied to a 5 mL HisTrap HP column (GE Healthcare) and washed with Buffer E (table 3.1). Bound protein was eluted using Buffer F (table 3.1) with a linear gradient from 5% to 100%. Fractions containing proteins were identified by SDS-PAGE and pooled. Mass and purity were subsequently verified by mass spectrometry at the Molecular Characterization and Mass Spectrometry service (CCTiUB) of the University of Barcelona.

3.10 Fragment Library

The fragment library was generated in house with the collaboration of Dr Sergio Ruiz (Barril Lab). The providers selected to build-up the library were Specs, Asinex and Princeton Biomolecular based on prices, number of compounds, easiness of order and shipping compared to other providers. LigPrep calculations (Schrödinger) were run and the output of LigPrep was used as input for QikProp (Schrödinger) and MOE (Chemical Computing group) to obtain physico-chemical properties of the compounds. The compounds were filtered according to the following features:

- $140 \leq \text{molecular weight} \leq 300$
- $-2 \leq \log P \leq 3$
- $\log S \geq -3$
- number of rotatable bonds ≤ 6
- hydrogen bond donor ≤ 3
- hydrogen bond acceptor ≤ 7
- topological surface area ≤ 70

After the compounds were filtered, using a MOE script the selected compounds were clustered by Tanimoto fingerprint and the centroids were selected, as the most representative compounds. The final fragment library obtained is composed by 700 fragments. Fragments contained within clinically evaluated compounds appear to have greater 3D conformations than do those in fragment libraries, which are all biased towards molecules with limited 3D shape. Building libraries around fragments with greater 3D is an area of intense current debate. The three-dimensionality in compound selection was calculated for the final selected compounds. The principal moment of inertia (PMI) is used to describe the “low resolution” shape of a molecule in the form of 2D

triangular plot. The PMI parameters of a molecule, normalized to the molecule's longest dimension, display rod-shaped molecules in the top left, disk-shaped at the bottom and spherical the top right of the triangle.

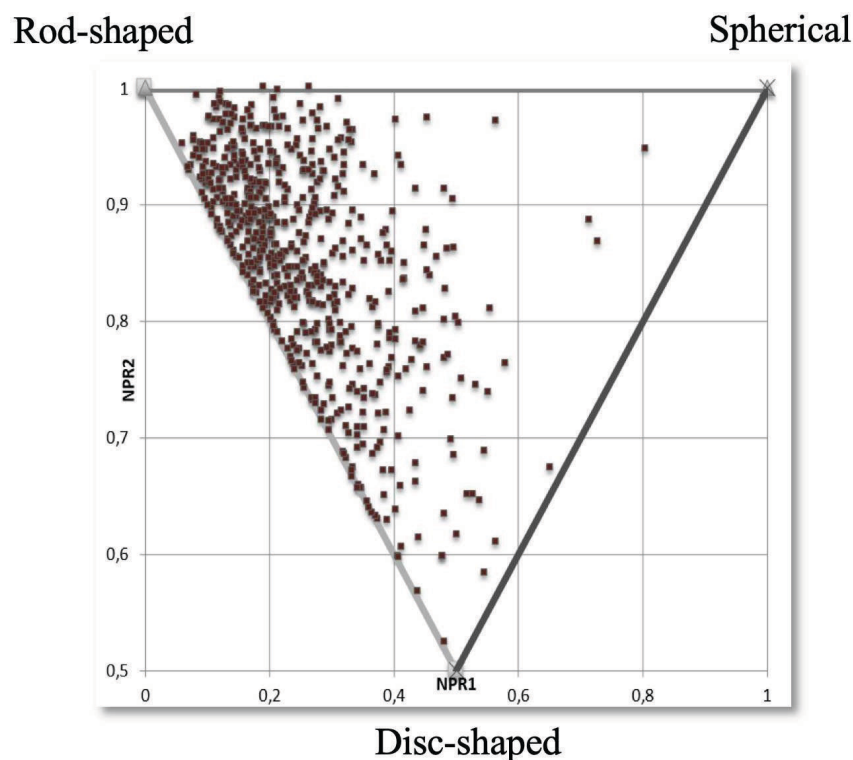


Figure 3.1. PMI of the fragment library.

3.11 Biophysical Techniques

3.11.1 DSF experiments

DSF was used to screen Fbw7-Skp1 complex against the in-house fragment library. The stock solutions were prepared at 100 mM concentration in 100 % DMSO from the commercial compounds. The SYPRO Orange dye (Sigma-Aldrich) was purchased as 5000x stock concentration and was diluted 1000x in the protein solution. The protein sample was then dispensed into the wells of a 96 well PCR plate (Thermo Scientific), followed by the addition of the fragments solution to give the final volume of 25 μ l in each well. The final protein concentration in the screen was 5 μ M and the final

volume 5 μl of Fbw7-Skp1 complex was screened at one single concentration of fragments 1 mM, 5% DMSO. The screen was run using a LightCycler 480 PCR instrument (Roche Life Science) at the Genomics Service (CCTiUB) of the University of Barcelona. The fluorescence intensity was monitored with excitation and emission wavelengths at 465 nm and 580 nm, respectively. Plates were scanned from 20 °C to 85 °C with a heating rate of 0,6 °C min^{-1} . Raw data were extrapolated using the LightCycler 480 software. Wells containing protein solution and dye were used as the reference T_m . The melting curves and thermal shifts were visualized using Microsoft Excel spreadsheet.

3.11.2 SPR experiments

All SPR studies were performed at 25 °C using Biacore T200 instruments (GE Healthcare) located at the Scientific and Technical Services of the University of Barcelona(CCTiUB). Prior to the start of each new experiment, the system was cleaned using Biacore maintenance kit and a fresh CM5 or CM7 chip was inserted, which was preconditioned and normalized according to the instructions displayed by the software. Proteins were immobilized on CM7 (Fbw7-Skp1 complex) or CM5 (CRBP proteins) sensor chip (GE Healthcare) using standard amine coupling procedure. The carboxymethyl dextran matrix of the sensor chip was activated with 0.1 M N-hydroxysuccinimide (NHS) and 0.4 M 1-ethyl-3-(3-(dimethylamino)propyl) carbodiimidehydrochloride (EDC) at a flow rate of 15 $\mu\text{l}\cdot\text{min}^{-1}$ for 7 min. The immobilization was performed with 10 $\mu\text{g}\cdot\text{ml}^{-1}$ protein in 10 mM sodium acetate at pH= 5.0 at a flow rate of 5 $\mu\text{l}\cdot\text{min}^{-1}$. It was performed using Phosphate Buffered Saline (PBS, 10 mM Phosphate, pH= 7.4, 150 mM NaCl) as immobilization running buffer. Unreacted activated groups of the dextran matrix were deactivated by injection of 1 M ethanolamine

hydrochloride for 7 min. For GST immobilization, Biacore GST Capture Kit Instructions were followed. Anti GST-antibody was immobilized at a flow rate of $5 \mu\text{l}\cdot\text{min}^{-1}$. Recombinant GST and protein were passed on the surface of the chip for every cycle, followed by regeneration of the surface of the chip with glycine-HCl, pH= 2.2.

3.11.2.1 Fbw7-Spk1 SPR screening

The SPR screen was performed using in-house fragment library, which contain 700 compounds. Protein immobilization was performed according to GE guidelines:

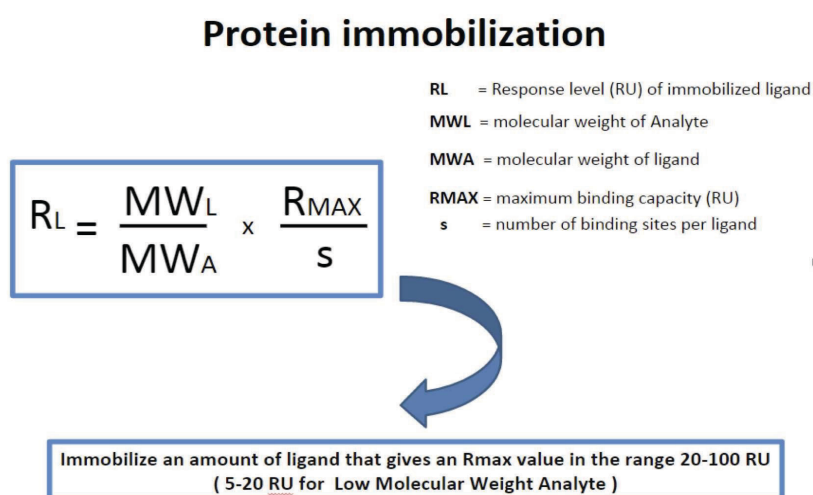


Figure 3.2 Formula applied to calculate the amount of protein to immobilize on the surface of the chip taking in account the molecular weight of the fragments.

In order to get a good signal/noise ratio and to have experimental duplicates, we decided to immobilize different amount of protein in two channels:

- *Low density* channel: 7000 RU to get a maximum response of 25 RU
- *High density* channel: 16000 RU to get a maximum response of 55 RU

The *High-density* channel resulted the best conditions to perform the

Methods and Materials

experiment. Additionally, channel 1 was taken as reference surface where no protein is presented but where the immobilization protocol was followed and finally, channel 4 was left empty to discriminate from unspecific dextran binding. Table 3.2 outlines the main parameters employed in immobilizing the Fbw7-Skp1 complex on the sensor chip's surface for the fragment screening.

Table 3.2. *Main parameters for protein immobilization*

Sensor chip	CM7
Channel 1	Reference surface (no protein, same immobilization protocol)
Channel 2	Low density channel (9000 RU)
Channel 3	High density channel (16000 RU)
Channel 4	Empty channel (to check unspecific dextran binding)
Flow rate	5 $\mu\text{l}\cdot\text{min}^{-1}$
Fbw7-Skp1 concentration	10 $\mu\text{g}\cdot\text{ml}^{-1}$
Protein buffer	50 mM Hepes pH=8.0, 50 mM NaCl
Immobilization Buffer	PBS, 10 mM Phosphate, pH=7.4, 150 mM NaCl

As primary screening, 10 μl of each fragment at 10 mM concentration in 100% DMSO were dispensed into 96 well plates. 190 μl of the sample preparation buffer were added later to give a final concentration of 500 μM 5% DMSO. The plates were then sealed and vortexed in a centrifuge to get solutions to the bottom of the wells and remove the air bubbles. The main solutions used to set up the fragment screen are presented in table 3.3.

Table 3.3 *SPR-Fragment screening Buffers*

Running buffer	1x PBS, 10 mM Phosphate pH=7.4, 150 mM NaCl, 0.05% tween-20, 5% DMSO
Sample preparation buffer	1.05x PBS, 10.5 mM Phosphate pH=7.4, 157.5 mM NaCl, 0.05% tween-20, 5% DMSO
Solvent correction samples	Sample preparation buffer containing 3%, 3.7%, 4.4%, 5.1%, 5.8%, 6.6%, 7.3%, and 8% DMSO
Preparations of fragments	Stock solutions at 10 mM in 100% DMSO were diluted to 500 μM in sample preparation buffer

3.11.2.2 Fragment hit characterization

Titration experiments were performed to confirm and characterize the hit fragments obtained from the single concentration screening. For determining the K_d of the fragments a 2-fold dilution series starting from 500 μM were analysed. A flow rate of 60 $\mu\text{l}\cdot\text{min}^{-1}$ was used, and the association was monitored for 60 s, while the dissociation time was monitored for 180 s. 100 mM stock solutions were prepared by dissolving the fragments (powder) in 100% DMSO. These stock solutions were diluted to 10 mM in 100% DMSO and then used to prepare several dilutions 100% DMSO from which final concentrations were obtained adding sample preparation buffer, resulting in 500 μM , 250 μM , 125 μM , 62.5 μM , 31.25 μM , 15.62 μM solutions. For data analysis, the Biacore T200 evaluation software 2.0 (GE Healthcare) was used. Signals were corrected for non-specific binding to the surface by subtracting signals from the reference surface from those of the protein surfaces. Reference surface interactions with DMSO were performed by using a series of solvents standard (solvent corrections). Moreover, signals were corrected from background by subtracting signals from an average of two blanks injections from those of compound injections (blank subtraction). The binding affinity was estimated by fitting the data to a single site interaction model and the steady-state values were extracted from the sensograms and plotted against the concentration. Furthermore, to increase the chances of finding positives hit, fragments that gave a $K_d < 200 \mu\text{M}$ were assessed by screening them through a chip surface covered by Fbw7-Skp1 complex immobilized with the GST-tag. Results were corrected and analysed as previously described for covalent immobilization.

3.11.3 Retinol Kinetic affinity characterization

Interaction assays were performed in running buffer consisting of 1x Phosphate Buffered Saline (PBS, 10 mM Phosphate pH=7.4, 150 mM NaCl), 0.05% (v/v) tween-20, 5% (v/v) DMSO. Retinol was dissolved as 50 mM stock solution in pure DMSO and diluted with 1.05x PBS 0.05% (v/v) tween-20. A flow rate of 60 $\mu\text{l}\cdot\text{min}^{-1}$ was used, and the association was monitored for 90 s, while the dissociation time was recorded for 240 s. For data analyses, the Biacore T200 evaluation software 2.0 (GE Healthcare) was used. Signals were corrected for non-specific binding to the surface by subtracting signals from a reference surface (same immobilization procedure without proteins) from those of the CRBPs surfaces (reference subtraction). In addition, corrections for minor differences between CRBPs surfaces and reference surface interactions with DMSO were performed by using a series of solvent standards (solvent correction). Moreover, signals were corrected for background by subtracting signals from an average of two blank injections from those of compound injections (blank subtraction). For determining the K_d of retinol a 2-fold dilution series starting from 250 μM was analysed. The binding affinity was estimated by fitting the data to a single site interaction model and the steady-state values were extracted from the sensograms and plotted against the concentration. The binding kinetics was estimated by taking in account the k_{on}/k_{off} ratios.

3.11.4 NMR experiments

NMR was selected as an orthogonal technique to confirm fragment hits and to check Fbw7-Skp1 folding and stability in the absence and in the presence of DMSO. A Bruker 600 MHz NMR spectrometer equipped with a cryoprobe of triple resonance TCI has been used to perform the

experiments with the help of Drs Margarida Gairi and Teresa Gonzalez of the NMR unit (CCiTUB) of the University of Barcelona.

3.11.4.1 1D-NMR of Fbw7-Skp1 complex

Proton spectra of Fbw7-Skp1 complex were recorded to ensure protein folding and stability during time. Standard pulse sequences were used and water suppression was achieved using the “Excitation Sculpting” scheme. Datasets were obtained after 1024 scans. 200 μ l of Fbw7-Skp1 complex, 40 μ M in PBS (50 mM NaCl, 94 mM K_2HPO_4 , 6 mM KH_2PO_4 , pH= 8.0) were transferred to a 2 mm NMR-tube. A first spectrum was recorded after water suppression and 5% DMSO final volume was added lately to confirm stability in the presence of DMSO. Furthermore, two spectra were recorder after 24 and 48 h incubation in the presence of DMSO.

3.11.4.2 Saturation transfer difference (STD)

Standard pulse sequences were used for 1D and STD data acquisition. Water suppression was achieved using the “Excitation Sculpting” scheme. STD-NMR experiments were carried out at 25 °C. The 4 s irradiation period consisted of a train of 50 ms gaussian pulses separated by a 1 ms delay on-resonance (-1 ppm) and off-resonance (50 ppm). A 40 ms T2 filter was used to suppress residual protein signals. STD datasets were obtained after 512 or 1024 scans. NMR data were processed in the TOPSPIN version 3.6.1 (Bruker Biospin). Hits coming from SPR were evaluated using STD-NMR. Different experiments were performed to get the optimal experimental conditions: protein concentration, % of D_2O , fragment concentrations and detergent concentrations. The best conditions were obtained with the following conditions. The stock fragment solutions were prepared in DMSO- d_6 at a 2 mM concentration. 12.8 μ l of Fbw7-Skp1 complex at 34.3 μ M in

PBS (pH = 8.0, 50 mM NaCl, 94 mM K₂HPO₄, 6 mM KH₂PO₄) was mixed with 319.7 μ l of D₂O. The resulting solution was then mixed with 17.5 μ l of the fragment stock solution to reach a final volume of 350 μ l. The final solution was then transferred to a 5 mm Shigemi tube (Cortenect). The protein and fragment concentration were 1.25 μ M and 100 μ M, respectively, in 90% D₂O, 5% DMSO. For detergent-based STD-NMR, 3.5 μ L of a 5x stock tween-20 solution was added in a final volume of 350 μ l to get a final concentration of 0.05% (maximum amount allowed to break aggregations). In parallel, a control without protein was prepared in order to discard false positives.

3.11.5 Fluorescence Polarization (FP) assay

The fluorescence polarization assays were performed on a ClarioStar reader (BMG LabTech) in NUNC 384-well plates using the excitation module 482-16 and the dichroic LP 504 filters and emission 530-40 filter. A fluorophore probe, a 16-mer diphosphorylated peptide (PEVPPpTPPGpSHSAFTK(FICT)) of DISC1, a natural substrate of Fbw7-Skp1, and its unlabelled form (PEVPPpTPPGpSHSAFTK) was purchased on demand. For the positive control, each well solution (20 μ l) contained a final concentration of 50 nM of the labelled natural substrate of the protein, 1% DMSO and serial protein (Fbw7-Skp1) concentrations from 28 μ M to 190 nM. The blank wells were prepared in the same way, but the protein was not added to the solution. For the competitive assay, the wells contained 20 μ l of a final concentration of 50 nM FITC-Disc1, 1 μ M of the unlabelled natural substrate of the protein (Disc1), 1% DMSO and serial protein concentrations from 28 μ M to 190 nM. The blank wells were prepared in the same way, but protein was not added to the solution. For the competitive assay with the confirmed fragment hits, the wells

contained 20 μl of a final concentration of 50 nM FITC-Disc1, 10 μM Fbw7-Skp1 complex and several concentrations of fragments: 0 μM – 6.25 μM – 12.50 μM – 25 μM – 50 μM – 100 μM – 200 μM – 400 μM – 800 μM at 8% DMSO. The blank wells were prepared in the same way, but protein was not added to the solution. The buffer used for these experiments was 50 mM Hepes pH= 8.0, 50 mM NaCl, 0.01% tween-20, 1 mM DTT. All the solutions were centrifuged at 0.2 rpm for 30 seconds and incubated at room temperature for 30 min after preparation before testing. Data was obtained in duplicate by Mars software (BMG LabTech) provided by the manufacturer and the polarization value obtained were plotted against concentration to get the corresponding dose-response curves. The dissociation constants (K_d) were calculated from these dose-response curves for each individual experiment.

3.11.6 TR-FRET assay

The TR-FRET binding assay was used to measure the disruption of interaction of Brd4(BD1) with [Lys(Ac)5/8/12/16]-Histone H4 (1-21)-GGK(Biotin) (Eurogentec) with several compounds, accordingly to a previously described protocol [54]. The experiments were performed in 384-well black plates (NUNC) in a buffer composed of 50 mM Hepes, pH= 7.4, 50 mM NaCl, 400 mM KF (Sigma), 0.5 mM CHAPS (Sigma), and 0.05% BSA (Sigma) in a final volume of 100 μl . Brd4(BD1) (100 nM) protein domain was mixed with 200 nM concentration of the biotin-labelled histone peptides and incubated for 30 min at room temperature. Two nM Eu³⁺ cryptate-conjugated streptavidin (CisBio) and 10 nM anti-His6-XL665 (Cisbio) were added to the mix and incubate for 30 min. Compounds were prepared in 250 mM Hepes pH= 7.4, 250 mM NaCl 20% DMSO at a single concentration and added to the solution to give a final concentration of 50

μM, 1% DMSO in a total well volume of 100 μl and incubated for 1 h. Compounds showing inhibition were then screened at several serial dilution concentrations. Inhibition (%) at 50 μM was calculated with the following formula:

$$100 \times \left[1 - \frac{\text{competitor FRET ratio } \frac{665}{620} - \text{FRET ratio } \frac{665}{620} \text{ of the negative control}}{\text{FRET ratio } \frac{665}{620} \text{ of the positive control} - \text{FRET ratio } \frac{665}{620} \text{ of the negative control}} \right]$$

where negative control was run without Brd4(BD1) and compound (inhibitor) and positive control without compound to give the bottom signal and top signal, respectively. Titration of a known inhibitor ((+)-JQ1) was used as control for each experiment. IC₅₀ values were calculated by plotting the log [competitor/compound] versus mean normalized response data from a single experiment measured in duplicates and fitting it to a four-parameter equation with variable Hill slope:

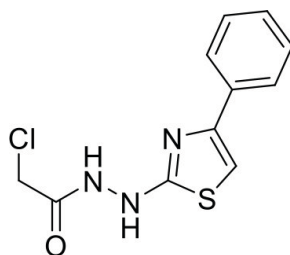
$$Y = \{bottom + \frac{top - bottom}{[1 + 10(\log IC_{50}x)]} \times Hill\ slope\}$$

using the GraphPad Prism software. The plates were measured in a CLARIOStar reader (BMG Labtech) using the homogeneous time-resolved fluorescence module (excitation, 337 nm with 200 flashes; emission, 620 and 665 nm). The 665 nm/620 nm ratios were converted to % normalized TR-FRET ratio = Top ratio signal as 100% and bottom ratio signal as 0% to normalize experiments performed at different days.

3.12 Synthesis of Brd4(BD1) compounds

Reagents, solvents and starting products were acquired from commercial sources. NMR spectra were recorded at the NMR unit (CCiTUB) of the University of Barcelona in CDCl₃ at 400 MHz (¹H and ¹³C). Chemical shifts are reported in δ values downfield from TMS or relative to residual CDCl₃ (7.26 ppm, 77.00 ppm), DMSO-d₆ (2.50 ppm, 39.52 ppm) or CD₃OD (3.31 ppm, 49.00 ppm) as an internal standard. Data are reported in the following manner: chemical shift, multiplicity, coupling constant (J) in hertz (Hz) and integrated intensity. Multiplicities are reported using the following abbreviations: s, singlet; d, doublet; dd, doublet of doublets; dm, doublet of multiplets; t, triplet; m, multiplet; q, quadruplet and bs, broad signal. Evaporation of solvents was accomplished with a rotary evaporator. Thin layer chromatography was done on SiO₂ (silica gel 60 F254). Chromatography refers to flash column chromatography was carried out on SiO₂ (silica gel 60, SDS, 230–400 mesh). IR spectra were performed in a spectrophotometer Spectrum Two FT-IR Spectrometer, and only noteworthy IR absorptions (cm⁻¹) are listed. The accurate mass analyses were carried out using a LC/MSD-TOF spectrophotometer at the Molecular Characterization and Mass Spectrometry service (CCTiUB) of the University of Barcelona. HPLC-MS (Agilent 1260 Infinity II) analysis was conducted on a Poroshell 120 EC-C15 (4.6 mm x50 mm, 2.7 μ m) at 40 °C. Mobile phase (A: H₂O + 0.05% formic acid and B: acetonitrile + 0.05% formic acid) using a gradient elution. Flow rate 0.6 mL/min. The DAD detector was set at 254 nm and the injection volume was 5 μ l and oven temperature fixed at 40 °C.

3.12.1 2-Chloro-N'-(4-phenylthiazol-2-yl)acetohydrazide (Pre-SSR1)



To a cooled solution (0 °C) of 2-hydrazinyl-4-phenylthiazole (1.50 g, 7.85 mmol) in THF (20 mL) was added dropwise a solution of chloroacetyl chloride (0.70 mL, 8.85 mmol) in THF (7 mL). The reaction mixture was stirred at 0°C for 30 min and the precipitate formed was collected by filtration and washed with water to give PreSSR-1 (1.95 g, 93%) as a white solid. ¹H-NMR Spectral data are coincident with the previously reported [61].

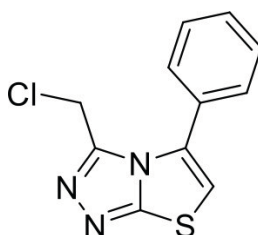
¹H NMR (400 MHz, DMSO-d₆) δ 4.22 (s, 2H, CH₂), 7.28 (s, 1H, SCH), 7.29-7.33 (m, 1H, ArH), 7.38-7.42 (m, 2H ArH), 7.80-7.82 (dd, J = 8.0 Hz, 2H, ArH), 9.17 (bs, 1H, NHCN), 10.92 (bs, 1H, NHCO).

¹³C NMR (400 MHz, DMSO-d₆) δ 41.0 (CH₂), 103.6 (SCH), 125.8 (2CHAr), 128.0 (CHAr), 128.7 (2CHAr), 133.5 (C-ipso), 148.7 (NCCH), 166.2 (CO), 171.8 (NCS).

IR (ATR) 3163, 2969, 2737, 1723, 1624, 1537, 1485, 1397, 1339, 908, 799, 741, 683 cm⁻¹.

HRMS C₁₁H₁₁ClN₃OS [M+H]⁺ 268.0306; found, 268.0307.

3.12.2 3-(Chloromethyl)-5-phenylthiazolo[2,3-c][1,2,4]triazole (SSR1)



A mixture of PreSSR-1 (1.22 g, 4.57 mmol) in xylene (10 mL) and phosphoryl

chloride (7.65 mL, 82.07 mmol) was heated at 110°C for 4 h. Then, the reaction mixture was evaporated and neutralized with saturated NaHCO₃ solution. The aqueous phase was extracted with EtOAc, the combined organic phases were washed with brine, dried over Na₂SO₄, and evaporated to give a residue, which was purified by flash column chromatography (EtOAc) to give SSR-1 as a yellowish solid (824 mg, 72%).

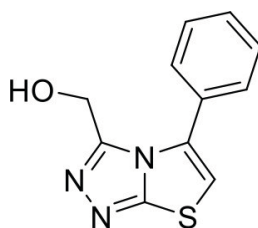
¹H NMR (400 MHz, CDCl₃) δ 4.42 (s, 2H, CH₂), 6.89 (s, 1H, SCH), 7.47-7.53 (m, 3H ArH), 7.73-7.75 (dm, *J* = 8.2 Hz, 2H, ArH).

¹³C NMR (400 MHz, HETCOR, CDCl₃) δ 46.6 (CH₂), 103.8 (SCH), 107.0 (C-ipso), 126.0 (2CHAr), 128.8 (CHAr), 129.3 (2CHAr), 129.5, 147.2, 167.8 (NCN, NCCH, NCS).

IR (ATR) 3164, 3111, 2924, 1724, 1603, 1486, 1194, 1095, 1001, 802, 756, 730, 684 cm⁻¹.

HRMS C₁₁H₉ClN₃S [M+H]⁺ 250.0200; found, 250.0205.

3.12.3 3-(Hydroxymethyl)-5-phenylthiazolo[2,3-c][1,2,4]triazole (SSR-2)



A mixture of SSR-1 (400 mg, 1.60 mmol), sodium acetate (400 mg, 4.88 mmol), sodium iodide (16 mg, 0.11 mmol) and DMF (6 mL) was heated at 100°C for 2 h. The solvent was removed under reduced pressure and the residue was partitioned between EtOAc and water. The organic phase was washed with brine, dried over Na₂SO₄, and evaporated to give the corresponding acetoxy intermediate (240 mg, 55%) as a brown solid. Then,

a mixture of acetoxy intermediate (110 mg, 0.40 mmol), MeOH (15 ml) and a solution of NaOH 3N (1 ml) was heated at reflux for 15 min. The MeOH was evaporated and neutralized with saturated NaHCO₃. The aqueous phase was extracted with EtOAc, the combined organic phases were washed with brine, dried over Na₂SO₄, and evaporated to give a residue, which was purified by column chromatography (1:0 to 97:3 EtOAc:MeOH) to give SSR-2 as a yellow solid (32 mg, 34%).

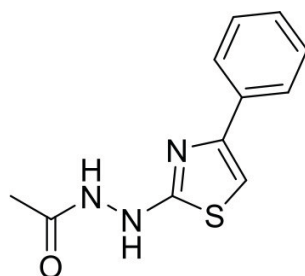
¹H NMR (400 MHz, CDCl₃) δ 4.61 (s, 2H, CH₂), 6.80 (s, 1H, SCH), 7.53-7.57 (m, 3H ArH), 7.58-7.61 (dd, J = 7.4 Hz, 2H, ArH).

¹³C NMR (400 MHz, CDCl₃) δ 54.7 (CH₂), 114.2 (SCH), 128.2 (C-ipso), 129.2 (2CHAr), 129.3 (2CHAr), 130.7 (CHAr), 131.6, 148.4, 158.6 (NCN, NCCH, NCS).

IR (NaCl) 3263, 3105, 2924, 2854, 1704, 1538, 1469, 1442, 1295, 1017, 775, 693 cm⁻¹.

HRMS C₁₁H₁₀N₃OS [M+H]⁺ 232.0539; found, 232.0534.

3.12.4 N'-(4-phenylthiazol-2-yl)acetohydrazide (Pre-SSR3)



To a cooled solution (0 °C) of 2-hydrazinyl-4-phenylthiazole (300 mg, 1.57 mmol) in THF (5 ml) was added dropwise a solution of acetyl chloride (0.13 ml, 1.82 mmol) in THF (3 ml). The reaction mixture was stirred at 0°C for 30 minutes and the precipitate formed was collected by filtration and washed with water to give PreSSR-3 (351 mg, 96%) as a white solid.

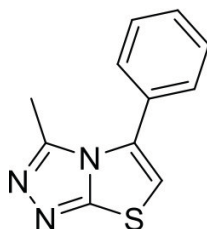
¹H NMR (400 MHz, DMSO-d₆) δ 1.91 (s, 3H, CH₃), 7.23 (s, 1H, SCH), 7.26-7.31 (m, 1H, ArH), 7.37-7.40 (m, 2H, ArH), 7.80-7.83 (dd, *J* = 9.2 Hz, 2H, ArH), 9.42 (bs, 1H, NHCN), 10.16 (bs, 1H, NHCO).

¹³C NMR (400 MHz, DMSO-d₆) δ 20.6 (CH₃), 103.3 (SCH), 125.9 (2CHAr), 128.1 (CHAr), 128.7 (2CHAr), 133.0 (C-*ipso*), 147.8 (NCCH), 169.3 (CO), 172.8 (NCS).

IR (ATR): 3365, 3108, 2949, 2835, 1687, 1616, 1499, 1285, 1171, 992, 903, 744, 689 cm⁻¹.

HRMS: C₁₁H₁₂N₃OS [M+H]⁺ 234.0696; found, 234.0699.

3.12.5 3-methyl-5-phenylthiazolo[2,3-*c*][1,2,4]triazole (SSR-3)



A mixture of PreSSR-3 (115 mg, 0.49 mmol) in xylene (2 ml) and phosphoryl chloride (1.65 ml, 17.7 mmol) was heated at 110°C for 4 h. Then, the reaction mixture was evaporated and neutralized with saturated NaHCO₃ solution. The aqueous phase was extracted with EtOAc, the combined organic phases were washed with brine, dried over Na₂SO₄, and evaporated to give a residue, which was purified by column chromatography (EtOAc) to give SSR-3 as a yellowish solid (51 mg, 48%).

¹H NMR (400 MHz, CDCl₃) δ 2.23 (s, 3H, CH₃), 6.72 (s, 1H, SCH), 7.44-7.46 (dd, *J* = 8.0 Hz, 2H, ArH), 7.49-7.54 (m, 3H ArH).

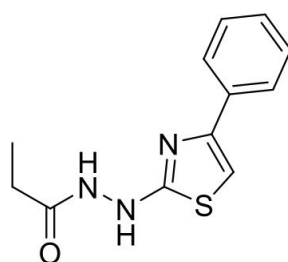
¹³C NMR (400 MHz, CDCl₃) δ 12.6 (CH₃), 113.7 (SCH), 128.2 (C-*ipso*), 128.9 (2CHAr), 129.6 (2CHAr), 130.5 (CHAr), 131.0, 144.9, 157.9 (NCCH),

NCN, NCS).

IR (ATR) 3028, 2928, 1716, 1492, 1473, 1434, 1378, 1031, 762, 717, 700, 666 cm^{-1} .

HRMS $\text{C}_{11}\text{H}_{10}\text{N}_3\text{S}$ $[\text{M}+\text{H}]^+$ 216.0590; found, 216.0588.

3.12.6 N'-(4-phenylthiazol-2-yl)propionohydrazide (Pre-SSR4)



To a cooled solution (0 °C) of 2-hydrazinyl-4-phenylthiazole (500 mg, 2.61 mmol) in THF (8 ml) was added dropwise a solution of propionyl chloride (0.26 ml, 2.97 mmol) in THF (4 ml). The reaction mixture was stirred at 0°C for 30 minutes and the precipitate formed was collected by filtration and washed with water to give PreSSR-4 (551 mg, 85%) as a white solid.

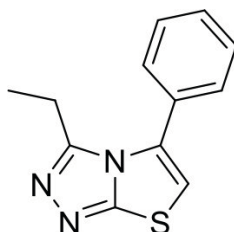
^1H NMR (400 MHz, DMSO- d_6) δ 1.07 (t, J = 7.6 Hz, 3H, CH₃), 2.20 (q, J = 7.6 Hz, 2H, CH₂), 7.24 (s, 1H, SCH), 7.29-7.32 (m, 1H, ArH), 7.38-7.42 (m, 2H, ArH), 7.80-7.82 (dm, J = 7.4 Hz, 2H, ArH), 9.70 (bs, 1H, NHCN), 10.36 (bs, 1H, NHCO).

^{13}C NMR (400 MHz, HETCOR, DMSO- d_6) δ 9.6 (CH₃), 26.6 (CH₂), 103.1 (SCH), 125.7 (2CHAr), 127.8 (CHAr), 128.6 (2CHAr), 133.7 (C-*ipso*), 149.0 (NCCH), 172.8* (CO), 173.0* (NCS).

IR (ATR) 3098, 2978, 2874, 1681, 1635, 1590, 1441, 1333, 1060, 778, 703 cm^{-1} .

HRMS $\text{C}_{12}\text{H}_{14}\text{N}_3\text{OS}$ $[\text{M}+\text{H}]^+$ 248.0852; found, 248.0854.

3.12.7 3-Ethyl-5-phenylthiazolo[2,3-c][1,2,4]triazole (SSR-4)



A mixture of PreSSR-4 (200 mg, 0.81 mmol) in xylene (4 ml) and phosphoryl chloride (2.65 ml, 28.43 mmol) was heated at 110°C for 4 h. Then, the reaction mixture was evaporated and neutralized with saturated NaHCO₃ solution. The aqueous phase was extracted with EtOAc, the combined organic phases were washed with brine, dried over Na₂SO₄, and evaporated to give a residue, which was purified by flash column chromatography (EtOAc) to give SSR-4 as a yellowish solid (83 mg, 45%).

¹H NMR (400 MHz, CDCl₃) δ 1.29 (t, *J* = 7.2 Hz, 3H, CH₃), 2.73 (q, *J* = 7.2 Hz, 2H, CH₂), 6.79 (s, 1H, SCH), 7.46-7.52 (m, 3H ArH), 7.72-7.74 (dm, *J* = 8.2 Hz, 2H, ArH).

¹³C NMR (400 MHz, CDCl₃) δ 32.7 (CH₃), 42.8 (CH₂), 104.0 (SCH), 126.1 (2CHAr), 128.3 (CHAr), 128.9 (2CHAr), 129.0 (C-*ipso*), 129.6, 134.0, 167.3 (NCCH, NCN, NCS).

IR (ATR) 3134, 2979, 1699, 1620, 1588, 1524, 1193, 1016, 951, 828, 757, 720 cm⁻¹.

HRMS C₁₂H₁₂N₃S [M+H]⁺ 230.0746; found, 230.0749

Chapter 4

Results

4.1 Identify and confirm the first hits binding to Fbw7-Skp1 complex

4.1.1 Background: The Ubiquitin Proteasome System (UPS)

The ubiquitin-proteasome system (UPS) is the major proteolytic system in eukaryotic cells through which approximately 80% of proteins are degraded [56]. Protein ubiquitination occurs through a cascade of enzymatic reactions, involving an E1 ubiquitin-activating enzyme, an E2 ubiquitin-conjugating enzyme and, finally, an E3 ubiquitin-ligase (of which >700 are known in humans [62]). Repeated iterations of this ubiquitination process, results in long chains of ubiquitin repeats on a given substrate and posterior recognition and degradation by the proteasome [63, 64] (Fig. 4.1). The therapeutic potential of this system has already been proved by the release in the market of two proteasome inhibitors, bortezomib and carfilzomib, for the treatment of multiple myeloma [65]. They act downstream the UPS cascade blocking all degradation process, therefore upregulating non-desired substrates that lead to a wide range of side effects. E3 ligases, instead, represent appealing targets in drug discovery; indeed, they dictate substrate specificity by holding the substrate in close proximity of the E2 ubiquitin-conjugating enzyme and thus offer the possibility for specific therapeutic intervention with less toxic events [66]. Despite the relevant therapeutic potential they offer, the identification of small-molecules against E3 ligases to develop novel therapies has been achieved with very limited success. Only three molecules (thalidomide, lenalidomide and pomalidomide) targeting the E3 ligase cereblon have been approved by the Food and Drug Administration (FDA) for the treatment of hematologic malignancies such as multiple myeloma (MM) and del(5q) myelodysplastic syndrome (MDS) [67]. The main stumbling block in finding ligands for E3 ligases is the need of

modulating Protein-Protein interactions (PPI). The contact surfaces involved in protein-protein interactions are large (1,500-3,000 Å²) compared with those involved in protein small-molecule interactions (300-1,000 Å²). In addition, the contact surfaces of proteins that interact with other proteins are generally flat and often lack the grooves and pockets present at the surfaces of proteins that bind to small molecules [68]. Furthermore, competing with post-translation modifications can result quite challenging, i.e. disrupting phosphorylation binding is arduous since a small-molecule should be hydrophobic enough to pass through the cell-membrane and polar enough to compete with it (drug-like properties not achievable, leading to the idea that E3 ligases are undruggable targets). Nevertheless, using different strategies, the scientific community has been able to find small-molecules to target several of these E3 ligases. For example following a peptidomimetic strategy, the team lead by Dr Alessio Ciulli was able to identify nanomolar ligands of the VHL E3 ligase [69]. Performing a virtual screening ligands of the Skp2 E3 ligase were identified [70]. FBDD has been already used as a strategy to target E3 ligases, scientists at GSK and Astex were able to identify fragments able to bind to the E3 ligase Keap1 with moderate potency [71]. Another approach that is yielding a number of small-molecules binding to the E3 ligases is the deep study of the mechanism of action or the structure of already identified drugs. This was the case of the discovery that thalidomide and other immunomodulatory drugs (IMiDS) bind to cereblon, [72] promoting the degradation of important non-natives substrates. Finally, using covalent ligands discovered with proteomic approaches, very recently,

E3 ligases such as RNF114 and RNF16 have been targeted. [73]

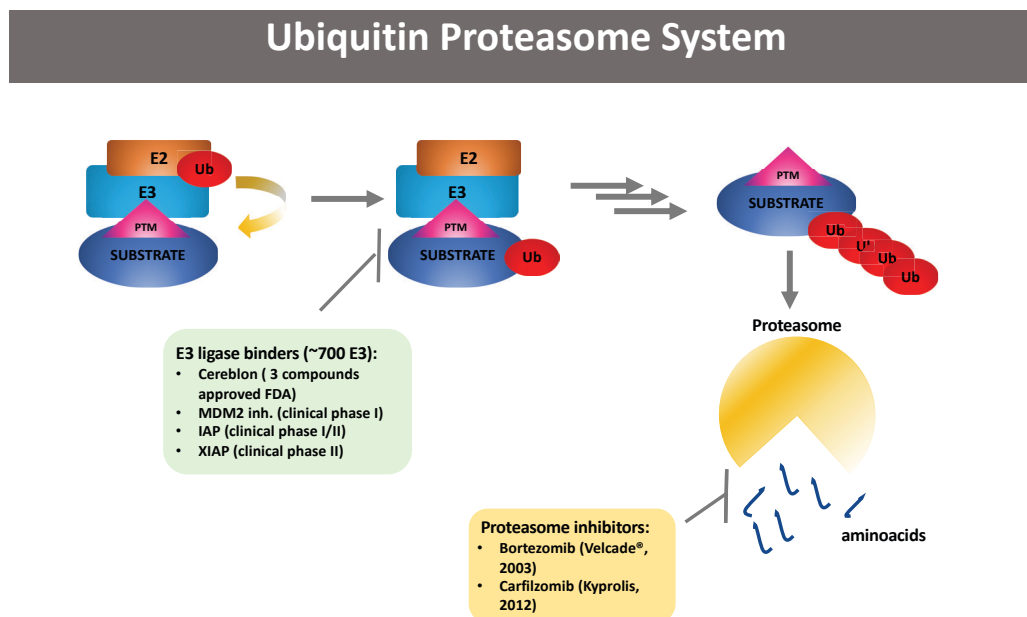


Figure 4.1. UPS pathway regulates proteins turnover. A substrate carrying a post-translation modification (PTM) (e.g. phosphorylation) binds to the recognition site (degron) of a specific E3 ligase. Once this happens the E2-ubiquitin conjugating enzyme transfer the ubiquitin (Ub) to a lysine of the substrate. After several iterations of ubiquitination, the substrate is recognized by the proteasome and destroyed.

4.1.2 Hijacking the UPS: PROTAC technology

In parallel, in recent years, controlling protein level with small molecules to regulate protein function has attracted much interest as a new therapeutic modality: PROteolysis TArgeting Chimera (PROTAC) technology. PROTAC molecules are heterobifunctional small molecules designed to induce intracellular protein degradation. One tail of the PROTAC molecule binds to a Protein Of Interest (POI), while the other tail recruits a specific E3 ligase forming a ternary complex that allows ubiquitin transfer from the E3 ligase to the POI leading to degradation [74, 75] (Fig. 4.2). PROTAC molecules, stand far from the classical pharmacology in which a compound binds to the POI and modulates their activity. In the case of classical inhibitors high target occupancy is required to affect target function (the so-called occupancy-driven pharmacology) and the effect of the drug will last as long as the

molecule occupies the protein. Instead, for PROTAC, sub-stoichiometric drug binding is sufficient and binding anywhere in the protein surface of the POI can potentially be exploited, as far as the formation of the complex is effective and leads to degradation. Furthermore, once the protein is degraded PROTAC molecules can bind another POI, in a catalytic way. In this case, biological regeneration (turnover) of a new POI is necessary to restore POI function (event-driven pharmacology) [76]. The potential to redirect protein degradation by artificially recruiting an E3 ligase has been demonstrated and several studies have reported the successful development of drug-like PROTAC molecules able to degrade therapeutically relevant proteins. The first proof-of-concept of the PROTAC approach was described in 2001 and 2004 as a result of a research collaboration between the groups of Drs Craig Crews and Raymond Deshaies [78, 79]. This first generation of PROTAC molecules were mainly peptides either for the E3 ligase or for the POI, thereby the PROTAC approach remained largely dormant for over a decade. However, the development of drug-like E3 ligases really flurried the field. Employing VHL and cereblon as E3 ligases, in May and June 2015, three teams, led by Ciulli, Bradner (Dana Farber) and Crews, published several independent studies describing active PROTAC molecules with drug-like properties [79–83]. Today, dozens of PROTAC molecules have been developed targeting a plethora of targets, mainly for the treatment of different cancers in both industry and academia. PROTAC molecules targeting B-cell lymphoma 6 (BCL6) from AstraZeneca [84], P300/CBP-associated factor and general control nonderepressible 5 (PCAF/GCN5) from GSK [85], Bruton's tyrosine kinase (BTK) from Pfizer [86], focal adhesion kinase (FAK) from Boehringer Ingelheim [87] and Interleukin-1 receptor-associated kinase 4 (IRAK4) from GSK [88]. In 2019 Arvinas, a biotechnology company founded around the PROTAC concept, launched the

first clinical-trial of two candidate-drugs based on targeted protein degradation: ARV-110 and ARV-471 that target the Androgen and Oestrogen Receptor, respectively [89]. The first-in-class PROTAC ARV-110 has clinical activity in pre-treated mCRPC patients who have progressed on anti-androgen therapy and despite this, ARV-110 demonstrated the first evidence of antitumor activity in this difficult-to-treat patient population, as communicated by Arvinas this year. Studies are on-going and scientific community is excited to see the outcomes of these clinical trials as well as other discoveries to expand the number of therapeutics to treat huge unmet medical need.

In parallel to the development of drug-like PROTAC molecules, pushed by the observation that IMiDs promote the degradation of neo-substrates, several novel molecular glues have been recently described. Molecular glues have to promote de novo PPIs, somehow restricting their utility for the development of drugs for part of the proteome. However, molecular glues are formed by a unique molecular scaffold and have more drug-like properties. Until now, the discovery and development of novel molecular glues have been largely serendipitous. Recently, several structural and functional strategies have been described to shed light in the discovery of this type of drugs. [90]. In the next years, surely, a considerable number of drugs based on target protein degradation (PROTAC molecules and molecular glues) will be developed for until now considered undruggable proteins, converting a big part of the undruggable proteome into a druggable one, and delivering completely novel therapies to cure challenging diseases, such as cancer or Alzheimer's disease [91].

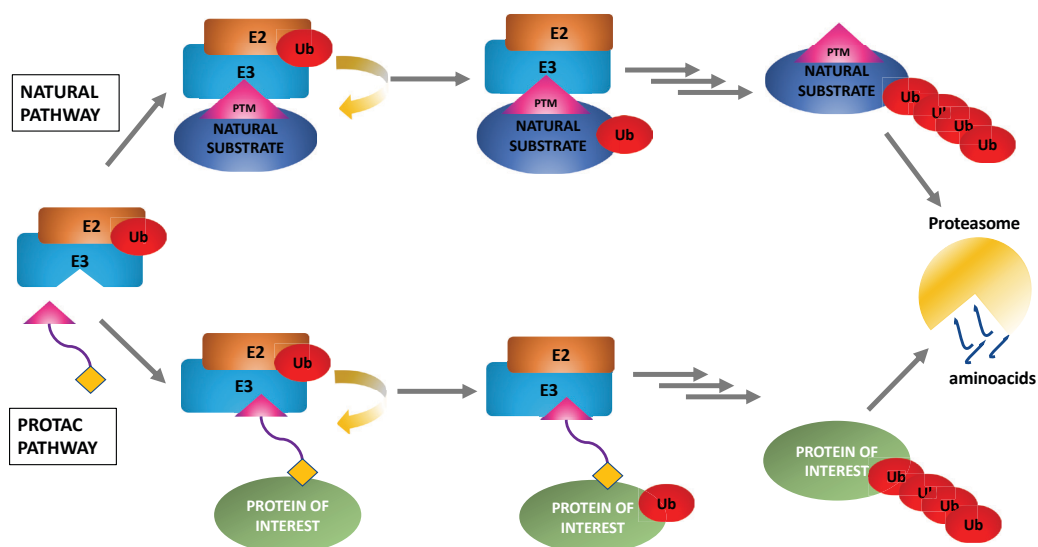


Figure 4.2. PROTAC molecule hijacks the ubiquitin proteasome system. On the top highlighted natural pathway of the UPS, on the bottom a PROTAC molecule binds in one side to the E3 ligase and on the other side to the POI that will be ubiquitinated and destroyed.

4.1.3 Fbw7: a challenging attractive target in drug discovery

E3 ubiquitin ligases can be divided into different families and subfamilies depending how they recognize and degrade the corresponding substrates. In mammals, the best characterized subfamily is the SCF family (S-phase kinase-associated protein 1 (Skp1)–Cull1–F-box protein). In this, the N-terminus of the Cull1 interacts with Skp1, which binds to the F-box domain of an interchangeable F-box protein [92] The F-box proteins typically recognize unique, short degradation motifs in their substrates (termed degrons).

Although many F-box protein–substrate pairings have been described and linked to potential biological functions, the first hurdle to generating SCF-directed pharmacology therapies remains the therapeutically validation and manipulation of each pair (for example, through the development of chemical probes). Nowadays, our expanding knowledge of the SCF family of ubiquitin ligases indicates that targeting a specific SCF complex may

result in effective therapies for various disorders such as cancer, sleep disorders, mood disorders, inflammation and acquired infections.

One of the most commonly deregulated proteins in human cancers is the SCF E3 ligase Fbw7, which targets a range of substrates for degradation, including key human oncoproteins such as cyclin-E, c-Myc, Notch and Junk. Indeed, on the one hand, recent progresses have illuminated Fbw7's central role in tumorigenesis: large-scale cancer genome studies have shown that Fbw7 gene is among the most mutated in cancer (6% of all cancers have mutations in this gene) and its tumour suppressor function has been demonstrated [93]. On the other hand, Fbw7 has showed an important role in sensitizing cancer stem cells to chemotherapies [94]. Given this relevance in cancer, several therapeutic strategies to pharmacologically manipulate Fbw7 have been envisaged; however, so far, no small molecules directly targeting Fbw7 have been reported. Another important function attributed to Fbw7 is its role in neuropsychiatric disease. Indeed, in 2018 it was demonstrated that Fbw7 dictates the degradation of Disrupted in Schizophrenia 1 (DISC1), a multifunctional scaffolding protein that assemble protein complexes to promote neural development and signalling. Stabilizing DISC1 by mean of Fbw7-DISC1 disruption could be beneficial to treat schizophrenia [95].

4.1.3.1 Fbw7-Skp1 structure

The first crystal structure of Fbw7 in complex with its adapter protein Skp1 was solved in 2007 by Dr. Bing Hao and collaborators (PDB code 2OVP) [96]. The Fbw7 structure consists of a F-box domain (helices H-1, H0, H1-H3), an α -helical linker domain (helices H4 and H5, and H4-H5 loop), a WD40-repeat domain and a five-residue tail (Fig. 4.3A). The WD40 domain forms a canonical eight-bladed β -propeller structure, with each blade

consisting of four antiparallel β strands (strands A through D) (Fig. 4.3B). As in other WD40-domain structures, the Fbw7 β -propeller resembles a cylinder with narrow and wide ends and a solvent-filled central channel (Fig. 4.3B). In order to exert its function, Fbw7 has to interact with the other partners of the SCF RING-E3 family (Fig 4.3C), interacting in one site with Skp1 (with the F-box domain) and in the other site with the substrate (WD40 domain).

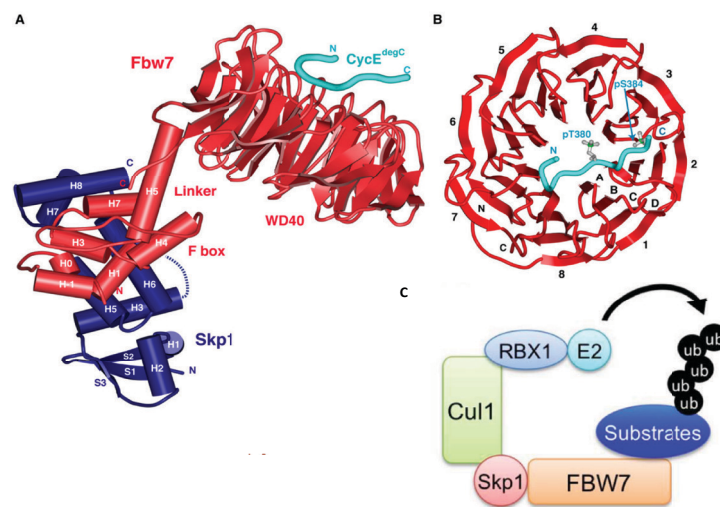


Figure 4.3. Overall structure of the ternary Fbw7-Skp1-C-terminal Cyclin E peptide complex. **A)** Ribbons representation shows Skp1 (blue), Fbw7 F-box, Linker and WD40 domain (red) and phosphorylated C terminal Cyclin E, CycE31pS372/pT380/pS384 (cyan). Dotted lines represent disordered regions. **B)** WD40 domain with the eight-bladed β -propeller structure forms a cylinder with a central channel. **C)** Substrate ubiquitination by RBX1-SCF E3 ubiquitin ligase: Cullin-1 at its N-terminus binds to Skp1 and an F-box protein, which recognizes protein substrates, and at its C-terminus binds to RBX1. RBX1, on the other hand, binds to Cullin-1 using its N-terminus and an E2 ubiquitin conjugating enzyme using its C-terminal RING domain. Together, RBX1-cullin-1 catalyses the ubiquitin transfer from E2 to protein substrates. Adapted from [96].

4.1.3.2 Fbw7-Skp1 substrate recognition and mutations

The WD40 domain contains the so-called degron site, given the ability to recognize the Fbw7 substrates. Most of these substrates harbour a conserved phosphorylation motif (CPD) [97], first identified in the yeast homologue

Cdc4 of Fbw7. This motif contains conserved residues (L)-X-pT/pS-P-P-X-pS/pT/E/D (X represents any amino acid) [98] found to be essential for E3 ligase recognition which typically includes threonine or serine (Fig. 4.4) that once phosphorylated engage three key arginine residues of the WD40 domain [99]. Crucially, according to COSMIC database, the most frequent mutations of Fbw7 occur in these three residues, Arg465, Arg479 and Arg505 (Fig. 4.5).

Fbw7 CPD consensus motif

Notch1	TPTLSPPL
CyclinE	TPPQSGK
C-myc	TPPLSPS
C-jun	TPPLSPI
SREBP1	TPPPSDA
SV40LTA	TPPPEPE
DISC1	TPPGSHS
Consensus	TPPXS

Figure 4.4. Conserved CPD consensus motif found in all known Fbw7 substrates [95].

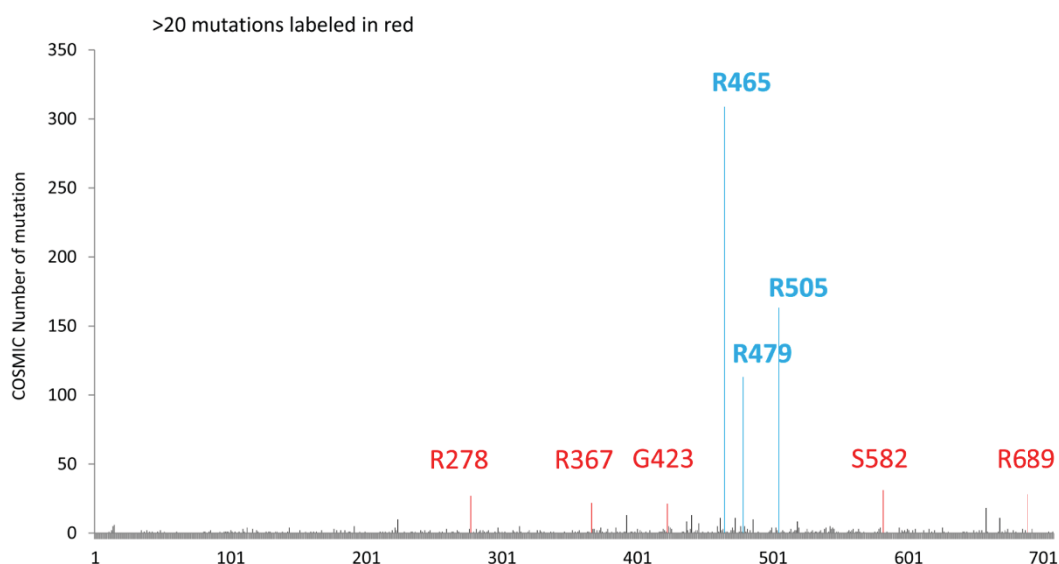


Figure 4.5. The distribution of Fbw7 mutations in the FBXW7 genome. The numbers of FBXW7 genetic alternations are retrieved from the COSMIC database. Fbw7 residues with a mutation number higher than 20 are highlighted in red and 3 Fbw7 hotspot mutations are labelled in blue. Adapted from [100].

Results Fbw7-Skp1

These mutations abolish protein-substrate interaction leading to dysregulation of substrate activity in cells and are frequently detected in a wide range of tumours (Fig. 4.6). As already mention above, Fbw7 controls several oncoproteins and a perfect balance of protein turnover regulated by UPS and hence by Fbw7 is crucial to maintain normal cellular homeostasis.

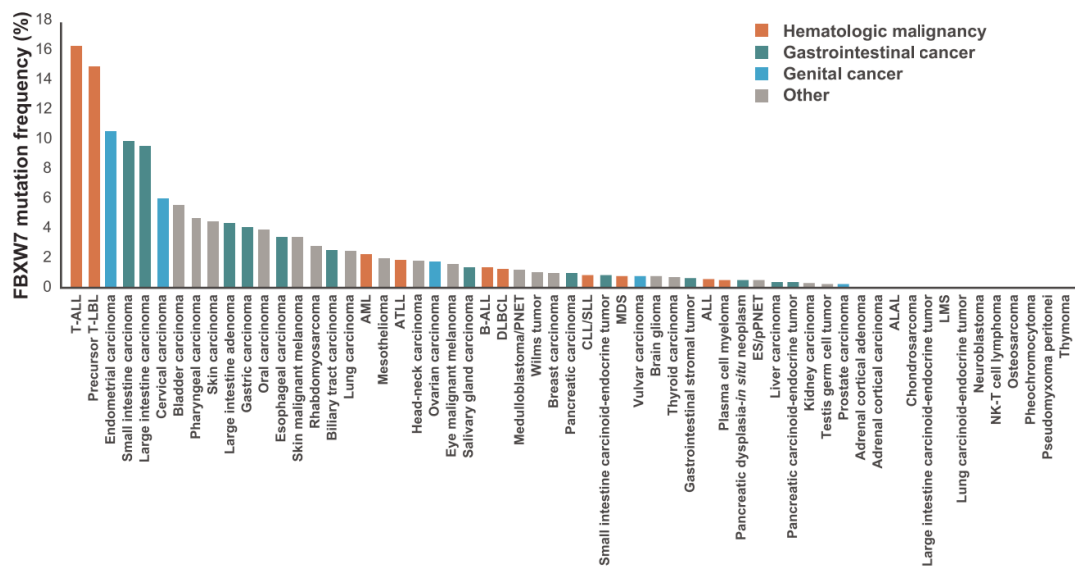


Figure 4.6. *FBXW7* gene mutation frequency for different human cancer types in the COSMIC database. Cancer types with >100 patients registered are listed. Adapted from [99].

4.2 Objective of the Chapter 4

In this chapter we aim to identify and characterize fragment molecules able to bind to the Fbw7-Skp1 multi-subunit complex allosterically and to develop them into chemical probes and as E3 ligases ligands for PROTAC development. Indeed, the Labs of Dr Galdeano and Dr Barril aim to expand the so-called druggable genome by targeting so-far unexploited sites that elicit a biological response through non-standard mechanisms of action. To achieve this goal, the Labs employ a multi-disciplinary and question-driven approach that combines computational, biophysical, chemistry and cell biology. Given its potential, we envisaged that FBDD could be an optimal drug discovery approach to find small-molecules of the Fbw7 E3 ligase.

4.3 Cloning, expression, purification and characterization of Fbw7-Skp1 complex

4.3.1 Cloning and expression of Fbw7-Skp1 complex in *E. coli*

The cloned glutathione S-transferase (GST)-tagged human Fbw7 (residue 263-707) and truncated Skp1 was co-expressed as a dicistronic message in BL21(DE3). It was necessary to co-express Fbw7 with Skp1, because Fbw7 alone was itself insoluble. The co-expression of the recombinant proteins in *E. coli* strain BL21(DE3) was started after induction with 1mM IPTG. (section 3.7.1). The first goal to express the protein was achieved as illustrated in Figure 4.7, although the overall yield of expression was very low (0,3 mg/L). A lot of trials were done to improve the yield of the expression:

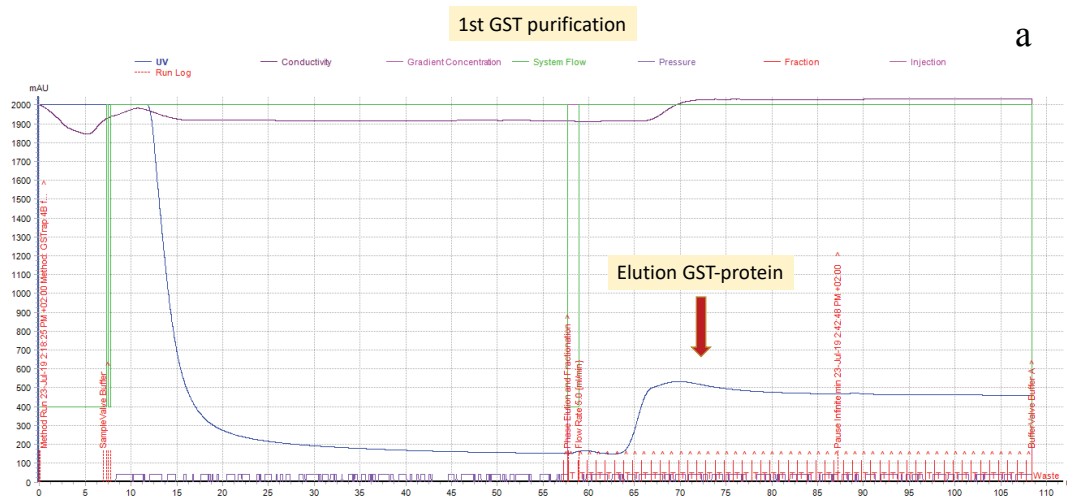
- Different bacterial strains such as *E. coli* BL21, BL21(DE3), BL21(DE3) plys, Rosetta, Rosetta(DE3)
- Different growth culture such as Terrific Broth or Luria Broth
- IPTG induction at 10 h, 18 h or 24 h
- Harvesting time of bacteria cells at OD 0.6 and 2.4

None of these conditions helped us to obtain a higher amount of protein. Furthermore, the removal of a GST-tag in change of a His-tag in the protein-construct, made by our collaborator Dr Bing Hao, did not help in achieving a better result. Therefore, one of the major bottlenecks of this project has been the low yield of the expression. While SPR has the advantage of a very low amount of protein needed (in the range of μg), DSF and NMR require a considerable amount of protein (in the range of mg). Huge efforts have been dedicated to express the sufficient amount of Fbw7, around 40 L of *E. Coli*, to perform all the biophysical experiments described in the thesis. Overall

this result reflects the difficulties on pursuing a fragment-based screening in the targets were the protein is produced in a low amount.

4.3 Purification and characterization of Fbw7-Skp1 complex

Fbw7-Skp1 complex was purified with the following protocol: cells were collected and subjected to lysis and the filtered solution was purified by a glutathione affinity chromatography. The GST-tag was then removed by an overnight thrombin cleavage. The cleaved complex was then purified again with glutathione affinity chromatography to separate the GST-tag and protein with glutathione affinity still present in the solution. Final purification was performed with an anion exchange chromatography using a heparin column. Chromatograms of the purification steps are illustrated in Figure 4.7.



Results Fbw7-Skp1

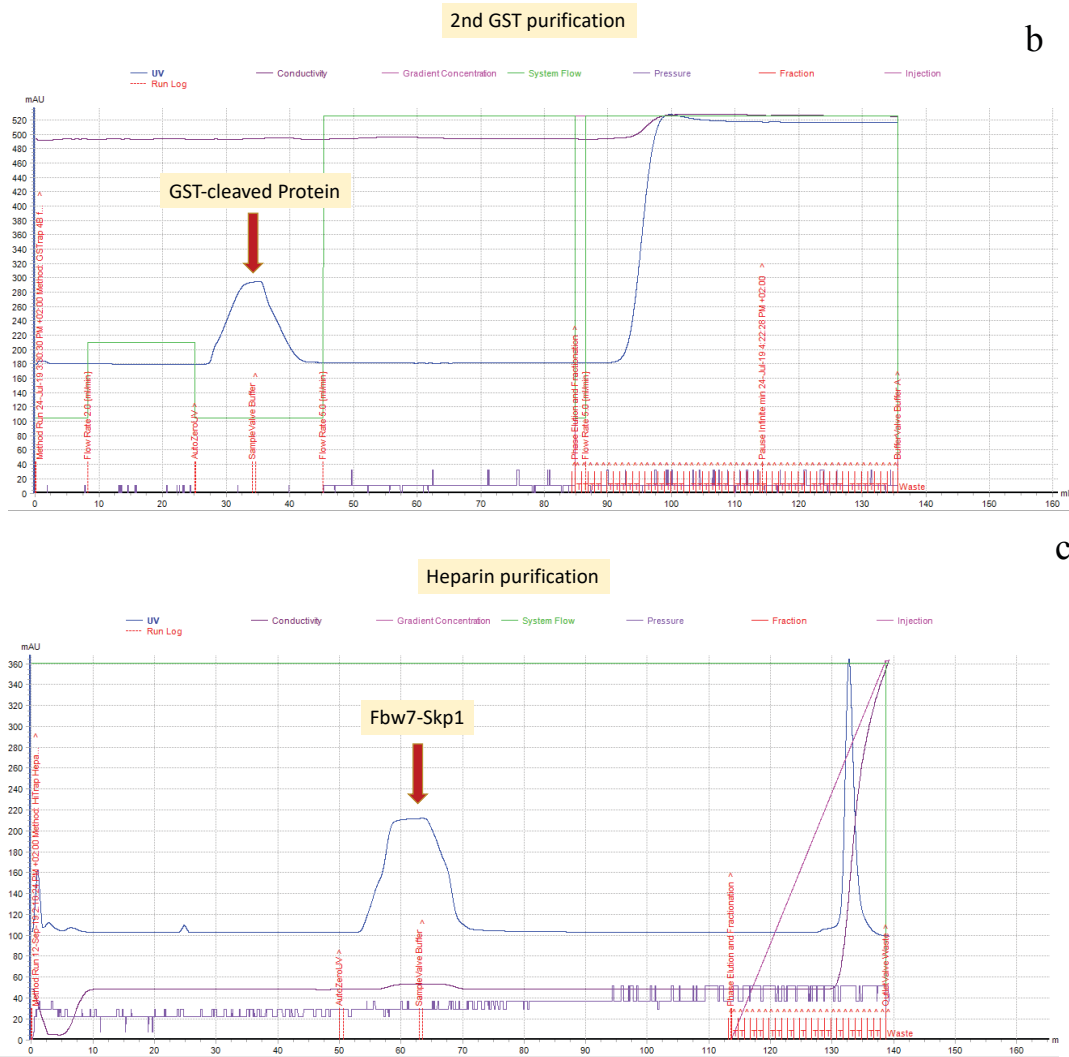
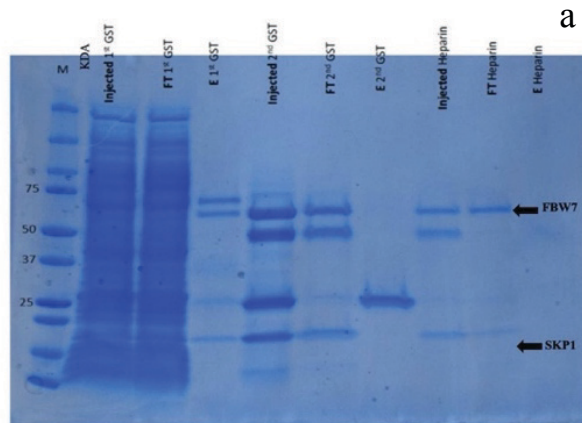


Figure 4.7. Chromatograms showing a) 1st GST purification b) 2nd GST purification c) Heparin purification.

The purification yield to 0.3 mg/l. Purification step were monitored by gel electrophoresis and protein purity was assessed by mass spectrometry. (Fig. 4.8)



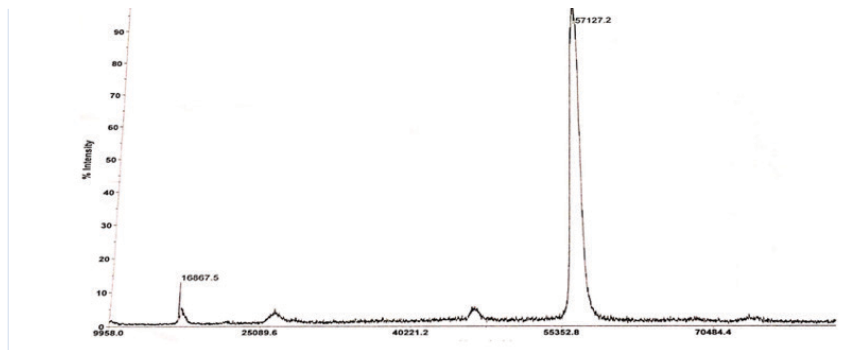


Figure 4.8. a) Electrophoresis gel showing the different steps of purification for Fbw7-Skp1, stained with Coomassie blue. b) Matrix-assisted Laser Desorption/Ionization (MALDI) shows Fbw7 (56 KDa) and Skp1 (16 KDa).

4.4 Fragment Screening

4.4.1 DSF

DSF experiment was selected as a frontline biophysical technique to screen the in-house fragment library. Since DSF conditions using Fbw7-Skp1 complex had never been described before, buffer screening was carried out to find the optimal conditions, by testing the stability of the protein using different buffers, salt concentrations and a range of pHs. Finally, 50 mM Tris pH= 8.0 and 50 mM NaCl buffer was considered the best buffer to perform the experiment. Therefore, a primary screening of five fragments at a single concentration (1 mM 5% DMSO) was performed in order to assess assay feasibility (Fig. 4.9 c, d, e, f, g). A 26-mer peptide (KAMLSEQNRASPLPSGLL[pT]PPQ[pS]GKK) of the natural substrate cyclin-E (K_d 70 nM [96]) was used as positive control (Fig. 4.9a and 4.9b). Crucially, the increasing of fluorescence was observed in two clear peaks in some cases, hence having two melting temperatures or just one peak in the other occasions, as illustrated in figure 4.9a and 4.9b, respectively. Notwithstanding, we were not able to determine which protein or domain corresponded to each melting temperature. The screening of the selected five fragments gave in all cases a

Results Fbw7-Skp1

negative shift. Therefore, since the reproducibility of the experiment was not consistent and given the preliminary and the inconclusive results obtained we decided to move forward using SPR to screen all the library.

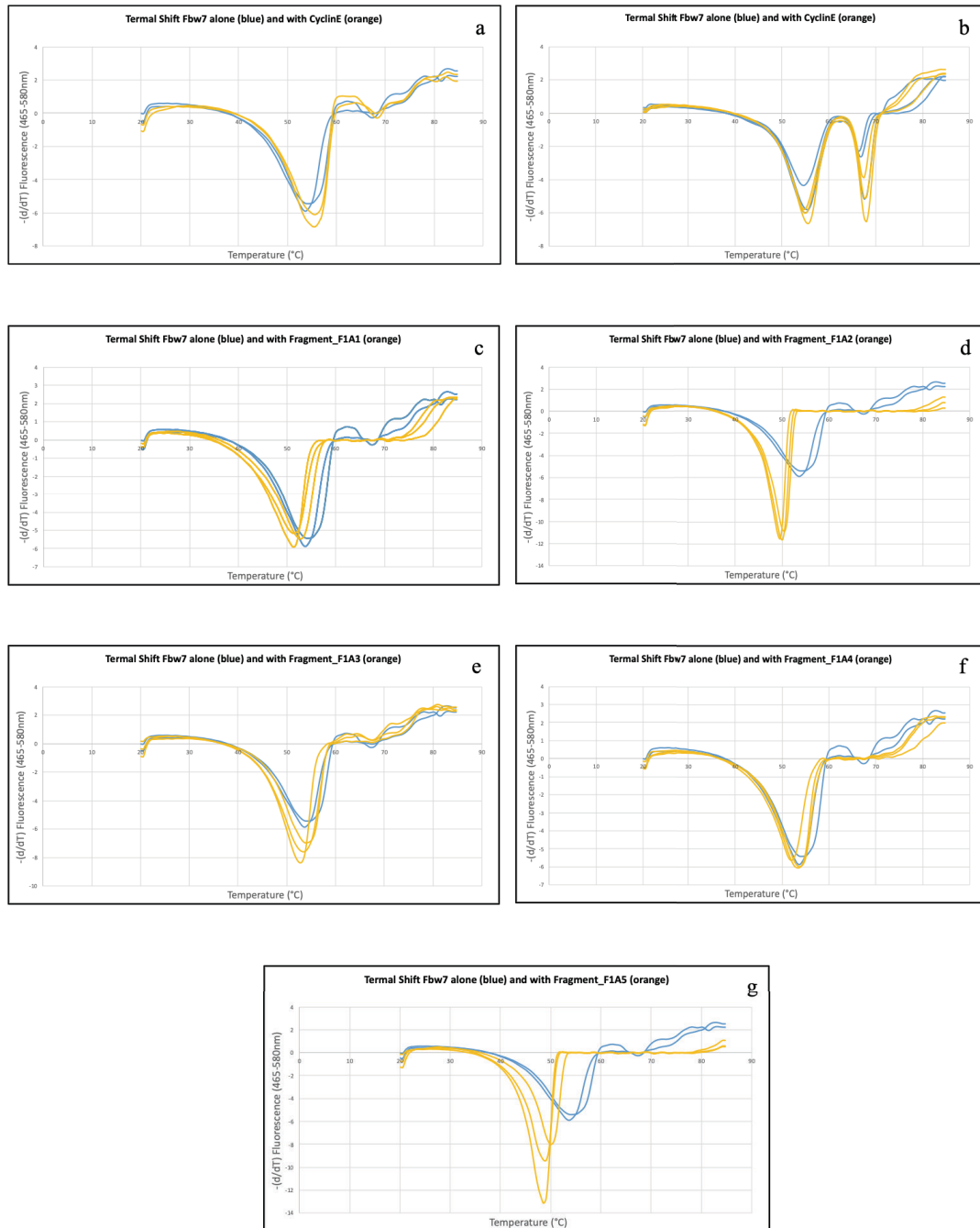
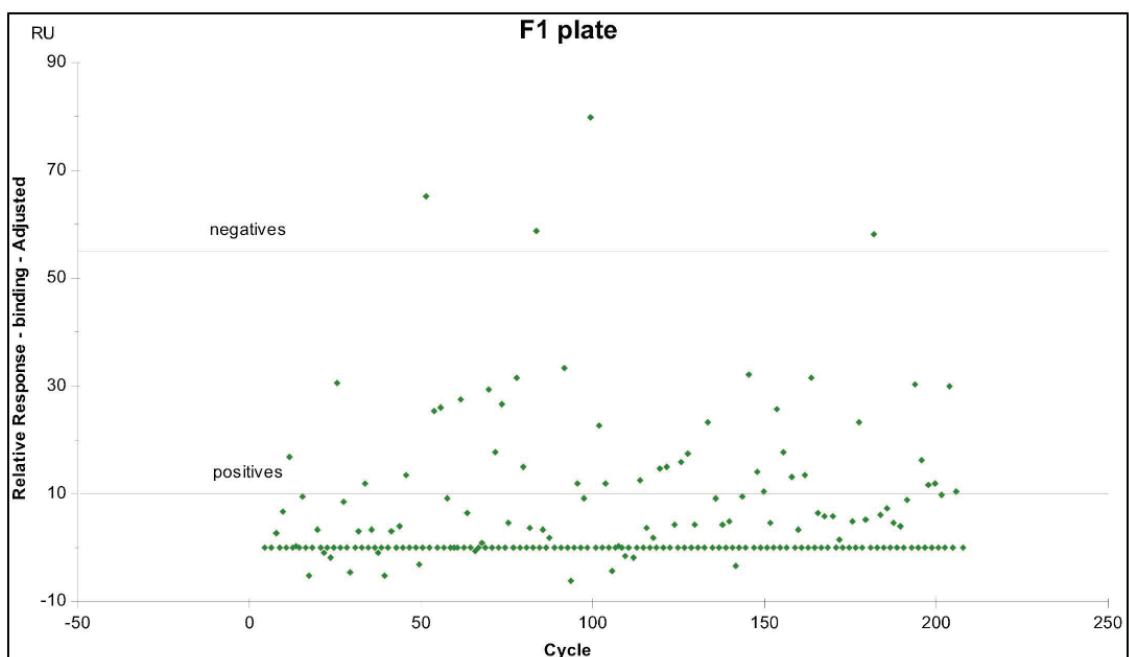


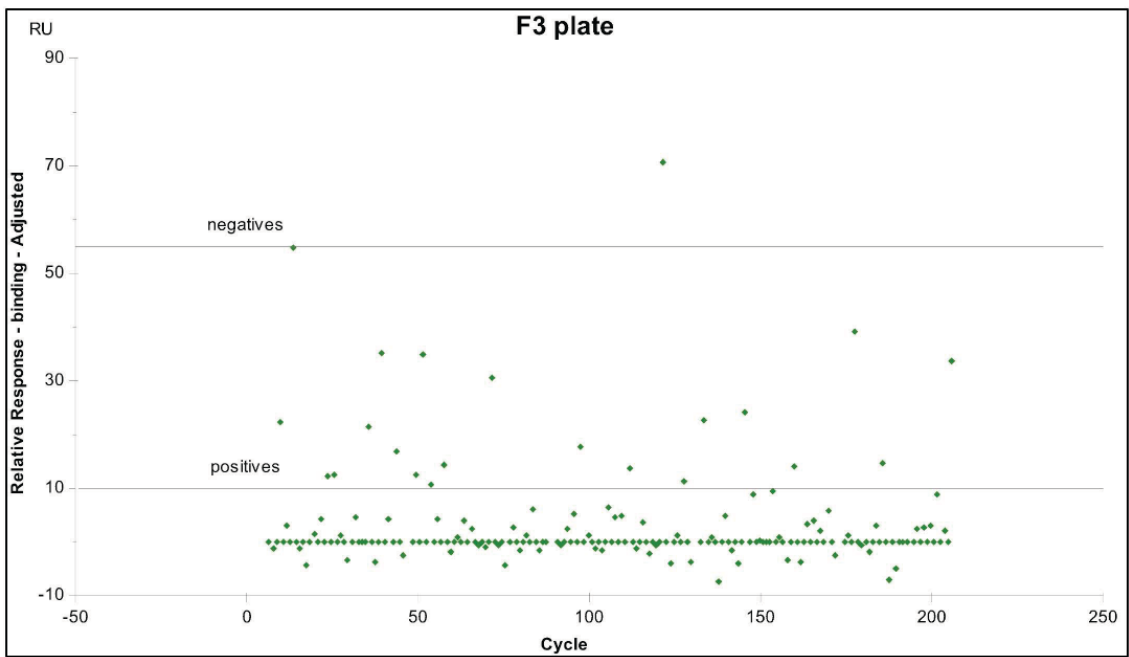
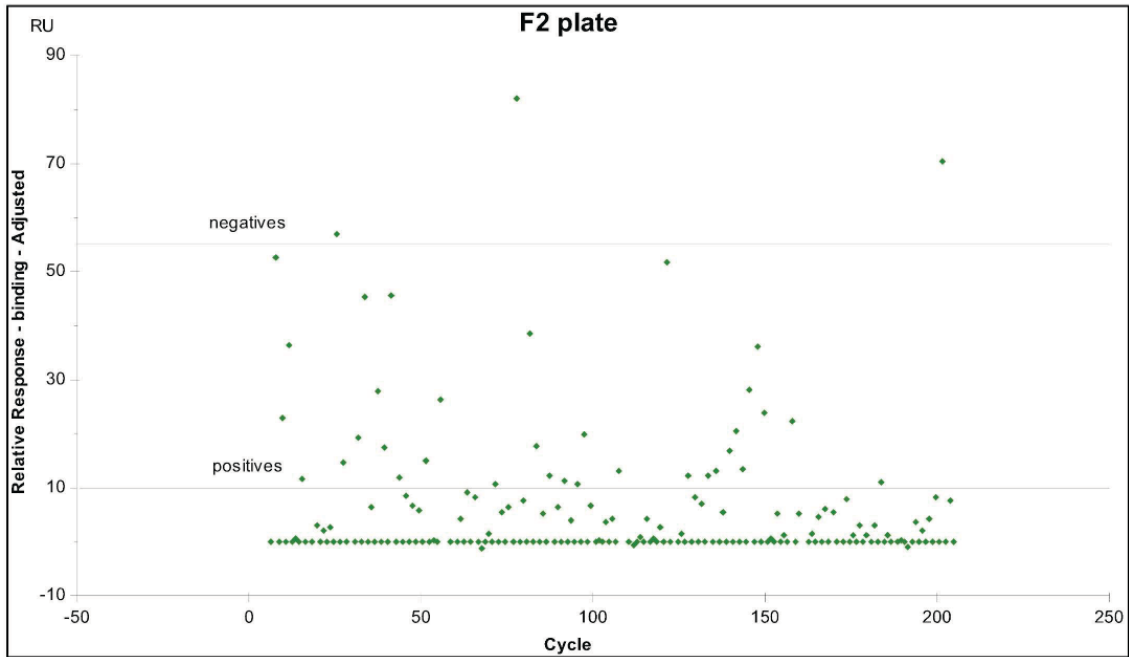
Figure 4.9. DSF Plots showing a) and b) the melting temperature of Fbw7-Skp1 in complex with a peptide of the natural substrate Cyclin E, c) with F1A1, d) with F1A2, e) with F1A3 f) with F1A4, g) with F1A5.

4.4.2 SPR experiments

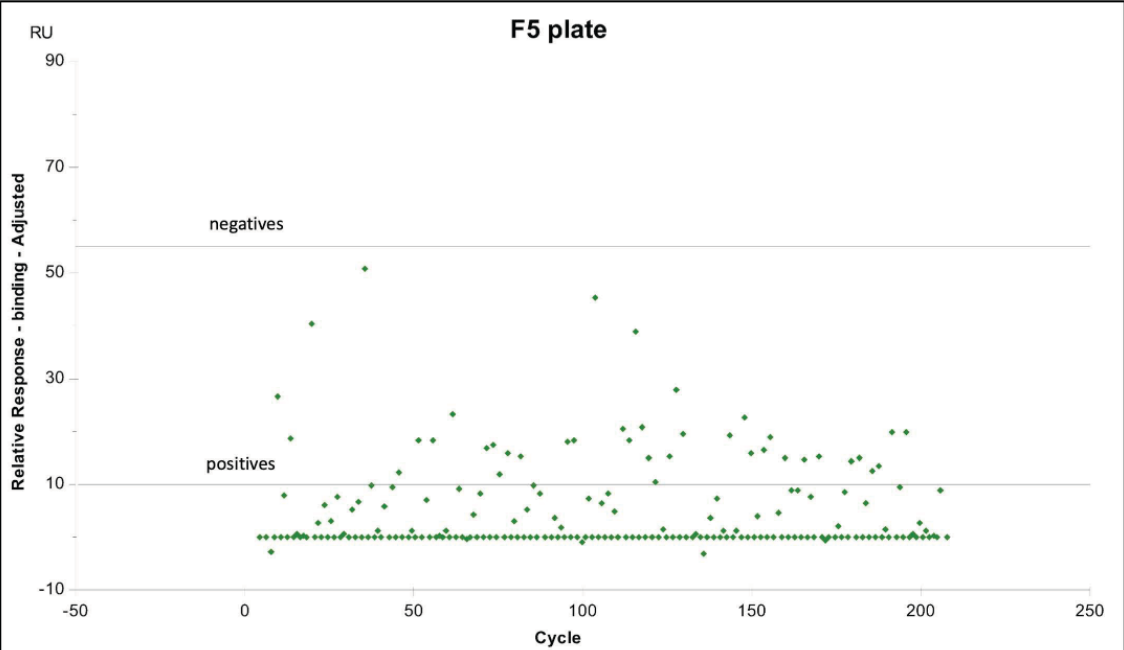
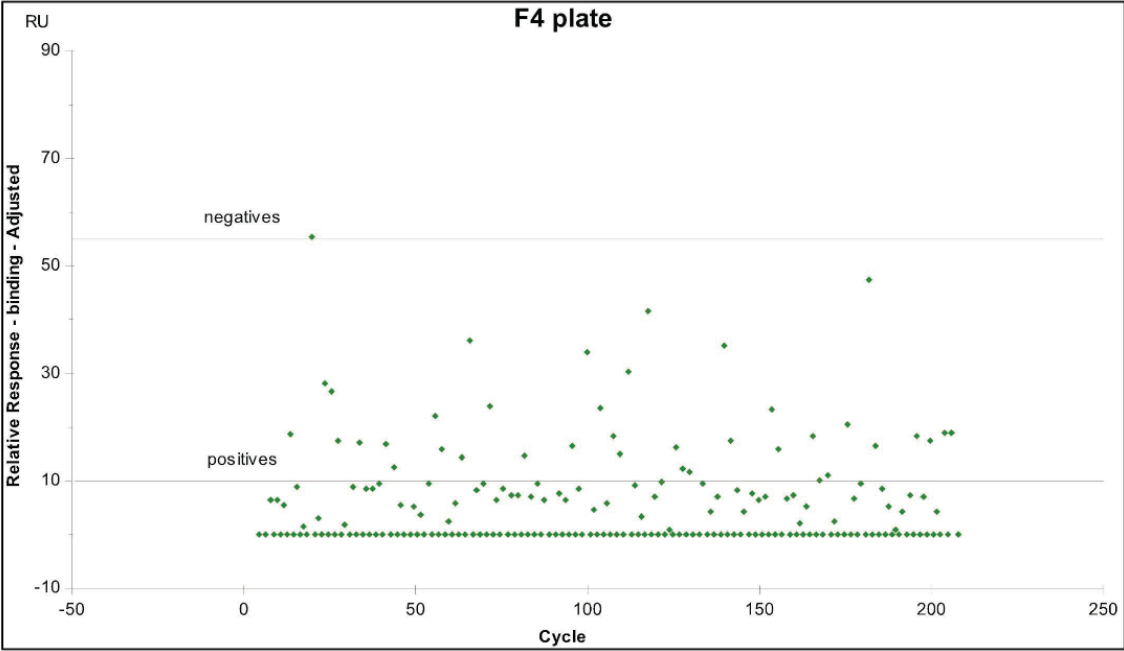
Fbw7-Skp1 complex was immobilized on the sensor chip's surface using the amine-coupling protocol as described in section 3.11.2. One channel was used for the immobilization of the protein with approximately 16000 RU. A blind screening was performed since the peptide of the natural substrate cyclin E did not give any response, perhaps due to the random immobilization of the protein in the surface of the chip, limiting the accessibility of the degron binding site for the peptide or due to solubility problems. Figure 4.10 shows the response of the fragment library divided in seven plates of 96 wells. According to the amount of protein immobilized in the chip and the molecular weight average of the library, 55 RU was determined as maximum response and the lower cut-off was set at 10 RU. The two horizontal lines show up the upper and lower cut-off levels. The green spots represent individual fragments at a single concentration of 500 μ M with 5% DMSO. The spots with a response upper than 55 exceed 1:1 stoichiometry, suggesting they either bind to multiple sites or aggregate non-specifically on the surface of the protein.



Results Fbw7-Skp1



Results Fbw7-Skp1



Results Fbw7-Skp1

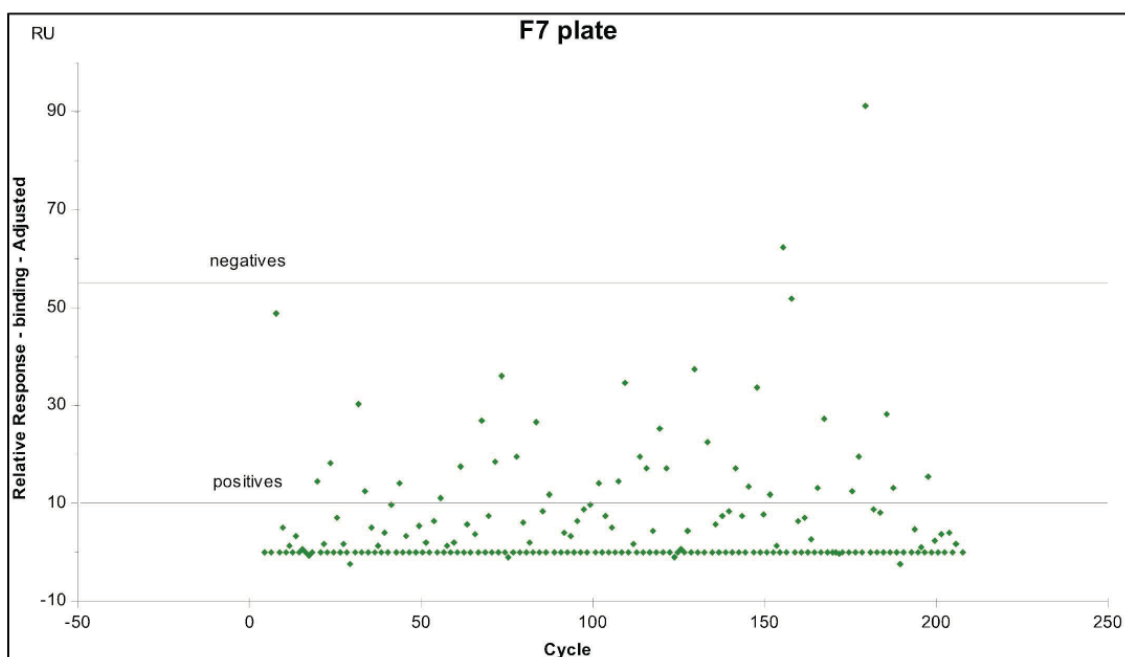
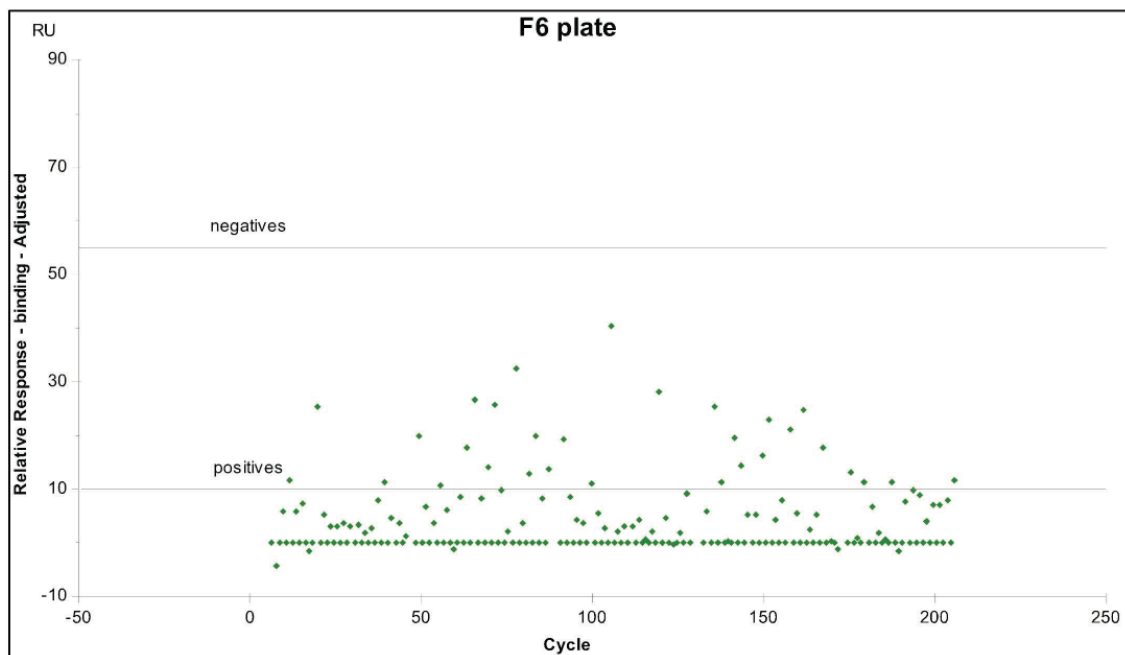


Figure 4.10. Each plot represents the screening of a 96-wells plate. Each fragment at a single concentration of $500 \mu\text{M}$ (green dot) corresponds to a different cycle (x axis) that gives a specific response unit (RU, y axis). Fragments giving a RU between 10 and 55 were considered positive hits.

4.4.2.1 Dose response curve assessed by SPR: Fbw7 linked by amine coupling reaction

220 fragments out of 672 fragments gave a positive response between 10 and 55 RU. In order to validate these fragments and to extrapolate the K_d , a titration curve was performed as specified in section 3.11.2.2. Of these 220 fragments 48 gave a $K_d < 2$ mM (table 4.1). The 21 most potent fragments with a $K_d \leq 200$ μ M are highlighted in table 4.1 and the dose-response curve illustrated in the following Figure 4.11. Example of the sensograms of the most potent fragments is illustrated in Fig. 4.12. Surprisingly, several of the fragments screened (F1A10, F1C8, F1C12, F1D5, F1G4, F1H7, F2A1, F3A8, F4E10, F4G7, F5G12, F6D5, F6F12, F6H3, F7E2, F7G9) showed a K_d in the two digits and one-digit micromolar range, which is very potent for a fragment-sized ligand. Moreover, it is important to point out that we found fragments with similar chemical scaffolds between the most potent ones. The overall hit rate of the screening was around 3%, an encouraging rate for a target that until now was considered undruggable.

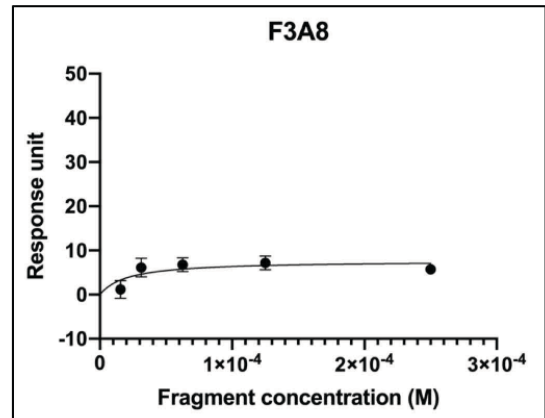
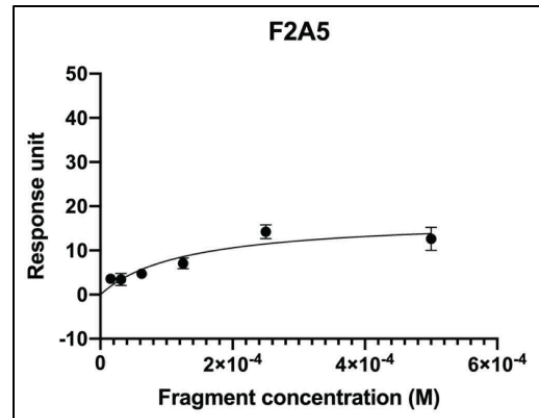
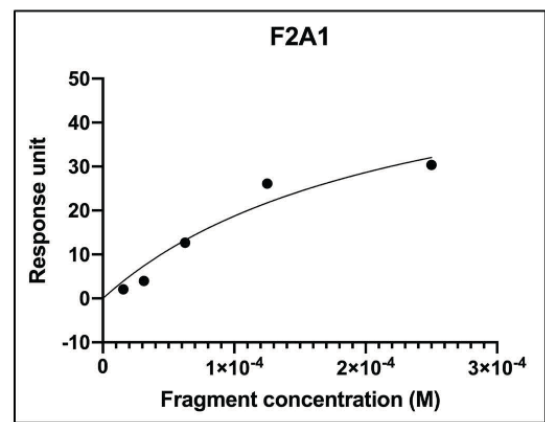
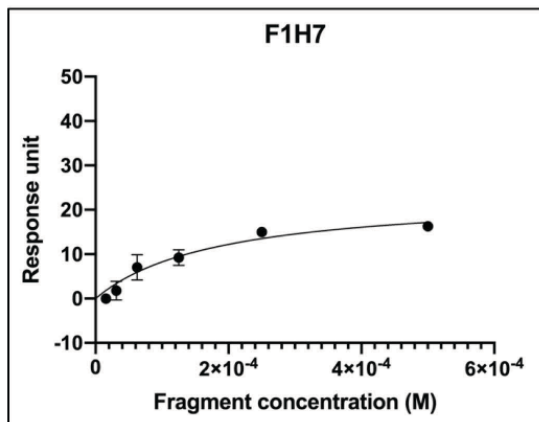
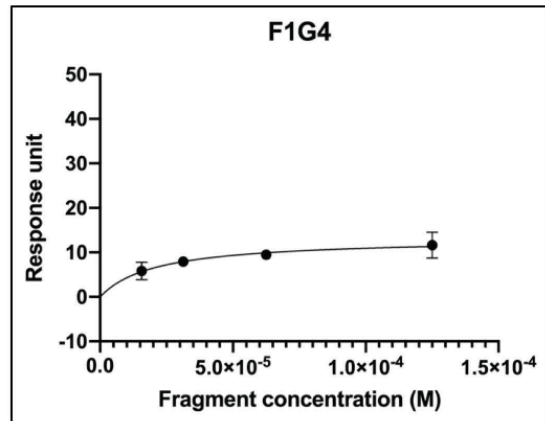
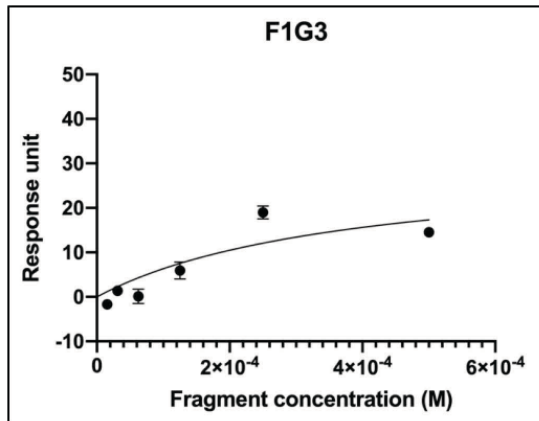
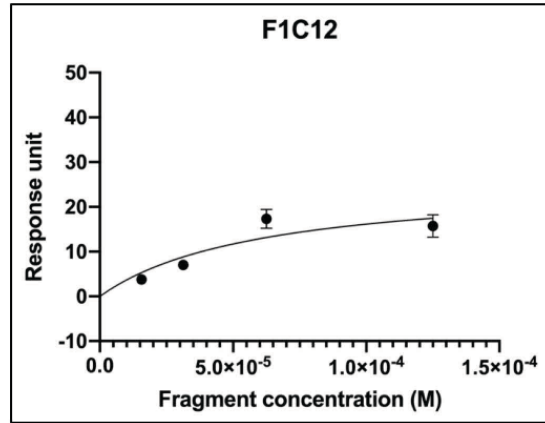
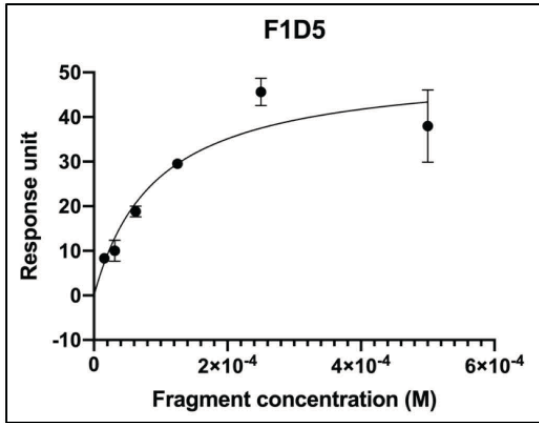
Table 4.1 SPR results

Fragment	Dissociation Constant (K_d ,M)	Maximum Response (Rmax)	R squared
F1A10	7,18E-05	43,10	0,8557
F1C8	2,21E-05	27,67	0,9826
F1C12	6,49E-05	38,56	0,8002
F1D5	7,96E-05	53,84	0,9060
F1G3	1,58E-04	31,01	0,7621
F1G4	4,48E-05	11,89	0,9840
F1H7	8,87E-05	25,00	0,9538
F2A1	9,80E-05	53,63	0,9423
F2A5	1,81E-04	17,41	0,8669
F2A5	1,81E-04	17,41	0,8669
F3A8	2,06E-05	7,67	0,5725
F4D1	2,01E-04	22,72	0,9684
F4E10	4,98E-05	27,95	0,8282
F4G7	7,20E-05	23,76	0,8482

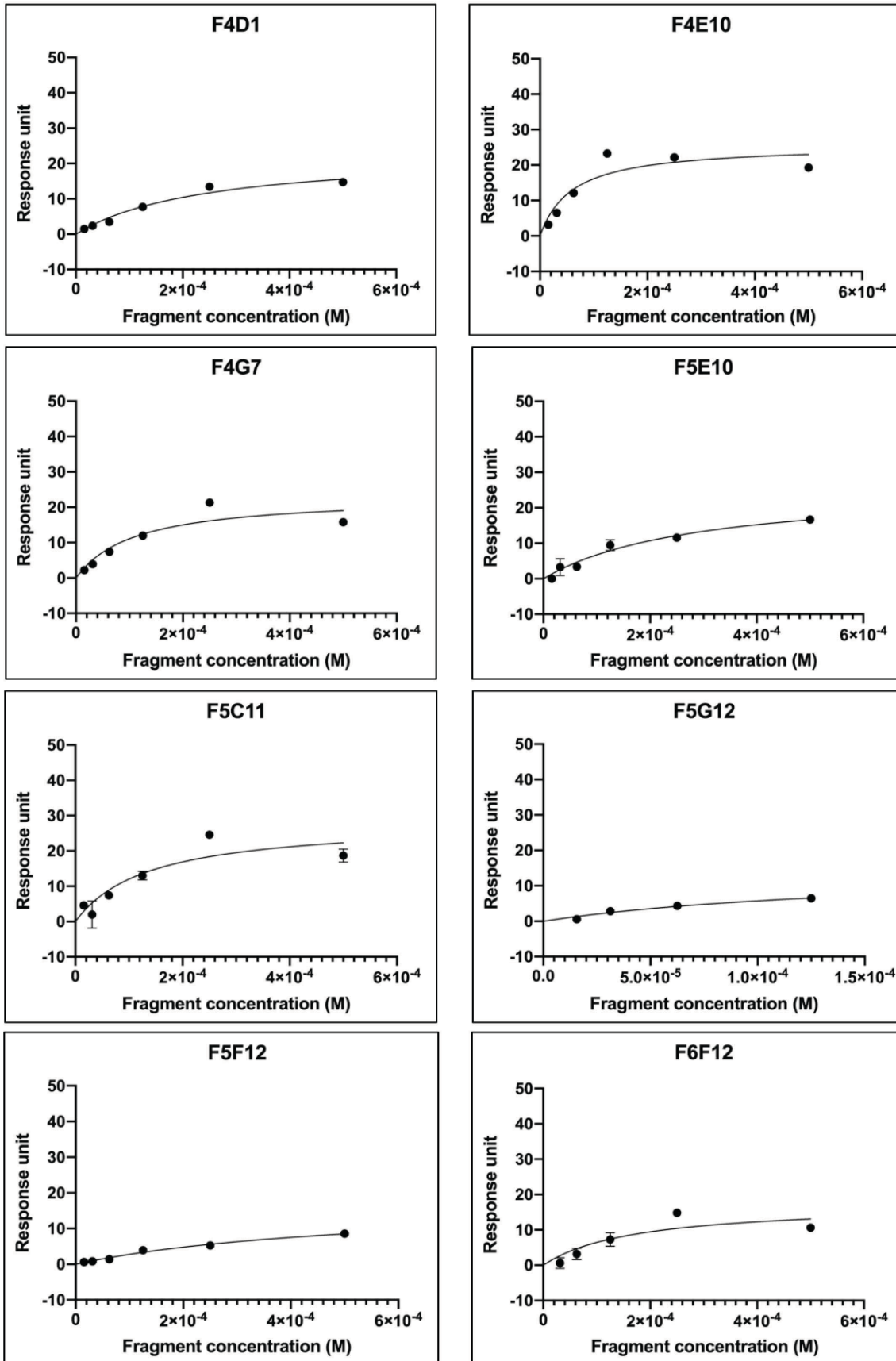
Results Fbw7-Skp1

F5C11	1,18E-04	29,51	0,8218
F5E10	2,02E-04	24,75	0,9658
F5F12	1,12E-04	17,74	0,9863
F5G12	8,01E-05	12,76	0,9602
F6B12	1,02E-03	64,09	0,9791
F6C5	5,48E-04	36,25	0,9291
F6D5	3,02E-05	43,94	0,7994
F6D12	4,25E-04	79,94	0,9915
F6F12	3,71E-05	16,14	0,7507
F6H3	9,54E-06	14,23	0,4399
F7A9	4,69E-04	43,24	0,9960
F7B1	5,00E-04	54,87	0,9938
F7E2	4,95E-05	14,46	0,9988
F7F7	3,89E-04	29,05	0,9747
F7G1	4,49E-04	71,62	0,9955
F7G5	3,84E-04	23,48	0,9909
F7G9	1,69E-05	22,31	0,9215
F7G10	4,34E-04	36,76	0,9938
F1C9	1,74E-03	192,41	0,9939
F1D11	2,74E-04	39,31	0,9731
F2A3	1,31E-03	133,09	0,9756
F2A10	8,09E-04	166,80	0,9911
F2D1	4,12E-04	63,91	0,9654
F2E8	1,57E-03	219,12	0,9995
F2F5	7,60E-04	42,41	0,9933
F2F10	7,12E-04	48,62	0,9869
F2G9	2,75E-04	15,68	0,9962
F3A2	5,96E-04	54,19	0,9970
F3F11	1,01E-03	46,16	0,9795
F3G2	1,29E-03	97,47	0,9967
F4C5	4,49E-04	73,07	0,9921
F5D7	9,38E-04	64,05	0,5258
F5E3	2,38E-04	31,86	0,9920
F5E6	9,24E-04	38,89	0,9360
F5B3	1,06E-03	189,49	0,9970

Results Fbw7-Skp1



Results Fbw7-Skp1



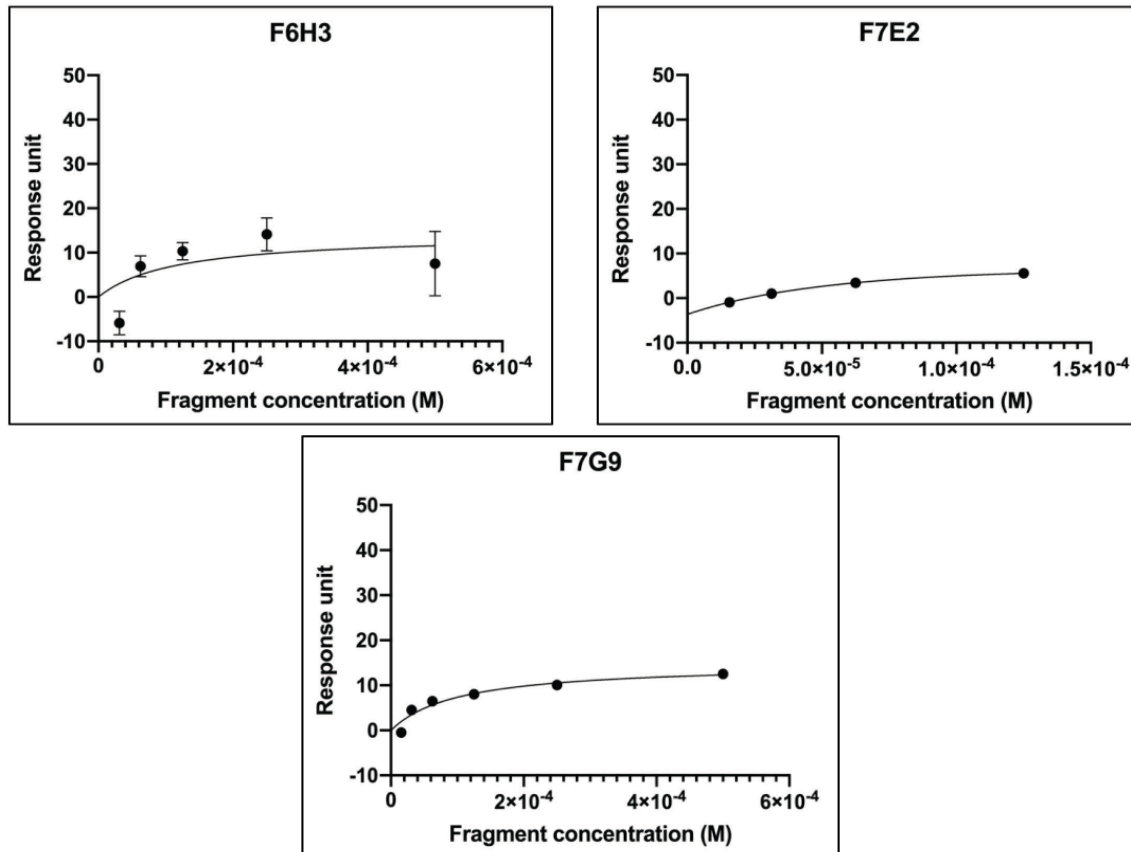


Figure 4.11. Dose-response curves of the most potent fragments

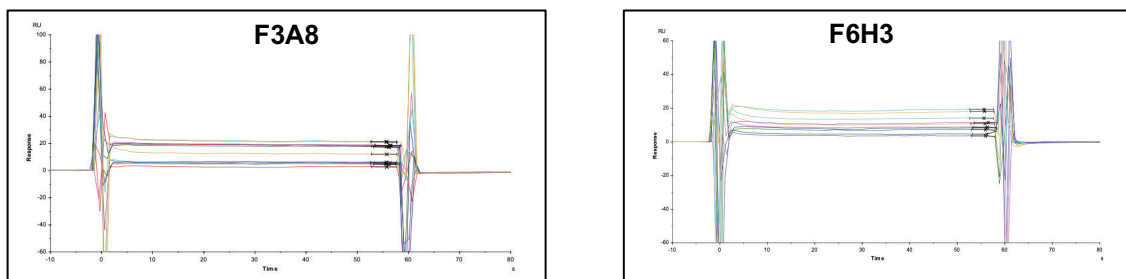


Figure 4.12. Sensograms of the two best fragments, F3A8 and F6H3.

4.4.2.2 Dose response curve assessed by SPR: Fbw7 immobilization with a GST tag

The most potent fragments identified in the previous screening were tested against Fbw7 immobilized in the SPR-chip, in this case through a GST-tag. An antibody anti-GST was first covalently linked to the surface of the SPR-chip and later on the GST-protein was passed over the surface of the chip to get captured by the antibody anti-GST. Molecules were therefore screened against the immobilized protein. Since the GST-tag is located in the Fbw7-

N terminal and far from the degron site, Fbw7 is free to expose its binding site without any restraint that can be generated immobilizing the protein covalently with its lysine. This strategy could seem better to identify potential fragments that bind to the degron site of Fbw7-Skp1 based in the previous result that the Cyclin E peptide did not give any response. GST immobilization offers the possibility of placing the protein in a specific conformation (depending on where the GST is located). Specifically, in our case, this immobilization allowed us to test the peptide positive control and to confirm previously identified fragments with an orthogonal immobilization system.

While the experiment is more insidious to establish i.e. the protein has to be captured every titration cycle, because the binding of GST with the antibody anti-GST is not strong enough to keep the protein in the chip permanently, it represents a good way of increasing the level of confidence on the results obtained with a covalent immobilization. Indeed, there is the risk that a random immobilization in the chip (covalent immobilization) can hide “sweet spots” on the surface of the protein and change its conformation state, giving rise to “new spot” not present in the native conformation of the protein.

To ensure the right immobilization in the surface of the chip and to confirm the correct binding site exposure, a 16-mer diphosphorylated peptide (PEVPPpTPPGpSHSAFTK(FICT)) of the natural substrate of Fbw7, DISC1, was tested. As mention the DISC1 recognition by Fbw7 was published during this PhD thesis and this is the reason why it was not assayed before. The K_d obtained (table 4.2), matches the one reported in the literature [101]. Furthermore, since for further FP assays a fluorophore-peptide would be needed, we purchased and also tested FICT-DISC1.

Once proved that the Fbw7-Skp1 protein was correctly immobilized in the

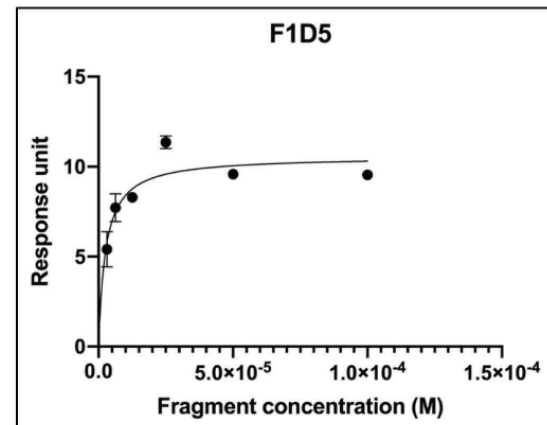
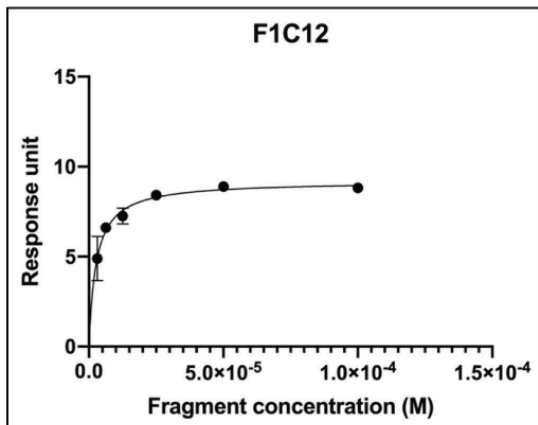
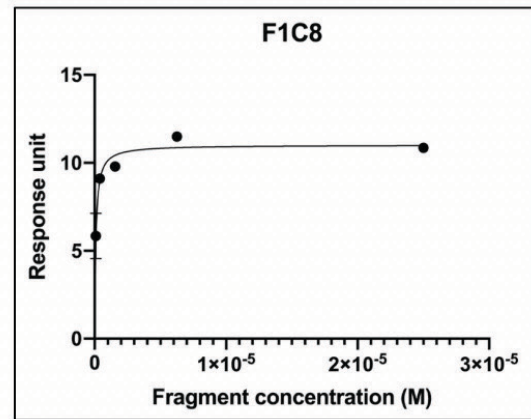
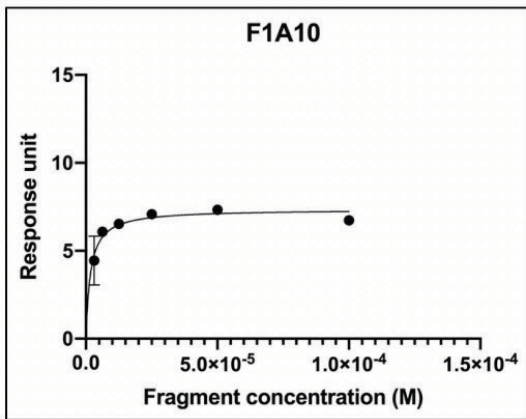
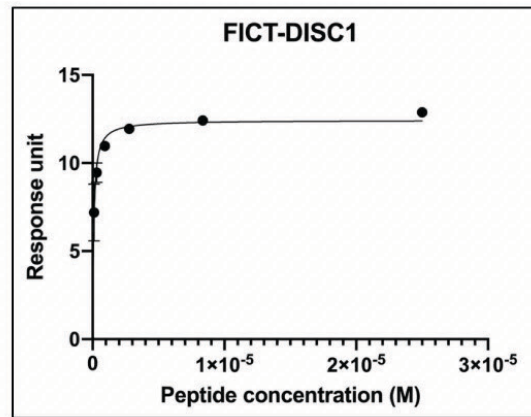
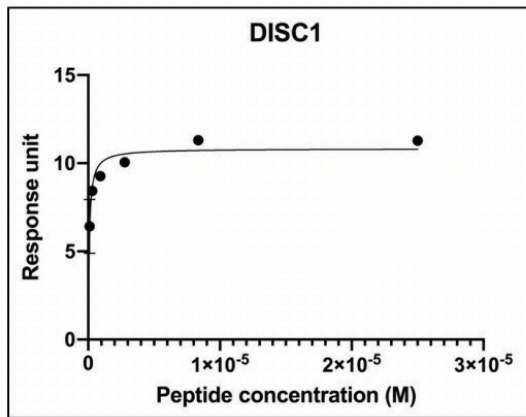
chip with GST tag, the 21 most potent fragments from the primary screening (SPR covalently linked) were screened at several concentrations starting from 100 μ M. Results are shown in table 4.2 and dose response curve in Figure 4.13. Example of two sensograms of the most potent fragments is illustrated in Figure 4.14. To note, compared to the sensograms of the same fragments showed in Figure 4.12, where the protein is covalently attached, there is not a stable baseline probably due to the leaking of the protein during time. Noteworthy, all the fragments bind to the protein in a dose-response manner and the observed K_d in all case improved by two orders of magnitude, compared with the previous K_d observed. However, since the protein is leaking all the time, the K_d values have not to be taken as precise number for the affinity of the fragment to the protein but more as clear demonstration that the fragments still bind to the protein when is exposing the native surface.

Table 4.2 GST-SPR results

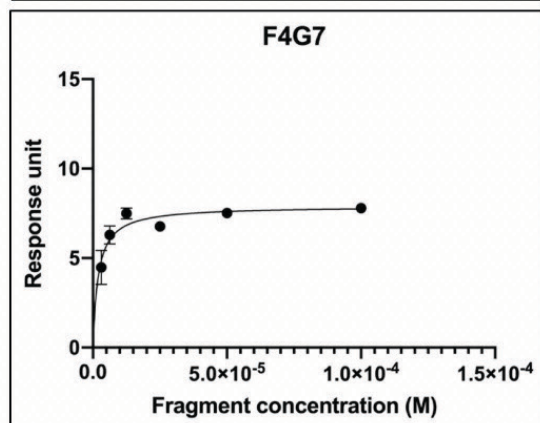
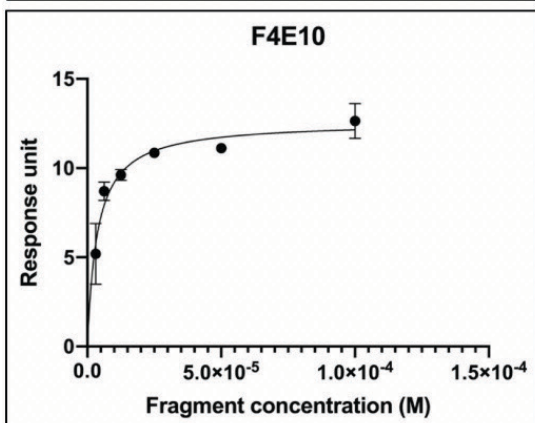
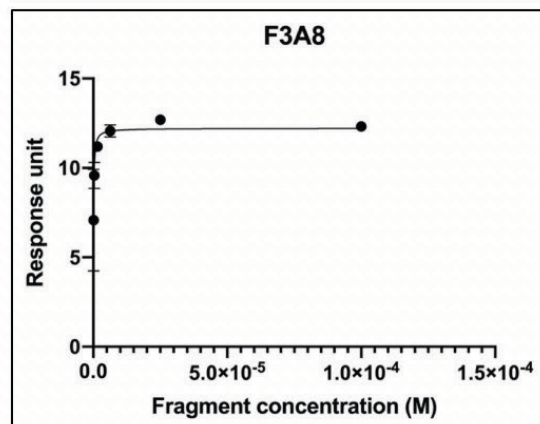
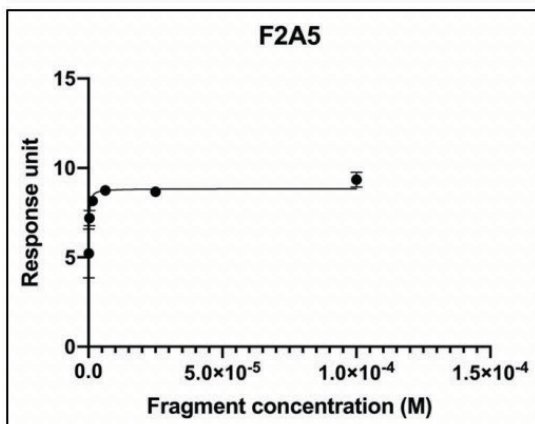
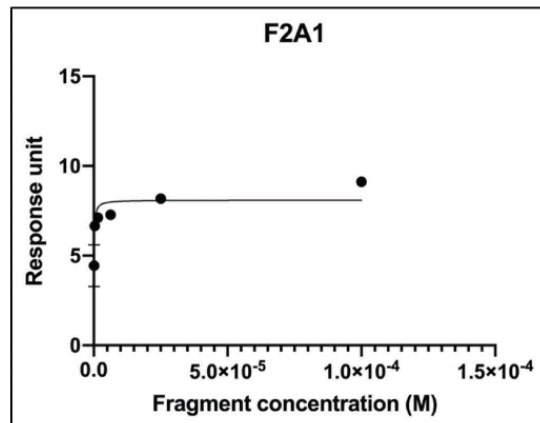
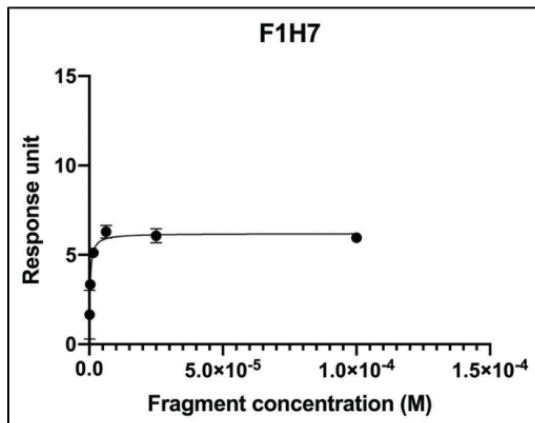
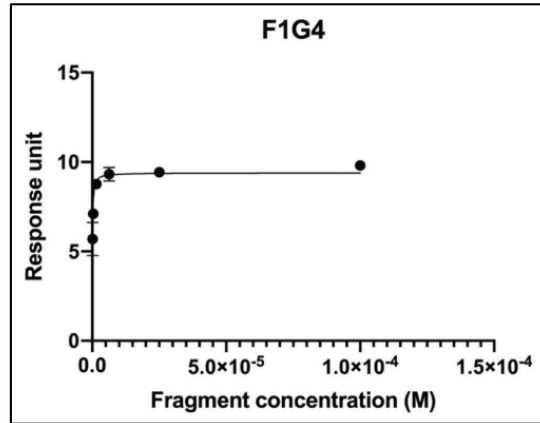
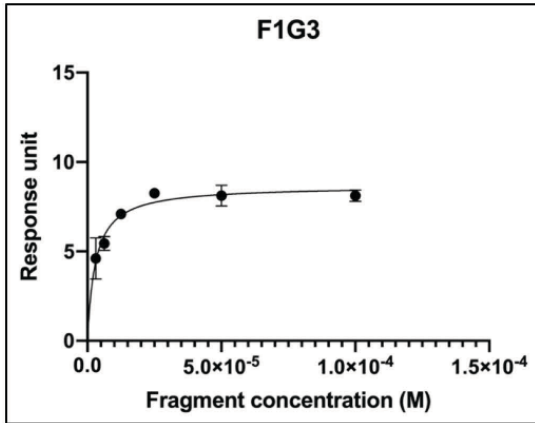
Peptide	Dissociation Constant (K_d ;M)	Maximum Response (Rmax)	R squared
DISC1	4,11E-07	10,83	0,9153
FICT-DISC	2,52E-07	12,43	0,968
Fragment	Dissociation Constant (K_d ;M)	Maximum Response (Rmax)	R squared
F1A10	1,72E-06	7,34	0,9007
F1C8	1,12E-07	9,73	0,9574
F1C12	2,69E-06	9,19	0,9846
F1D5	2,74E-07	10,57	0,787
F1G3	2,98E-06	8,66	0,953
F1G4	3,96E-07	9,39	0,9262
F1H7	3,58E-07	6,19	0,9855
F2A1	2,40E-07	8,10	0,8494
F2A5	2,02E-07	8,84	0,9594
F3A8	2,77E-07	12,21	0,9607
F4E10	3,72E-06	12,60	0,9517
F4G7	1,96E-06	7,90	0,8694
F5C11	7,52E-07	10,27	0,8641

Results Fbw7-Skp1

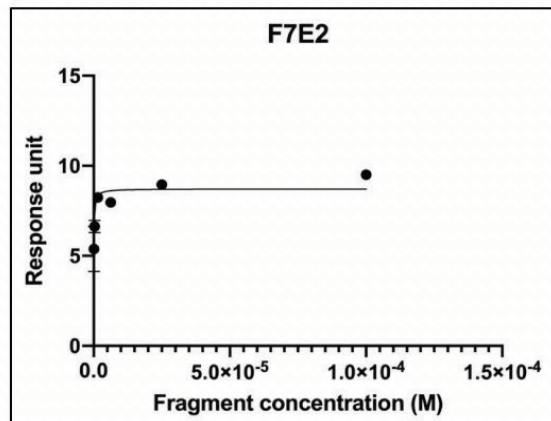
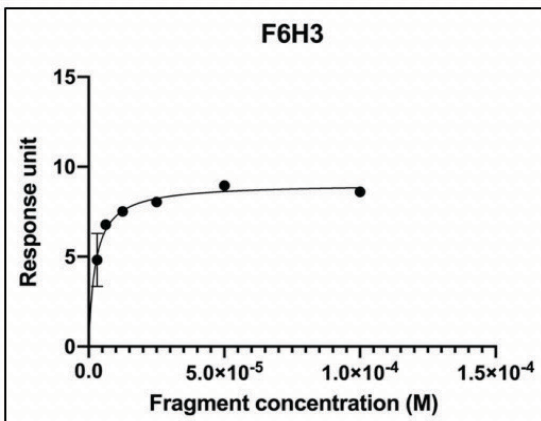
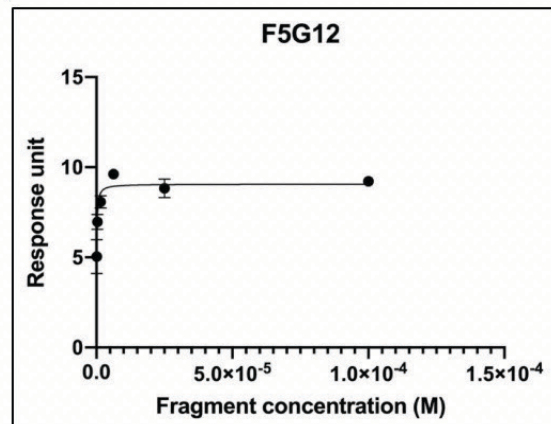
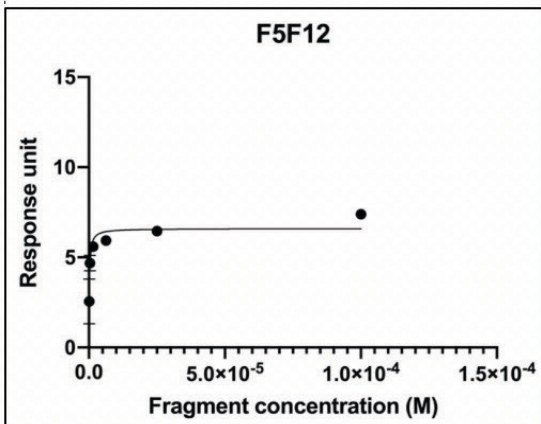
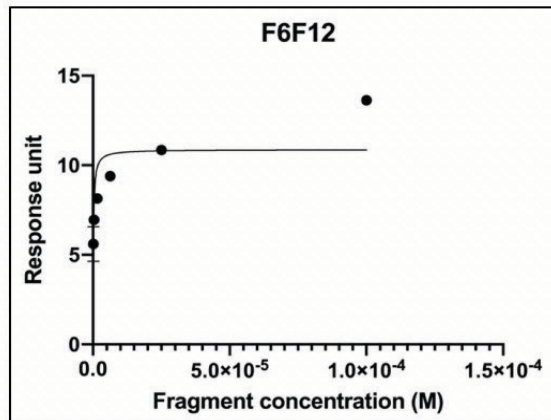
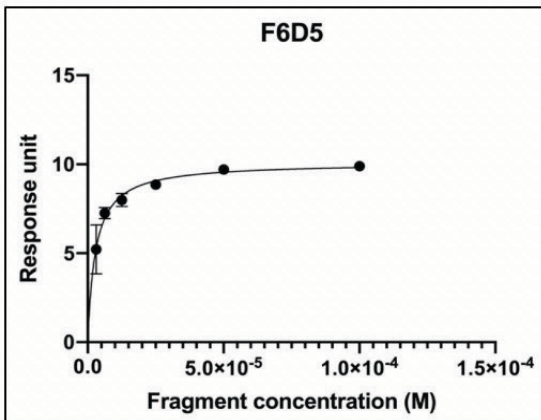
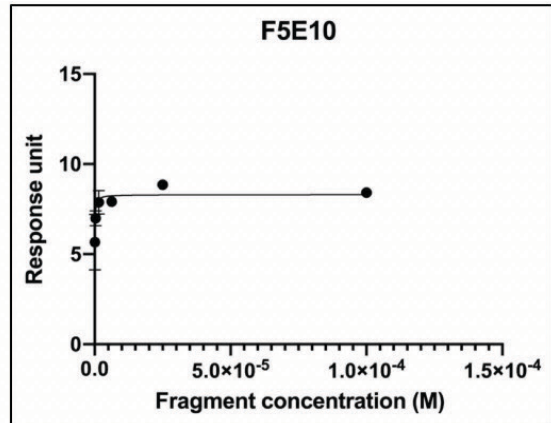
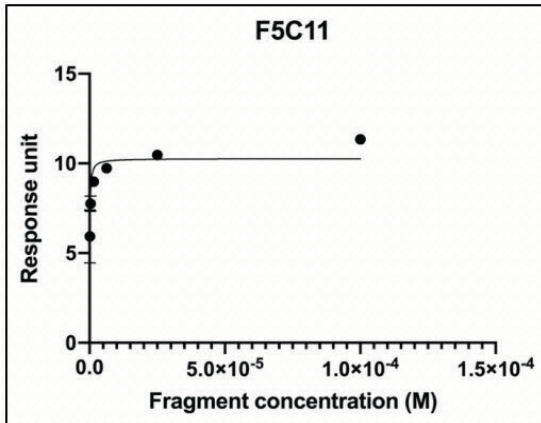
F5E10	2,81E-07	8,31	0,9015
F5F12	2,79E-07	6,60	0,9247
F5G12	2,85E-07	9,07	0,9291
F6D5	2,83E-06	10,10	0,985
F6F12	1,04E-05	10,88	0,6412
F6H3	2,50E-06	9,06	0,97
F7E2	4,99E-07	8,72	0,8466
F7G9	4,26E-08	10,73	0,9774



Results Fbw7-Skp1



Results Fbw7-Skp1



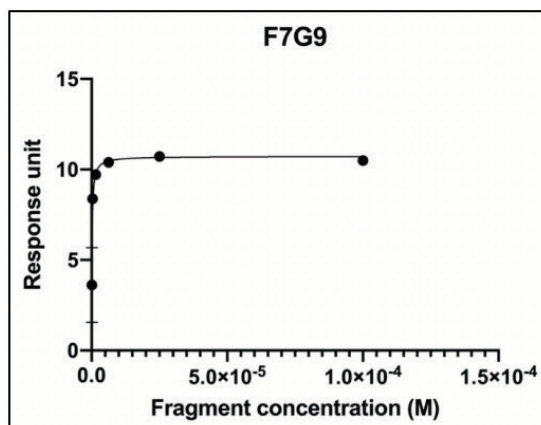


Figure 4.13. Dose-response curves of the most potent fragments with a GST immobilization protocol.

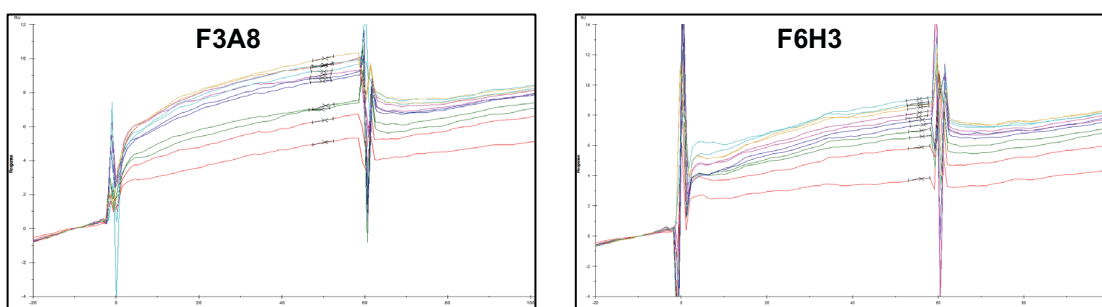


Figure 4.14. Sensograms of the two best fragments F3A8 and F6H3 with a GST-immobilization protocol.

4.4.3 STD-NMR to confirm binding

To further confirm binding of the hits identified with SPR, we decided to carry out NMR studies and we selected STD as orthogonal technique of SPR to screen the fragments. 13 out of the 22 most potent fragments were picked for further investigation and screened as described in section 3.11.4.2. As outcome of a preliminary assessment of the technique we found solubility problems with a fragment concentrations $>250 \mu\text{M}$. To overcome this problem we decreased the concentration of the fragments to $125 \mu\text{M}$. Additionally, some of the fragments required the addition of a detergent (tween-20) to avoid the formation of aggregates, as reported in the literature [101, 102]. The following figures illustrate the results obtained for the fragments giving an STD signal (red spectrum) with the STD % calculated overlaying the STD spectrum with the off-resonance spectrum.

Results Fbw7-Skp1

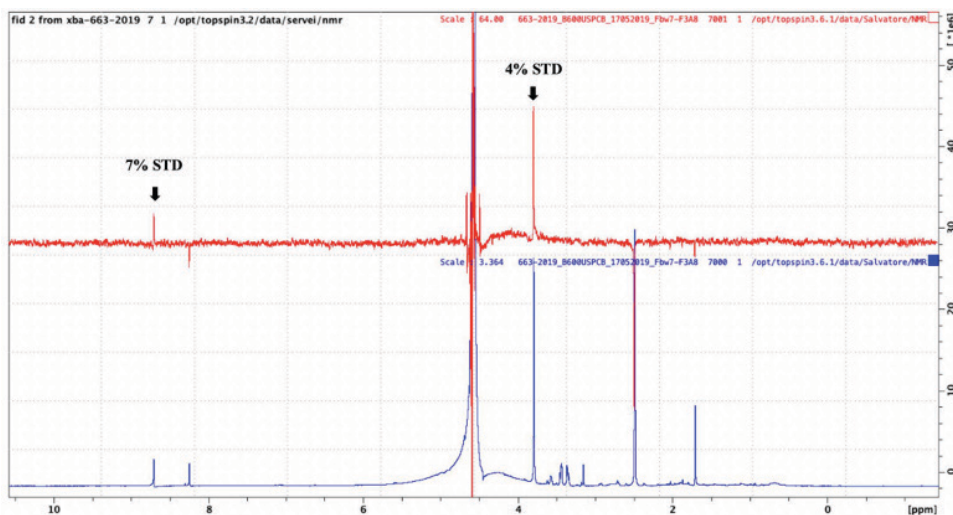


Figure 4.15. F3A8 (250 μ M) in complex with Fbw7-Skp1 (10 μ M).

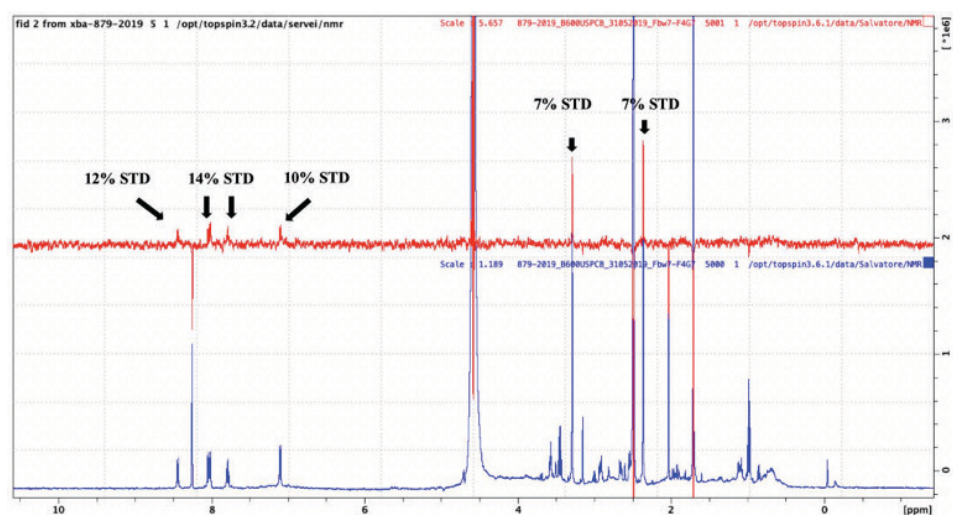


Figure 4.16. F4G7 (125 μ M) in complex with Fbw7-Skp1 (2.5 μ M).

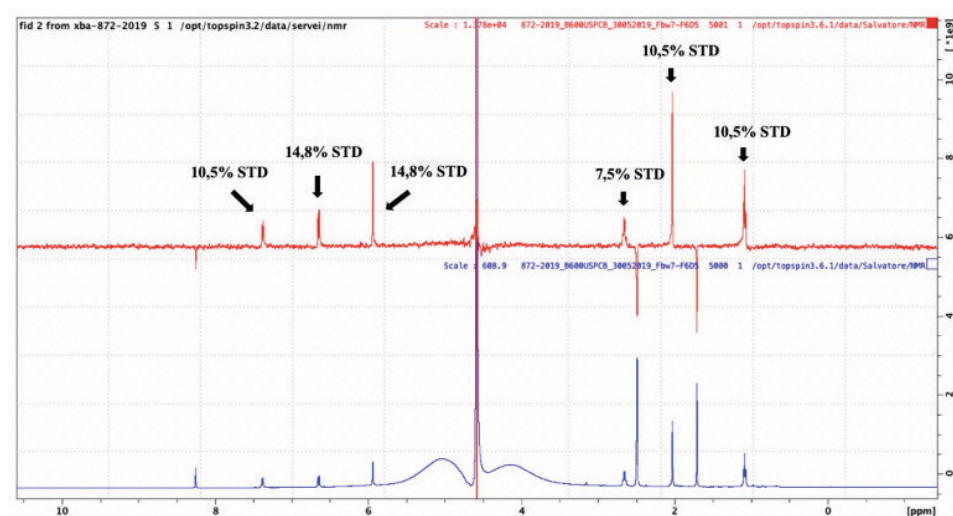


Figure 4.17. F6D5 (250 μ M) in complex with Fbw7-Skp1 (2.5 μ M).

Results Fbw7-Skp1

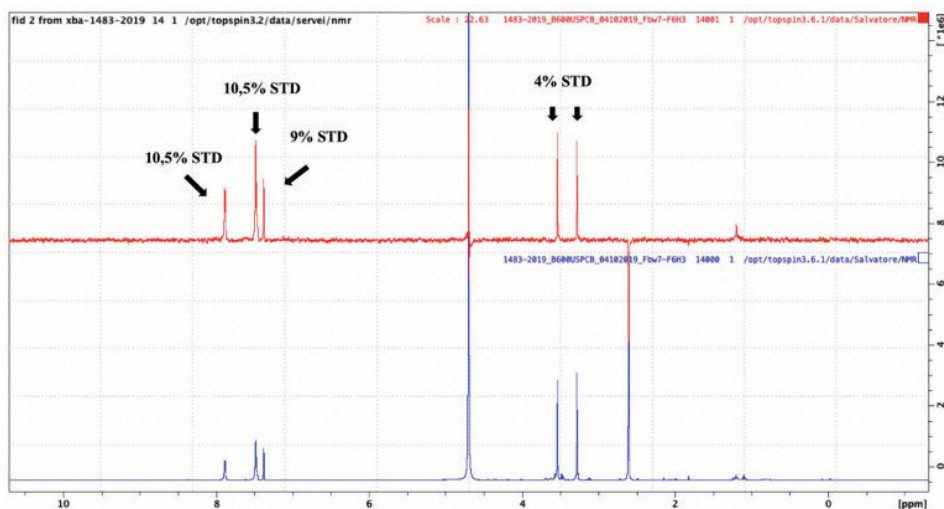


Figure 4.18. F6H3 (100 μ M) in complex with Fbw7-Skp1 (1.25 μ M).

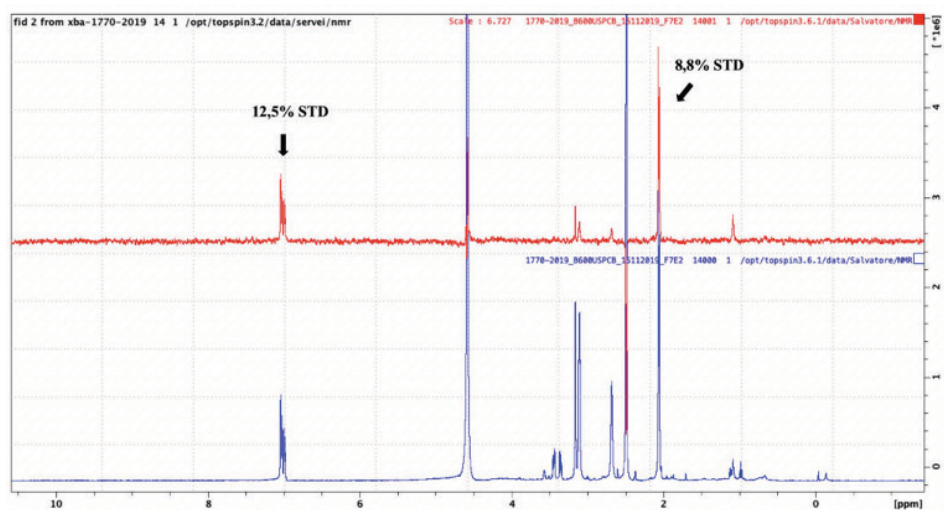


Figure 4.19. F7E2 (100 μ M) in complex with Fbw7-Skp1 (1.25 μ M).

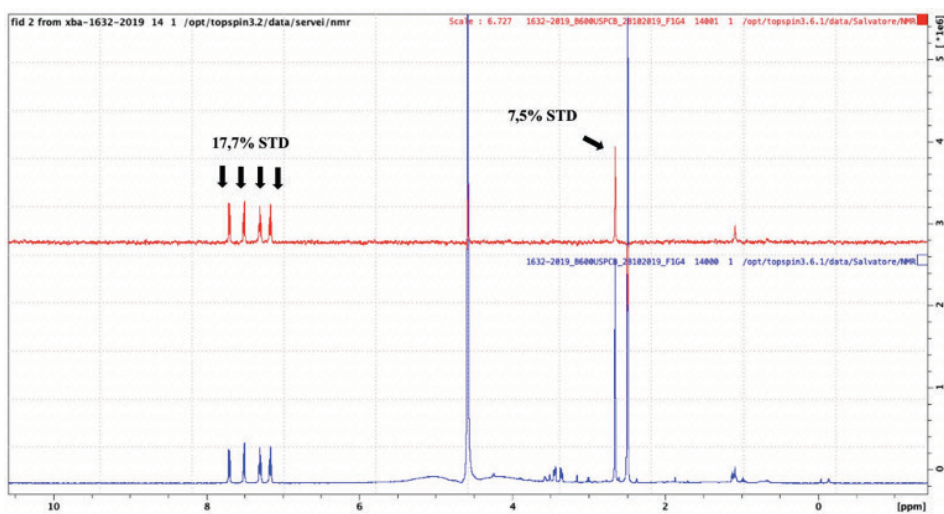


Figure 4.20. F1G4 (100 μ M) in complex with Fbw7-Skp1 (1.25 μ M).

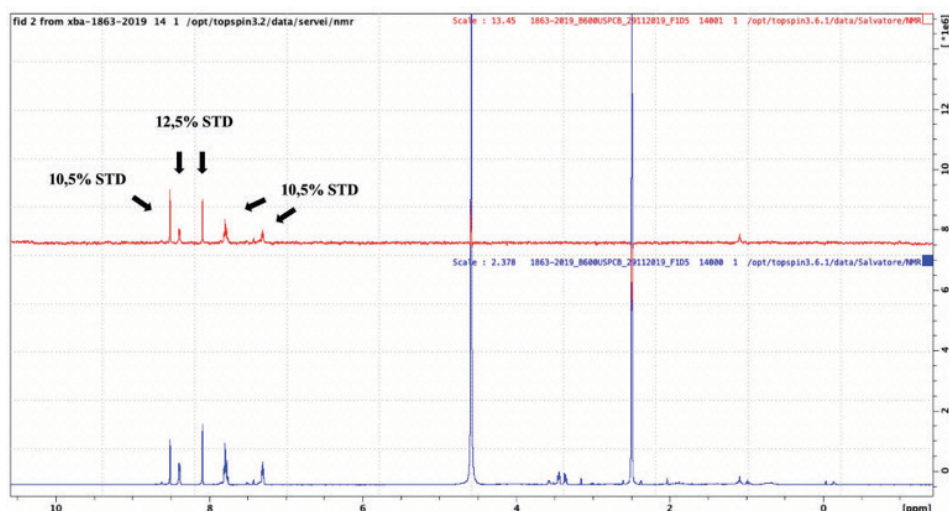


Figure 4.21. F1D5 (100 μ M) in complex with Fbw7-Skp1 (1.25 μ M).

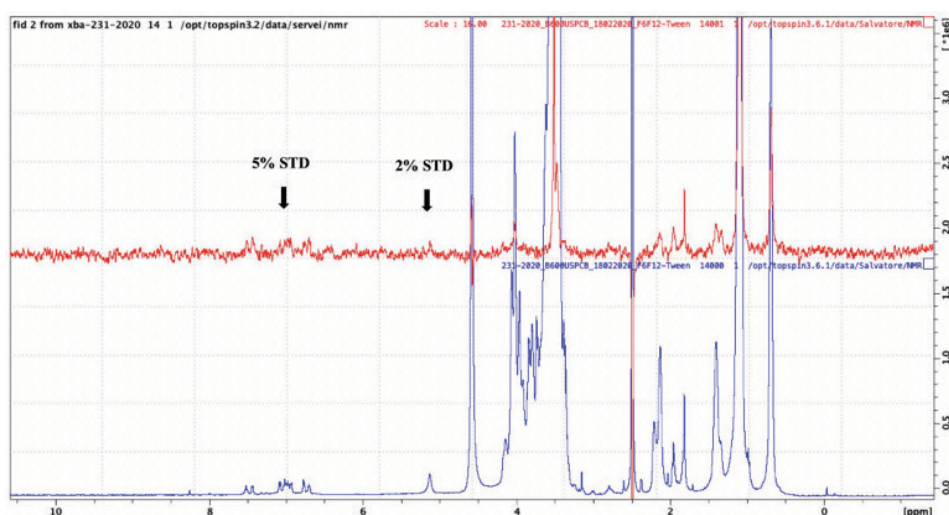


Figure 4.22. F6F12 (250 μ M) in complex with Fbw7-Skp1 (1.25 μ M) 0.05% tween-20.

The presence of tween-20 allowed us to discriminate between false positives and true positives. An example is given in Figure 4.23, in red STD spectrum of fragment F4E10 without protein and in blue off-resonance spectrum. The presence of aggregates give rise to STD signals, most probably due to the fact that aggregates get irradiated at the same magnetic field of the protein. STD signal decrease and almost totally disappear when aggregates break when adding tween-20 (Fig. 4.24). Therefore, after adding Fbw7-Skp1 complex in the solution, we observed STD signal that comes from the binding of the fragment F4E10 to the protein (Fig. 4.25)

Results Fbw7-Skp1

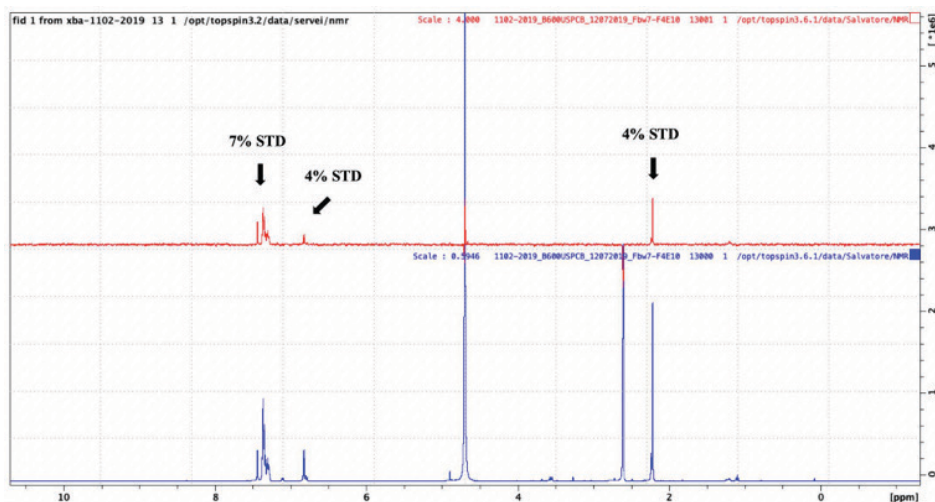


Figure 4.23. F4E10 (100 μ M) in PBS.

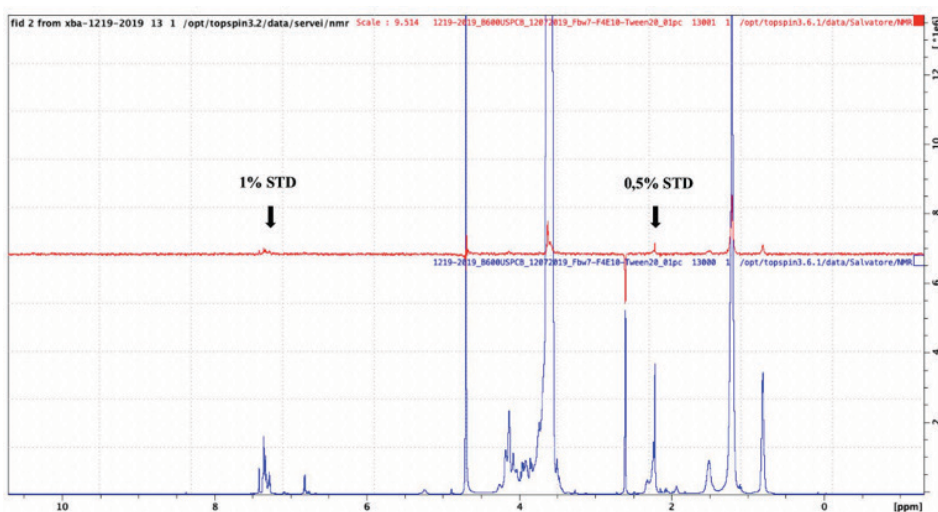


Figure 4.24. F4E10 (100 μ M) in complex with 0.05% Tween20.

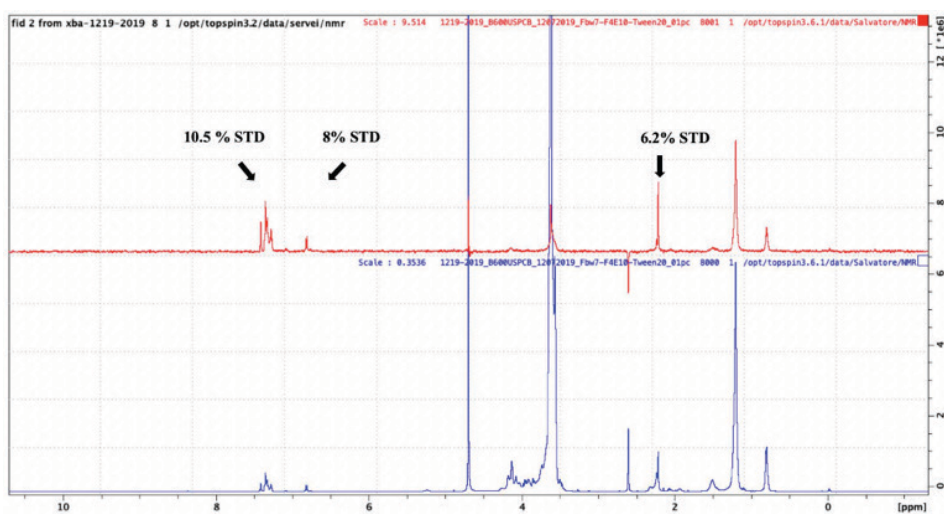


Figure 4.25. F4E10 (100 μ M) in complex with Fbw7-Skp1 (1.25 μ M) 0.05% tween-20.

4.4.4 Elucidation of the binding site of the fragments

4.4.4.1 X-ray crystallography

In order to identify the binding site of the hit fragments, a research collaboration with the group of Dr Bing Hao (Uconn-Health, USA) was launched. Based in the affinity and the chemical structure, we sent fragments F3A8, F4G7, F6F12, F6D5, F7G9, F6D12, F2A5 and F4E10 to Dr. Hao. First attempts to obtain the X-ray structure of the fragments bound to the Fbw7-Skp1 protein with both co-crystallization and soaking were unsuccessful. With some fragments they did not get crystals and with some fragment's crystals did not diffract well enough. However, so far, they have collected 19 datasets including a native data set so that they can use it as a negative control. Unfortunately, they did not find any meaningful density that can be assigned to the fragments. As an example, Figure 4.26 shows the density map of the degron site of Fbw7 soaked with fragment F3A8. It is possible to observe in green an unresolved density that do not correspond to the density of the protein but in this case also do not fit with the fragment molecule and is probably consistent with a sulphate ion. Work is on-going to optimize conditions to obtain these structures.

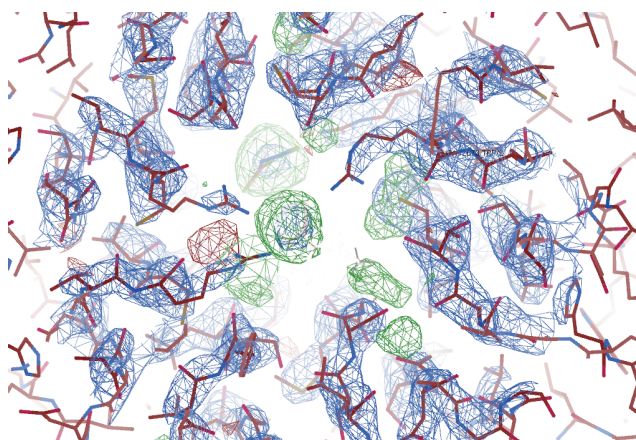


Figure 4.26. Density map of Fbw7-Skp1 soaked with F3A8. It is possible to observe the density map of the degron site of the protein with the corresponding aminoacid and in green the unresolved density map that in this case does not fit with our fragment F3A8.

4.4.4.2 MD simulations

In parallel to crystallography, we run molecular dynamic (MD) simulations to identify probable binding sites. Ten microsecond MD was performed in collaboration with a PhD student of Barril's Lab, Moira Rachman, using a common scaffold of 3 fragments hits in order to see if it has preference for any binding site of the protein. In more details, the fragments were analysed for common patterns and were subsequently deconstructed to contain common features. Fpocket [103-104] (default settings) identified 13 probable pockets in the surface of Fbw7-Skp1 in which a copy of the fragment was added to each of these putative pockets for the simulation. Parameters for the fragments were derived with a MOE [105] SVL script developed to automate DUCK calculations [106]. It calculates AM1-BCC charges [107] and assigns parm@Frosst [108] atom types and non-bonded parameters to the ligand. The System was solvated with the help of pyMDmix [109] which also provided the equilibration and minimization setting to perform 10 microseconds of unconstrained MD. Finally, all trajectories were aligned and cpptraj (dbscan) was used to cluster the density of the fragments throughout the simulation. Surprisingly we found fragments binding close to the degron site of the complex as shown in Figure 4.27.

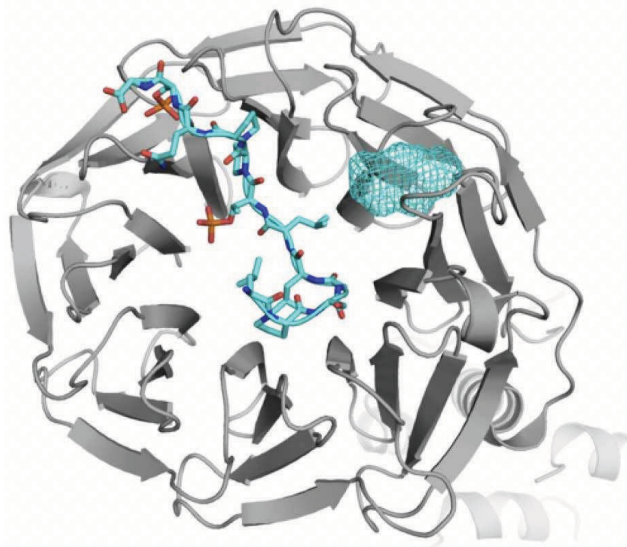


Figure 4.27. Most probable binding site of the fragments with common scaffold. In grey degron sites of the WD40 domain of Fbw7 binding to the peptide of the natural substrate cyclin-E (blue sticks) and fragment density (blue surface).

4.4.4.3 2D-NMR (CSP)

We also applied to the iNEXT discovery program to access Utrecht NMR facility, in collaboration with Dr Hugo Van Ingen, to perform 2D-NMR (CSP), to obtain a full mapping where the fragments bind. We now have ready a sample of ^{15}N -labeled Fbw7-Skp1 complex (see protein expression and purification below) for characterization and feasibility study. The goals of this first experiment are to i) assess spectral quality of the Fbw7-Skp1 construct, including temperature and long-term stability, ii) test addition of two candidate fragments ($K_d \sim 40 \mu\text{M}$ and $1 \mu\text{M}$) to the sample to see whether spectra changes can be detected.

We routinely express Fbw7 in complex with Skp1 (149 residues), an essential adapter protein within the E3 complex. The overall mass of this multisubunit protein is around 70 kDa, quite challenging for NMR, since the spectrum can result crowded and difficult to interpret. Fortunately, the backbone and side chain assignments of Skp1 have previously been determined (BMRB id 26765). Note that even without assignments,

compounds can likely be clustered based on their induced spectral changes, which would be very useful to target the crystallisation efforts.

4.4.4.3.1 Cloning, expression and purification of Fbw7-Skp1 complex ^{15}N labelled in E.coli

The expression of the ^{15}N labelled protein was required in order to elucidate the binding sites of the fragments using 2D-NMR. The same vector was used to express Fbw7-Skp1 complex ^{15}N labelled. M9-Marley expression was followed as describe in section 3.7.2. We obtained the same yield of the unlabelled protein expression, 0.3 mg/l.

For this first assessment to perform a sample characterization to probe the feasibility of studying Fbw7-drug interactions by NMR, 200 μl at a concentration of 100 μM were needed, which corresponds to 1.4 mg. Since this protocol was not performed before, optimization was carried out. As primary trial we followed a standard ^{15}N labelling procedure as specified in section 3.7.2, but with the only exception of not including solution Q. In fact, we first wanted to assess whether the protein could be expressed using a standard M9 protocol. This first trial led to a very low yield (0.08mg/l). In order to boost the production of the protein we decided to use a micronutrient solution, known as solution Q. As result we managed to reach at least the same yield of the unlabelled peptide (0,3mg/l). The purification was performed following the same procedure described in section 4.4. Purification steps were monitored by gel electrophoresis and protein purity was assessed by mass spectrometry (Fig. 4.28). To date, 12 L of E. coli have been produced with a total amount of 3,6 mg of protein, which ensure us that at least two fragments at one single concentration (200 μM) will be assayed by our collaborators.

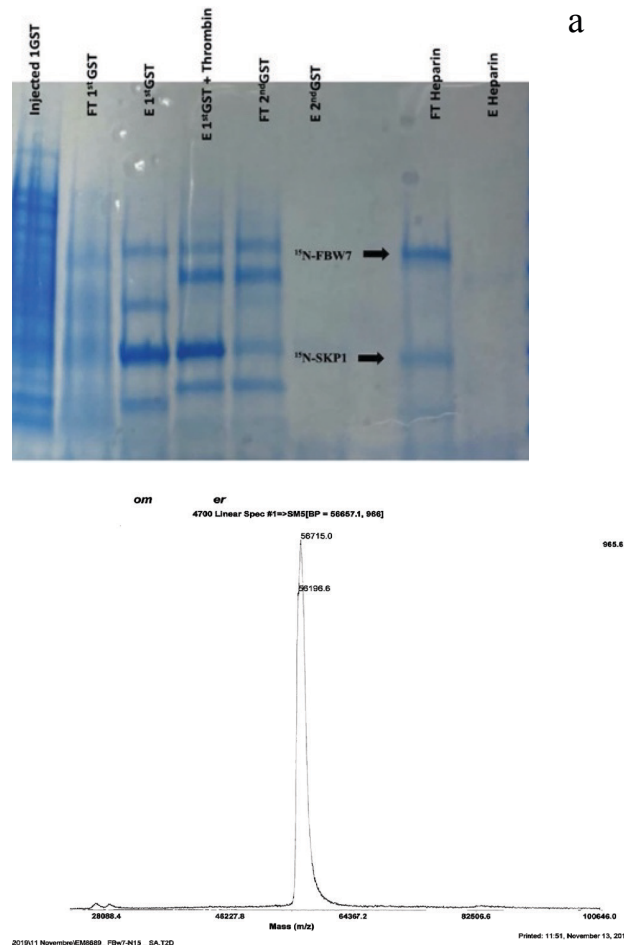


Figure 4.28. a) Electrophoresis gel showing protein bands for Fbw7-Skp1 ¹⁵N labelled, stained with Coomassie blue. b) Matrix-assisted Laser Desorption/Ionization (MALDI) shows Fbw7 (56 KDa) and Skp1 (16 KDa).

4.5 Mechanism of action elucidation

4.5.1 Fluorescence polarization Assay

FP assays were performed in order to evaluate the mechanism of action of the hit fragments. This technique would help us to disentangle if the hit fragments are able to disrupt the PPI between the Fbw7 E3 ligase and their substrates by direct interaction in the degron site or allosterically. In fact, Orlicky *et. al* by screening a library of compounds using FP discovered SCF-I2, a biplanar di-carboxylic acid fragment that is able to bind to Cdc4 protein, the yeast homolog of Fbw7 [97]. Structural studies revealed that SCF-I2 intercalates between the adjacent blades 5 and 6 of the conserved WD40

domain of Cdc4 and in this way induces formation of its own binding pocket located at 25 angstroms of distance from the degron-recognition site (PDB 3MKS). Interestingly, this new pocket is not found in the apo Cdc4-Skp1 complex (PDB code 1NEX), but the interaction of the SCF-I2 with this new allosteric pocket is enough to inhibit the recognition of the corresponding degrons in the degron-recognition site. The authors propose that similar allosteric approaches could be employed for developing ligands of other similar WD40 domains proteins that serve as substrate recognition subunits within the SCF-E3 ligase family.

Since the FP experiment was not still implemented in the Lab, first we needed to set up the optimal conditions, in collaboration with a PhD Student in the Lab, Andrea Bertran. We tested PBS, Hepes, BSA, tween-20, DTT and different pHs. At the end, the established conditions were: 50 mM Hepes (pH= 8.0), 50 mM NaCl, 0.01% tween-20 and finally 1 mM DTT, the last to be added to avoid aggregation and protein absorption onto the plate wells. The incubation time was imposed at 30 min because it was enough not only to let the components homogenize between them, but also to avoid fluorescence loss. In order to extrapolate the K_d of the labelled peptide (substrate of Fbw7, FICT-DISC1), a dose response curve was performed with a fixed label-ligand concentration of 50 nM and different increasing protein concentrations: 0 μM – 0.19 μM – 0.39 μM – 0.78 μM – 1.56 μM – 3.125 μM – 6.25 μM – 14 μM – 28 μM , as it is shown in Figure 4.29. The resulting K_d obtained from different experiments performed at different days was between 1 and 4 μM , according to the one of the unlabelled peptide reported in the literature [95].

Results Fbw7-Skp1

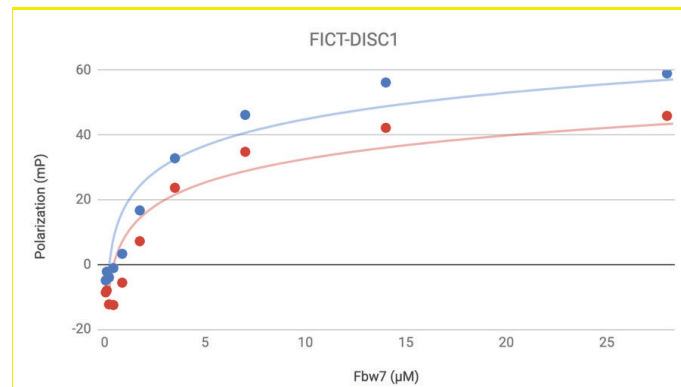


Figure 4.29. Curve replicates showing binding of the labelled peptide FICT-DISC1 to Fbw7-Skp1.

The labelled peptide (or positive control) was also tested at 50 nM with a serial protein concentration from 28 µM to 190 nM in the presence of the unlabelled substrate DISC1 at 1 µM (in excess) to achieve a notable displacement of the labelled one during the competition. The results are shown in Figure 4.30, proving the feasibility of the experiment once a competitor is added.

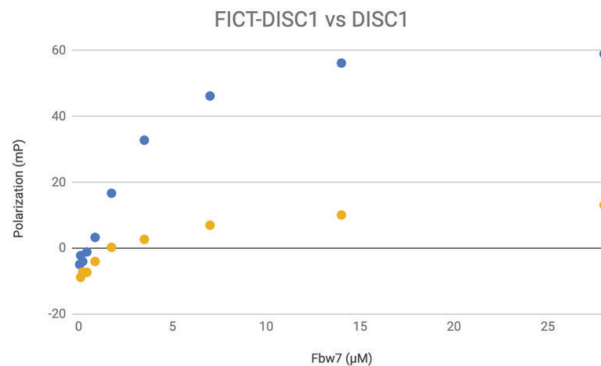


Figure 4.30. The plot shows the average of the two replicates, the curve of the labelled peptide alone (blue curve) and the competition of labelled and the unlabelled peptide (yellow curve).

To assess fragments competition, a dose-response curve was performed by fixing Fbw7-Skp1 concentration at 10 µM according to the quality of the assay calculated using Z' factor extracted from previous experiments performed and several concentrations of fragments: 0 µM – 6.25 µM – 12.50 µM – 25 µM – 50 µM – 100 µM – 200 µM – 400 µM – 800 µM. No inhibitions

were recorded for the fragments tested as illustrated in the following picture (Fig. 4.31).

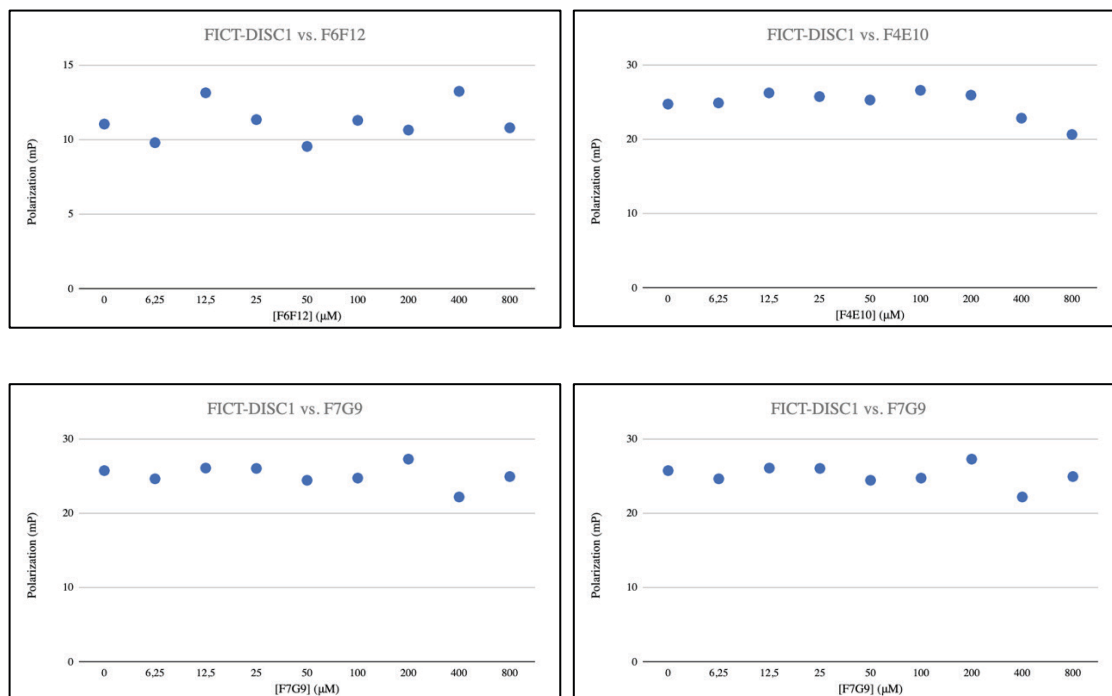


Figure 4.31. *IC₅₀ of FICT-DISC1 in the presence of different fragments.*

To avoid fragments aggregation and precipitation at the highest concentration (800 μM), 8% of DMSO was used in this case. As a positive control of the assay, an IC₅₀ of FICT-DISC1 was run at different concentration of DISC1 with 8% DMSO: 0 μM – 0.023 μM – 0.047 μM – 0.093 μM – 0.1875 μM – 0.375 μM – 0.75 μM – 1.5 μM – 3 μM. FICT-DISC1 was maintained at 50 nM and the protein was fixed at 10 μM. Result is shown in Figure 4.32 and proves that the increase in DMSO does not affect either the stability or the protein or the binding of the natural substrate DISC1.

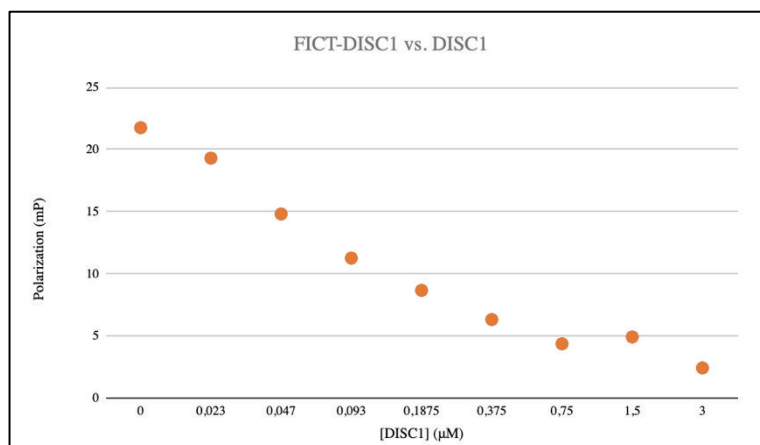


Figure 4.32. IC_{50} of FICT-DISC1 in the presence of different concentrations of DISC1 once using a high percentage of DMSO (8%).

4.6 Folding and stability of Fbw7-Skp1 complex

DMSO is the standard solvent for preparing stock solutions of compounds for drug discovery. The assay concentration of DMSO is normally 0.1% to 5% (v/v) or 14 to 715 mM. Thus, DMSO is often one of the principal additives in assay buffers. This standardization of stock solutions does not eliminate possible pitfalls associated with the effect of the DMSO-containing solutions on individual proteins, leading to a partial or total protein denaturation [110]. Therefore, a proton spectrum of 40 μ M Fbw7-Skp1 complex in PBS (50 mM NaCl, 94 mM K_2HPO_4 , 6 mM KH_2PO_4 , pH= 8.0) with 10% D_2O was recorded at different hours in order to check the stability and folding of the protein, in the absence and in the presence of 5% of DMSO. As it appears in Figure 4.33, the proton spectrum is not affected by the addition of DMSO, indeed the signals remain unchanged. After 24 h of the addition of DMSO another proton spectrum was acquired, showing that the signals start to give a different shape of proton signals. Since the presence of sharper picks are common to smaller system and they become clearly visible at this time, they may suggest that there is a partial unfolding after certain time or even partial degradation, confirmed then by a further acquisition of the spectrum at 48 h after the addition of DMSO (Fig. 4.34).

Results Fbw7-Skp1

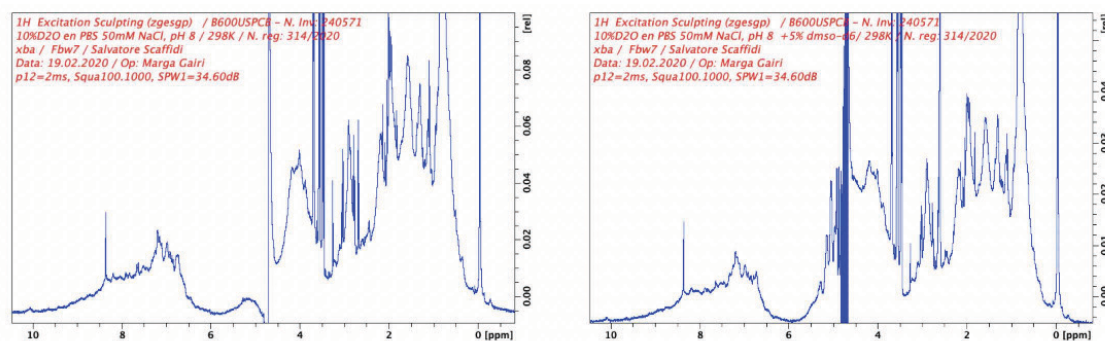


Figure 4.33. On the left ^1H -Fbw7-Skp1 complex in PBS pH= 8.0. On the right ^1H -Fbw7-Skp1 complex in PBS 5% DMSO pH= 8.0.

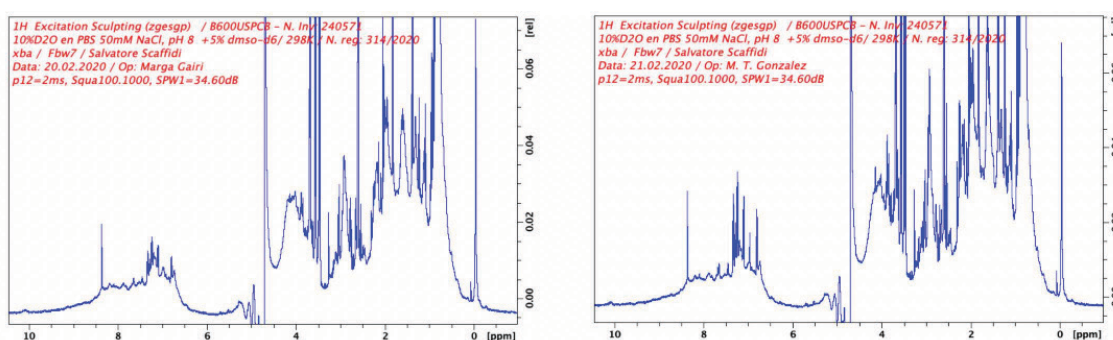


Figure 4.34. On the left ^1H -Fbw7-Skp1 complex in PBS 5% DMSO pH= 8.0 after 24h. On the right ^1H -Fbw7-Skp1 complex in PBS 5% DMSO pH= 8.0 after 48h.

4.7 Summary and future perspective

FBDD represents a useful tool to assess the *ligandability* of a protein since fragments are efficient probes to explore the chemical space at protein surfaces [111] and fragment hits can be an effective starting point for a drug discovery programs [112]. The success of the fragment-based approach is depicted by the release in the market of four fragment-derived compounds treating patients: erdafitinib, pexidartinib, vemurafenib and venetoclax. Many other fragment-derived drugs are under investigation in clinical phase 1 and 2 and another drug, asciminib is close to touch the market, being already in clinical phase 3.

Given the importance of Fbw7 E3 ligase, we embarked on a fragment campaign in order to explore the propensity of Fbw7 protein surface to bind

fragments that can be used as starting point to obtain chemical probes or PROTAC molecules.

In parallel, Miriam Martinez, a PhD in the Galdeano Lab, has designed a computational approach to identify small-molecules able to bind to Fbw7 and other undruggable E3 ligases. She identified druggable sites at the Fbw7-Skp1 protein surface using MDMix [109]. Virtual screening of the previously found regions was performed with the open-source docking software rDock [50]. 7 vendor libraries (around 7 million drug-like compounds) were docked with pharmacophoric restraints derived from MDmix simulations. The top-scoring molecules were re-evaluated using dynamic undocking (DUck), a recently developed method based on molecular Dynamics that removes a large proportion of docking false positives [106].

In this PhD thesis, to screen the in-house fragment library, we decided to pick as primary technique DSF, that has the advantage of being a fast (high-throughput) and cost-effective (real-time thermal cycler widely available and cheap reagents) technique and the equipment easy to use. DSF represents a powerful technique in early drug discovery programs to assess protein *ligandability* as it is described by Molly Chilton et al [113]. Unfortunately, DSF, as showed in Figure 4.9, was not the best technique to explore Fbw7-Skp1 complex either for the complexity of the protein since it has different domains and two different proteins that make difficult the interpretation of the melting temperature or because DSF is not sensitive enough to reliably detect the binding of compound with a K_d close to 1 mM.

We then selected SPR to screen the in-house library of approximately 700 fragments. SPR, indeed, together with NMR and X-ray crystallography is one of the most applied biophysical techniques in FBDD since it is fast, cost-efficient and quantitative [114]. Since we were interested in finding

fragments binding in all possible sites of Fbw7-Skp1 complex and not specifically to the degron site, we first decided to immobilize Fbw7-Skp1 complex with an amine coupling reaction. With the aid of a blind fragment-screening we were able to find 220 fragments as positive fragments to be further investigated with a dose-response curve. All of them were screened at several dilutions and 500 μ M 5% DMSO was selected as the maximum concentration to use to avoid aggregation at higher concentration. Surprisingly 21 fragments gave a $K_d < 200 \mu$ M with an overall hit rate of the screening of approximately 3%, an encouraging number for a target that until now was considered undruggable.

Motivated by these promising results we then decided to change the immobilization protocol of the protein in SPR. Indeed, since in the previous screening most of the fragments showed saturation with a lower response than expected ($R_{max} = 55$, see Fig 4.10), we thought that, due to the random immobilization, part of the protein present on the surface of the chip could be not active for that specific binding site, therefore we proceeded immobilizing Fbw7-Skp1 complex with a GST-immobilization protocol. In this way the protein is free to expose its entire surface avoiding the constraint generated by linking the protein with its lysine. With that, it was possible to detect binding of the peptide of a natural substrate of Fbw7 (DISC1) and its labelled fluorescent form (FICT-DISC1) since GST-immobilization allows the correct solvent exposure of the degron site of the protein and as expected all the fragments displayed binding (table 4.2 and Fig. 4.13). Best practice for finding fragments includes the use of orthogonal techniques and for this reason 13 hit fragments out of 21 were also tested with STD-NMR. Impressively, nine fragments confirm binding, making us envisage that they can be good starting point for the optimization into lead compounds. Noteworthy, in order to ensure that the high percentage of DMSO employed

(5%) was not affecting the folding of the protein, we carried out a proton spectrum analysis of Fbw7-Skp1 complex in the absence and presence of 5% DMSO. We observed that after 24 h of the addition of DMSO there is a slight difference in the proton signals suggesting a partial unfolding of the protein that seems to remain unchanged after 48 h.

In order to elucidate a putative mechanism of action of the fragments, we asked whether the most potent hit fragments could affect the binding of the peptide of the natural substrate DISC1 to Fbw7. Unfortunately, we did not observe disruption with DISC1 peptide, probably due to the fact that fragments are too small to inhibit directly or by an allosteric modulation the peptide or they may bind to a different domain of Fbw7 or to Skp1, therefore not affecting the activity of the protein.

Finally, we should not discard the possibility of a conformational change of Fbw7 that may not affect the binding of the natural substrate but it may affect the orientation of substrates that will be reflected in an increase or decrease of ubiquitination made by the E2 enzyme. Work is on-going in the lab to dive deep into this mechanism with an *in vitro* ubiquitin-assay or cell-based assays.

One of the mayor questions that arises after identifying a fragment is what to do with it. Before answering this question, it could be worth it to know exactly where the fragment binds, how it binds and if it could influence the function of the protein. Based on this, optimization can be followed up to have a “lead compound” able to modulate the function of the protein as inhibitor or enhancer. Willing to assign a binding site and a binding mode of the fragments we established a collaboration with Dr Bing Hao (UconnHealth, USA) that first elucidated the crystal structure of Fbw7 in 2007 [96]. Furthermore, we applied to the iNEXT program in order to establish a collaboration with the group of Dr Hugo Van Ingen (Utrecht) with

the aim to obtain a full mapping where the fragments bind by 2D-NMR (CSP). Work is undergoing and probably soon important information will be unveiled.

In parallel we run MD simulations to identify probable binding sites. Ten microsecond MD was already performed in collaboration with Dr Moira Rachman and as first outcome we found a probable binding site for a fragment close to the degron site of the complex as showed previously in Figure 4.27. Based on this preliminary results coming from MD simulation we could assume that growing these particular fragments with common scaffold to a larger molecules could somehow interfere with the binding of the natural substrate or it could be a good starting point to develop molecular glues able to enhance the binding of the natural substrates, as demonstrated recently by researchers at Nurix, where they were able to discover and rationally design, small-molecule molecular glues that enhance a substrate-ligase interaction to induce substrate degradation. These molecules potentiate the binding of S37A mutant β -catenin to β -TrCP [115].

In this scenario we can also consider the possibility to use the fragments to hijack the ubiquitin proteasome system by recruiting Fbw7 with PROTAC molecules where in one hand we will have a fragment binding to Fbw7, probably allosterically, and in the other hand a ligand of a POI connected by a linker.

PROTAC opens up the advantage of employing fragments although they do not show any activity for the modulation of the protein function. What matters is the possibility to anchor PROTAC to an E3 ligase for POI ubiquitination. Indeed, once the ternary complex is formed (E3-PROTAC-POI) the UPS machinery will lead to the destruction of the POI. Furthermore, it has been proven that although a ligand is promiscuous with no high affinity or high selectivity for a protein, once assembled in a PROTAC molecule is

able to selectively destroy a POI due to the unique protein-protein interactions between the E3 ligase and the POI [116].

Allosteric ligands of E3 ligases would confer unique opportunities in the targeted protein degradation field since they (i) do not compete with the natural substrates of E3 ligases, and hence do not require a potent affinity, (ii) are less likely to interfere with the natural function of the E3 ligase, (iii) can be as selective as the ligands that bind to the substrate binding site. The use of allosteric ligands can increase the number of E3 ligases used for targeted-protein degradation purposes pushing the development of tissue- and disease-specific PROTAC molecules.

Chapter 5

Results

5.1 Examine the thermodynamics and kinetics of the retinol binding to Cellular Retinol Binding proteins (CRBPs)

5.1.1 Background: Cellular retinol binding protein

Retinol (Vitamin A) is an important micronutrient that plays an essential role in physiological processes like vision, embryonic development, cell growth and differentiation. Since retinol has a poor aqueous solubility, it needs to be bound to carriers in body fluids and in cells for its transport and delivery to specific target tissues. The two most abundant intracellular retinol-binding proteins (CRBP; isoforms I and II) have distinct tissue distribution and binding affinity for retinol, reflecting the specialized adaptation of CRBP-I as retinol storage in the liver and the uptake of retinol from the intestinal lumen and release into the blood by CRBP-II in epithelial cells. The two proteins have a high structural identity: 56% residue identity and 70% residue homology and the structure fold consists of a β -barrel formed by two almost orthogonal five-stranded β -sheets (AE and FJ) and two short helices (I and II) inserted between β A and β B strands. The entry portal site is a crucial element formed by helices I and II and turns β C- β D and β E- β F that enables retinol to enter into the cavity (Fig. 5.1). Despite the high structural identity between CRBP-I and CRBP-II the retinol dissociation constant (K_d) for CRBP-I is smaller relative to CRBP-II, the ratio between binding affinities varying from 100-fold according to NMR measurements to 3.3-fold based on fluorometric assays [117]. It is unclear whether the affinity difference between the two isoforms stems from the few residue substitutions that line the binding pocket in the interior of the β -barrel or alternatively to differences in the dynamics of CRBP-I and CRBP-II, which might affect the entry/release of retinol to/from the binding cavity.

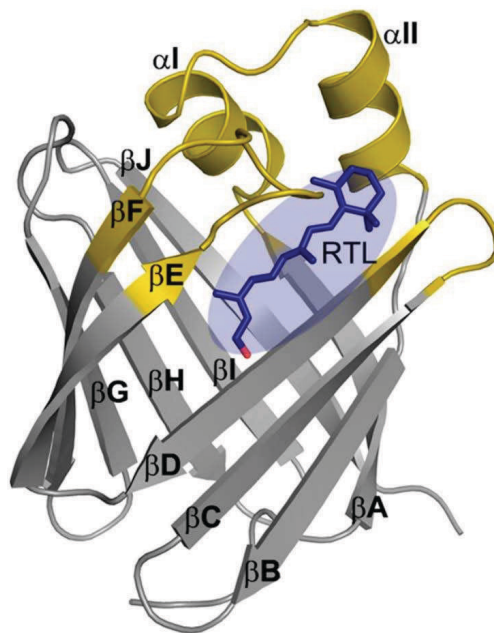


Figure 5.1. The structural fold of CRBP consists of 10 antiparallel β -strands (A-J) and two short helices (I and II). The regions that define the entry portal site are highlighted in yellow. Retinol (RTL) is shown as blue sticks. Adapted from [118].

5.2 Objective of the Chapter 5

To investigate the binding mechanisms in rat CRBP-I and-II and explore their functional implications, in collaboration with the group of Dr F. Javier Luque (University of Barcelona) a detailed analysis of apo- and holo- forms was performed by combining extended atomistic molecular dynamics (MD) simulations and parallel-tempering metadynamics (PT-metaD) (published article, see Annex 1). Overall, the results point out that the difference in the interaction energy with retinol can be mainly attributed to the conservative mutation of Ile78 in CRBP-I to Leu78 in CRBP-II.

In order to validate these calculations experimentally, SPR was used to obtain the kinetic and binding information of retinol for CRBP-II and its corresponding mutation Leu78→Ile, that recalls the structure and the main difference of CRBP-I, to confirm this unexpected effect related to the methyl isomerism between the side chains of these two residues.

5.3 Cloning, expression, purification and characterization of CRBPs

5.3.1 Site-directed mutagenesis

The Leu78→Ile single-mutated plasmid variant of CRBP-II was obtained performing a single site-directed mutagenesis as described in section 3.2.2.1. Primers were designed following the recommendations in the QuickChange Manual (Agilent). Both of the primers contained the desired mutation and anneal to the same sequence on opposite strands of the plasmid. In order to increase the efficiency of the mutagenesis reaction, primers of 39-mer were designed carrying the mutation site in the central region of the primer and presenting a GC content of 40% and ending in C or G bases to ensure a strong annealing. The leucine corresponding codon CTG, was mutated to the isoleucine codon ATT. Sequencing of the mutated plasmid is showed in Figure 5.2. The predicted sequence with the mutated codone (highlighted in blue) overlaps the sequence coming from the sequencing of the new plasmid.

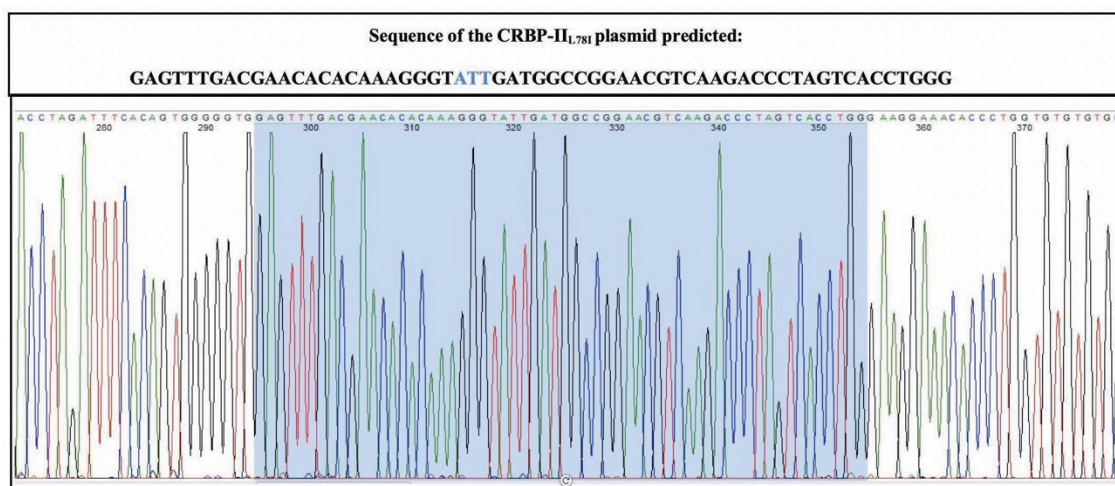


Figure 5.2. In the top the predicted sequence of CRBP-II_{L78I} with the codon in blue corresponding to the L78→I78 mutation. In the bottom part the results of plasmid sequencing after single point mutation with the matching sequence highlighted.

5.4 Purification and characterization of CRBPs

CRBP-I, CRBP-II and its Leu78→ Ile mutant were purified with the following protocol: cells were collected and subjected to lysis and the filtered solution was purified by a histidine affinity chromatography. The His-GST tag was then removed by an overnight TEV cleavage. The cleaved complex was then purified again with His affinity chromatography to separate the His-GST tag and protein with Ni³⁺ affinity still present in the solution. Final purification was performed with an anion exchange chromatography using an Hltrap QFF column. The purification yields to a final amount of 3 mg/l for CRBP-II, 0.2 mg/l for CRBP-II Leu78→ Ile single-mutated variant and 0.1 mg/l for CRBP-I. Purification steps were monitored by gel electrophoresis and protein purity was assessed by mass spectrometry (Fig 5.3)

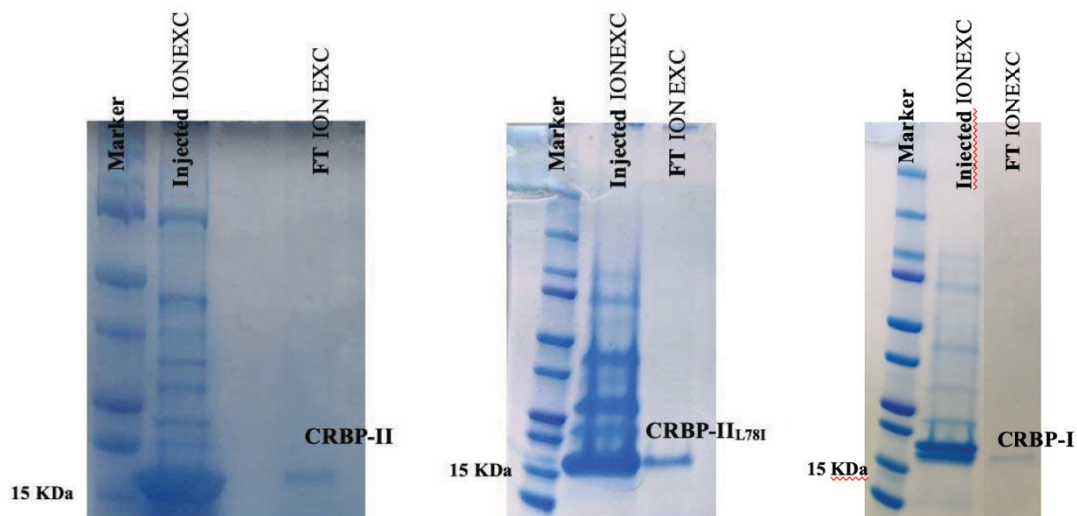


Figure 5.3. Purification of CRBP proteins. Electrophoresis gel showing the presence of the different purified CRBP proteins.

5.5 SPR: titration of retinol

The binding affinity and kinetic were analyzed using SPR. A CM5 chip was used to immobilize the protein. Retinol was dissolved as 50 mM stock solution in pure DMSO and diluted with 1.05x PBS 0.05% (v/v) tween-20 at

Results CRBPs

different concentration to get a dose response curve. The results depicted in Figure 5.4 show that the k_{on} remains essentially unaltered for both CRBP-II and its mutated variant. However, the k_{off} of retinol is slowed down by a factor of 2.2 in the mutated protein. The increased residence time originated from the single-point mutation Leu78→Ile agrees with the expected strengthening of the interaction of retinol with the mutated residue in CRBP-I (Ile) relative to CRBP-II (Leu), as deduced from the PT-metaD simulations and the decomposition analysis performed by the Javier Luque's group. Furthermore, the dissociation constant (K_d) decreased by 2.8-fold in the mutated CRBP- II, which compares with the lower limit of the experimental ratio between CRBP-I and CRPB-II (3.3-fold).

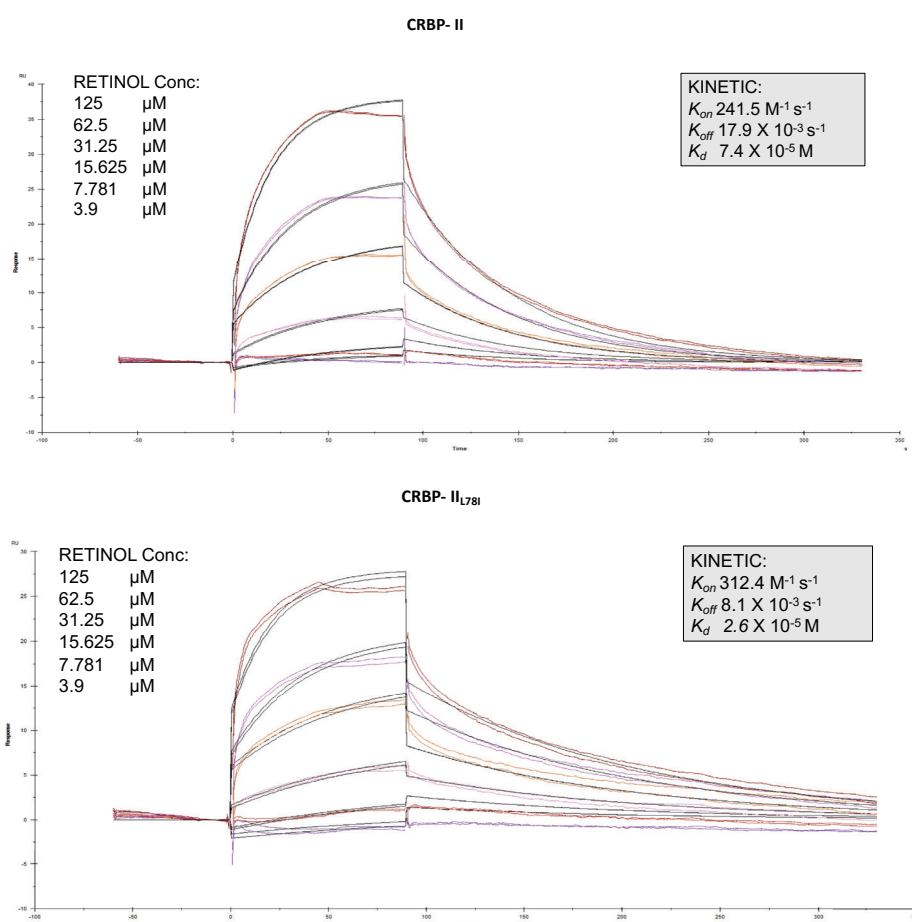


Figure 5.4. Top) Kinetic fitting of retinol binding to CRBP-II wild-type. Bottom) Kinetic fitting of retinol binding to CRBP-II L78→I.

In order to extrapolate the kinetic and affinity data for the CRBP-I isoforms, SPR experiment was also performed with these proteins. Not conclusive results could be obtained, probably due to a different immobilization of the protein in the surface chip as illustrated in Figure 5.5. Indeed, according to the amount of protein immobilized, the RU coming from retinol binding was expected to be in the same range of the one of CRBP-II and its mutant form (around 40 RU), instead we observed an 8-fold decrease in RU. Definitely, the shape of the sensogram made us envisage that the protein could not behave as well as the corresponding isoforms (blocked entry portal for retinol binding). Indeed, the retinol binds to a hydrophobic cavity of the protein that is shielded from solvent exposure. Once retinol is close to the protein surface, the portal needs to be opened to allow retinol binding.

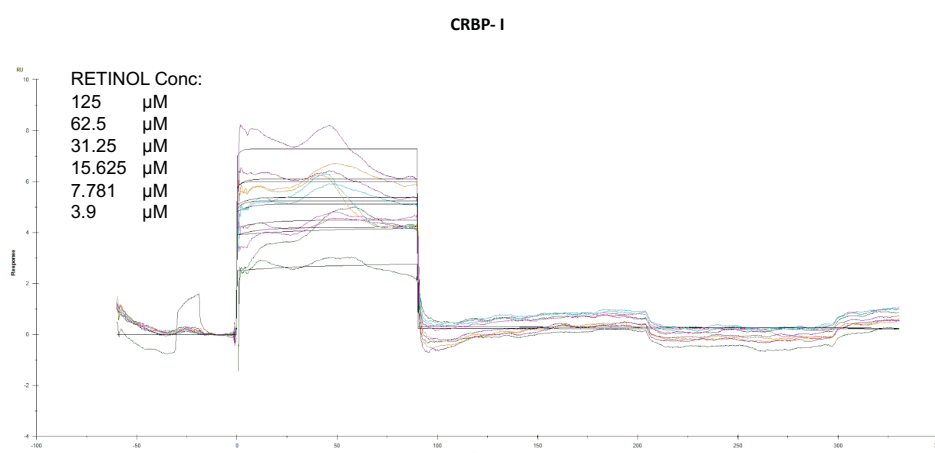


Figure 5.5. Sensogram of retinol binding to CRBP-I.

5.6 Summary

To confirm the impact of the Ile/Leu mutation at position 78 on the binding of retinol to CRBP-I and -II, CRBP-II was expressed and purified and its mutant form was obtained successfully following the procedure reported in the site mutagenesis kit. SPR was used to characterize the kinetic rate constants for the association (k_{on}) and dissociation (k_{off}) of retinol to CRBP-II and its Leu78 \rightarrow Ile single-mutated variant. As reported in the paper

(Annex 1), the results show that the k_{on} remains essentially unaltered for both CRBP-II and the mutated variant. However, the k_{off} of retinol is slowed down by a factor of 2.2 in the mutated protein. The increased residence time originating from the single-point mutation Leu78→Ile agrees with the expected strengthening of the interaction of retinol with the mutated residue in CRBP-I (Ile) relative to CRBP-II (Leu), as deduced from the PT-metaD simulations and the decomposition analysis. Furthermore, the dissociation constant (K_d) decreased by 2.8-fold in the mutated CRBP-II, which compares with the lower limit of the experimental ratio between CRBP-I and CRBP-II (3.3-fold). The net effect is the enhanced free energy penalty associated with the closed → open transition, which would disfavour the release of the ligand and increase the residence time of retinol in the interior of the β -barrel. Noteworthy, the affinity for the two isoforms is finely modulated by the differential interaction of the β -ionone unit of retinol with the residue (Ile/Leu) at the top of the loop EF, suggesting an unexpected role of the methyl isomerism between the two similar residues. Finally, these findings demonstrate that conservative changes in specific residues at remote sites distinct from the binding pocket, which should not alter the gross structural and physicochemical features of the protein, may result in fine-tuning of the ligand's binding properties.

Chapter 6

Results

6.1 Automated fragment evolution platform: Validating small-molecule binding to Brd4(BD1)

6.1.1 Background

There is an abundance of software that have been useful in FBDD to guide fragment optimization, for instance programs like PINGUI [57], ALTA [58] or DOTS [59] can be used for a fragment to lead optimization. However, *in silico* methods are most often used to rank optimized fragments, whereas the design strategy is guided by expertise and intuition of the medicinal chemists that ensure the synthesizability of the designs. This common practice may unconsciously lead to biased and scaffold centric chemical designs, thereby limiting the chemical search space. *In silico* strategies for chemical space exploration involve *de novo* drug design and virtual screening methods that address synthetic accessibility implicitly or explicitly.

Our group has developed an automatic protocol of iterative virtual screening using massive commercial libraries to evolve fragments to more potent drug-like compounds. Since the automatic protocol has been developed by two PhD students in the Barril Lab, Serena Piticchio and Dr Moira Rachman, the computational protocol will not thoroughly be described here. In abstract, given an initial fragment, which binding mode is known (e.g. by X-ray crystallography), the protocol searches virtual libraries for molecules that are chemically related and slightly bigger in size. These are then *tethered docked* to the target protein to identify those that are complementary. Dynamic undocking [106] is then applied to filter out false positives and top candidates are selected. The process is repeated until drug-sized molecules are attained. This protocol has been applied to several and different drug

discovery programs in-house or in collaboration, including the target of interest in this PhD thesis, the epigenetic reader Brd4(BD1).

In recent years, there has been an explosion of interest in targeting epigenetic proteins, exemplified by the developing of small molecules binding to bromodomains. Bromodomains are well known for their roles of readers of acetyl-lysine modifications and their functions have been largely reviewed [116, 117, 118, 119]. This post-translational modification (PMT) occurs in histones and regulates DNA replication. Thus blocking the acetyl-lysine binding site of bromodomains prevents recognition of acetylated histone tails and alters the process of chromatin remodelling [123], opening up the possibility to treat cancer, cardiovascular diseases and many other diseases [121, 124, 125]. The human genome encodes 46 of such bromodomain-containing proteins (BCPs). Each of the 46 proteins contains one to six bromodomains, giving a total number of 61 unique individual human bromodomain sequences. Over the past 5 years, many inhibitors that target the BET subfamily of bromodomains have emerged. The BET subfamily, which comprises four members in humans (Brd2, Brd3, Brd4 and BrdT), take their name from the presence of two related tandem bromodomains named BD1 and BD2, able to specifically recognize different acetylation patterns in H3 and H4 histone tails. Noteworthy, deregulation of Brd4 has been strongly linked to acute myeloid leukaemia [126], ovarian carcinoma [127], chronic obstructive pulmonary disease [128] and siRNA knock down of Brd4 induced upregulation of apolipoprotein A1 (ApoA1), which protects from atherosclerosis progression and other inflammatory processes [129]. Bromodomains bind to acetylated lysine residues through a network of interactions that involve the acetyl group, the conserved side chains of the Bromodomains and a set of structurally conserved and surprisingly inert water molecules [125] as illustrated in Figure 6.1.

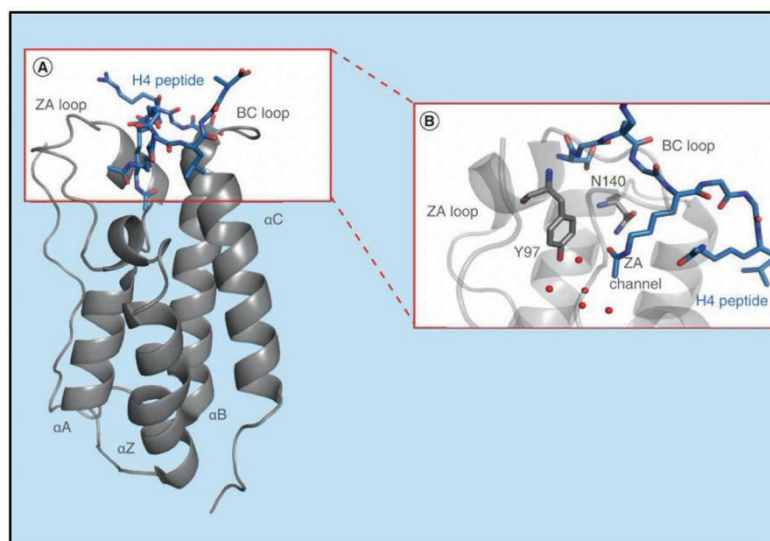


Figure 6.1. Structure and molecular recognition of BET bromodomains **A)** X-ray structure of the di-acetylated H4 peptide (double acetylation at H4K5acK8ac, in blue) bound to the BET bromodomain Brd4(BD1) (in grey, PDB 3UVW). **B)** Highlighted the conserved Y97, N140 and the ZA channel of BRD4-BD1(PDB 3UVW). Adapted from [87].

As mentioned above, two PhD students in our Lab, Dr Moira Rachman and Serena Piticchio developed an automated platform in order to generate new chemical structure from a known fragment binding to the protein of interest. The general aim of the project is to find novel and different scaffolds for Brd4(BD1) and validate the automate protocol using a known, reported fragment able to bind Brd4(BD1) as reference. Currently, there are a lot of examples in the literature of studies reporting bromodomains inhibitors containing several chemical scaffolds. The major part of these inhibitors disrupts the main key interaction of histone acetylated lysine with Asn140. One of this example includes fragment 1XA (PDB 4RL6 Fig. 6.2), showing an IC_{50} of 33 μ M [130]. Applying the automated fragment platform evolution, starting from fragment 1XA, the protocol searched in a database for molecules that are chemically related and slightly bigger in size. These molecules were then tethered docked to the target protein to identify those that were complementary. DUck was then applied to filter out false positive and top candidates were selected. The process was repeated until drug-size molecules were obtained and a total of 29 new compounds were generated.

6.2 Objective of the Chapter 6

From the fragments identified by the automated platform with the different iterations, 23 of these fragments were commercially available and six required in-house synthesis. Therefore, in the present thesis the objectives are to i) synthesize the non-commercial fragments and ii) set-up a TR-FRET experiment and assay all the optimised compounds, purchased and synthesised, for their ability to bind to the Brd4(BD1) protein.

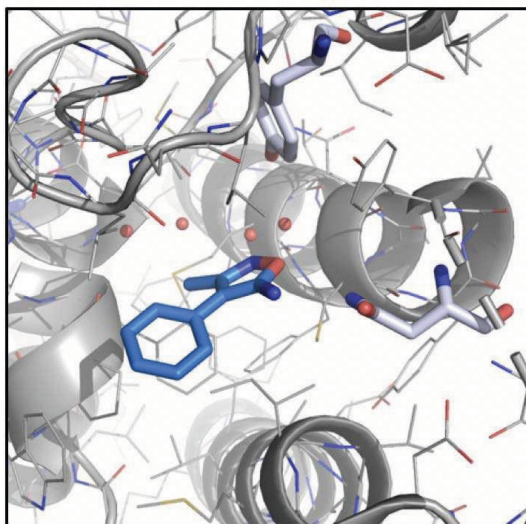


Figure 6.2. X-ray structure of 1XA in complex with Brd4(BD1) (PDB 4LR6).

6.3 Synthesis of SSR3, SSR4, SSR1 and SSR2

Synthetic procedure for compounds SSR3, SSR4, SSR1 and SSR2 is described in chapter 3 and general overview is given in Figure 6.3. Synthesis was obtained taking as reference the reported procedure in the literature [61]. Compounds were characterized as specified in section 3.12.

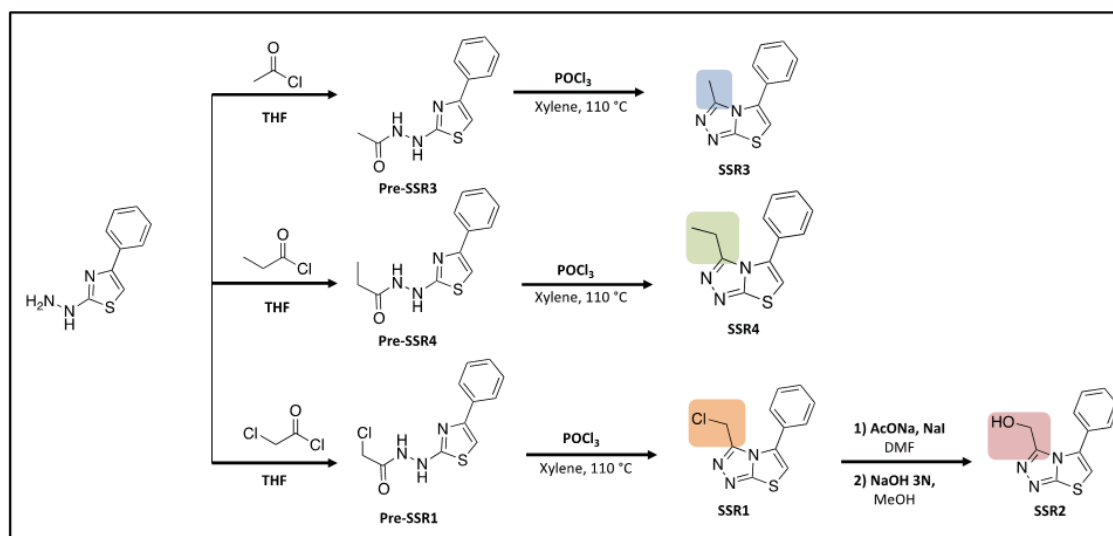


Figure 6.3: Synthetic route of SSR3, SSR4, SSR1 and SSR2.

In the case of Pre-SSR3 and Pre-SSR4 the synthesis was straightforward, the high nucleophilicity of the hydrazine group of the starting material and the high reactivity of the acetyl chloride and propionyl chloride allow us to achieve the desired pure compounds with excellent yields, 96% and 85 %, respectively. However, we encountered several problems in the synthesis of Pre-SSR1. We first decided to perform the reaction between the hydrazine and the 2-chloroacetyl bromide to ensure the nucleophilic attack of the starting material to the carbonyl group of the reagent, which bears a good leaving group (bromide). Surprisingly, we observed the formation of different chemical species coming from the addition of the hydrazine group either to the carbonyl group or to the alkyl chain (Fig. 6.4).

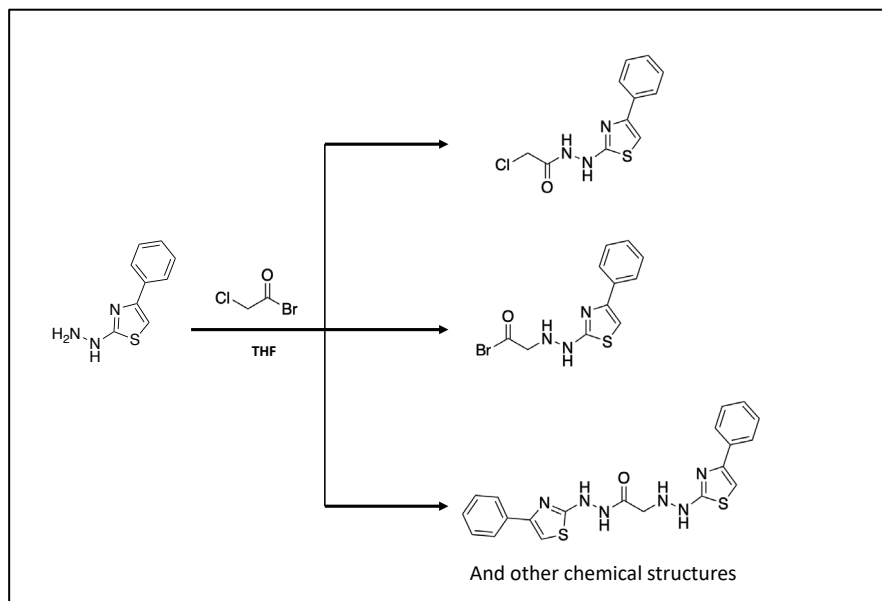


Figure 6.4. Possible side reactions during Pre-SSR1 synthesis.

The purification of the mixture obtained was tedious. For that, we then decided to change the 2-chloroacetyl bromide for a less reactive reagent, 2-chloroacetyl chloride. Amazingly the reaction allows us to obtain the desired compound, Pre-SS1, with an excellent yield (93% of yield) without further purification.

The next step was to obtain the final SSR1, SSR3 and SSR4 through a cyclization reaction. Using a protocol described in the literature [61] which employs exclusively POCl_3 and high temperature $110\text{ }^\circ\text{C}$, we attempted to force the cyclization process at several times (4, 8 and 12 h). Unfortunately, it was not possible to obtain the desired compound. Since the reaction was conducted without solvent, we then thought to select a solvent with a high boiling point to perform the reaction. Therefore, we decided to use either toluene (boiling point $110\text{ }^\circ\text{C}$) or xylene (boiling point $139\text{ }^\circ\text{C}$). Xylene resulted in the best choice to succeed in this reaction. Cyclization of Pre-SSR1, Pre-SSR3 and Pre-SSR4 gave SSR1, SSR3 and SSR4 with moderate to good yields (72% for SSR1, 48% for SSR3 and 45% for SSR4) after column chromatography purification. Noteworthy, in order to understand the low yields obtained for SSR3 and SSR4 compared to SSR1, we chose SSR3

as the compound to further investigated and to analyse its side products. According to NMR spectra and mass spectrometry we found as main product, the side compound that is showed in figure 6.5. The mechanism of the reaction remains unknown.

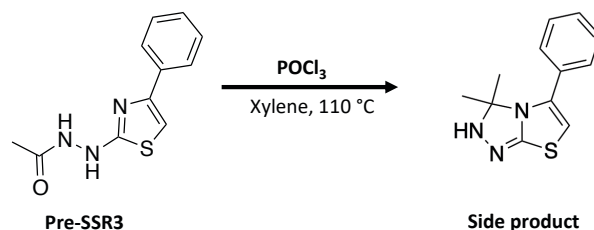


Figure 6.5. a) Side product of Pre-SSR3.

SSR1 was also used as starting point for the synthesis of SSR2. First with a mixture of SSR1, sodium acetate and sodium iodide in DMF at 100 °C for 2 h the corresponding acetoxy intermediate was obtained with 34% yield. The acetoxy group (good leaving group) in the presence of NaOH facilitated the formation of the corresponding desired alcohol, SSR2.

6.4 Cloning, expression and purification of Brd4(BD1) in *E. coli*

The cloned Histidine (His)-tagged BRD4(BD1) was expressed in Rosetta(DE3). The expression of the recombinant proteins in *E. coli* strain Rosetta(DE3) was started after induction with 0.4 mM IPTG as described in section 3.7.4. Brd4(BD1) was purified with the following protocol: cells were collected and subjected to lysis and the filtered solution was purified by a histidine affinity chromatography. Fraction 24 to 27 (Fig. 6.6) of elution were collected and subjected to buffer exchange (50 mM Hepes, pH= 7.4, 50 mM NaCl) to remove the high concentration of imidazole. The overall yield was of 4 mg/l of growth broth. Given the high amount of protein achieved with 1 L of growth broth, only 4 L were expressed. Furthermore, only one purification step was needed since to perform Tr-FRET experiment the

protein needed His-tag to interact with the labelled antibody anti-his.

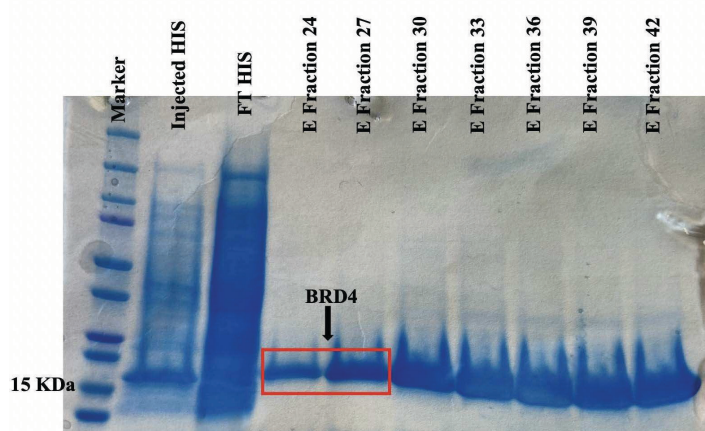


Figure 6.6. Brd4(BD1) his-tag purification.

6.5 Time-resolved Fluorescence Resonance Energy Transfer (TR-FRET) experiment

The experiment was set-up following a procedure described in the literature [131]. Several attempts were done in order to find the right conditions to perform the assay. One of the biggest problems encountered in the assay was the preparation of the buffer. It resulted really important how the different components are mixed together and at which time.

As a first attempt, Brd4(BD1) (100 nM) in 50 mM HEPES pH= 7.5, 50 mM NaCl buffer, was mixed with 200 nM of the tetra-acetylated histone H4 peptide biotinylated in the same buffer and the mixture was left equilibrating for 30 min. Subsequently, 3 nM of europium cryptate labelled streptavidin and 6.67 nM XL-665-labeled anti-6His antibody (Cisbio 61HisXLA) in 50 mM HEPES pH= 7.5, 50 mM NaCl buffer were added and incubated for 30 min. Finally, a solution of the compound in 250 mM HEPES pH= 7.5, 250 mM NaCl 20 % DMSO was added and the final volume of the assay (100 μ l) was achieved in the 96-wells plate with a buffer containing 4 M KF, 5 mM CHAPS and 1 % BSA. Several trials were done but not consistent results could be observed using a known

inhibitor ((+)-JQ1) as positive control.

To optimize the experiment, we change the 96-well plates used. We replaced the black NUNC plate (ref 10692202) by white GREINER plates (ref. 784904). In spite of the change of these plates, we could not obtain the desired result.

Since the preparation of the buffer with KF, CHAPS and BSA resulted often in the formation of a white colloidal solution hard to remove, we decided to change the preparation of the buffer. As a result of this change, Brd4 (BD1) was prepared in a buffer containing 50 mM HEPES pH= 7.5, 50 mM NaCl but with the addition of 2.5 mM CHAPS. The compounds to test were prepared adding 250 mM HEPES pH 7.5, 250 mM NaCl 20 % DMSO and 1 % BSA and, finally, 8 M KF was added alone in the last step of the assay set-up assay. These changes resulted in the right conditions to perform the assay. Once the experiment was set-up, a competition assay between Brd4(BD1) and a peptide of a natural substrate [Lys(Ac)5/8/12/16]-Histone H4 (1-21)-GGK(Biotin) treated with different compounds coming from the automated platform has been performed to get the IC₅₀. In table 6.1 are summarized TR-FRET results. Compounds were first assessed at a single concentration (50 μM) and measurements were done immediately after preparing the solutions and later on after equilibration of approximately 2 h. Potent compounds are highlighted in table 6.1 in orange and the starting fragment 1XA (of the automated platform) in red and were then screened at several concentration in order to obtain a dose-response curve. Each IC₅₀ value is presented with its own 95% Confident Interval (CI₉₅). Dose-response curves are depicted in Figure 6.7.

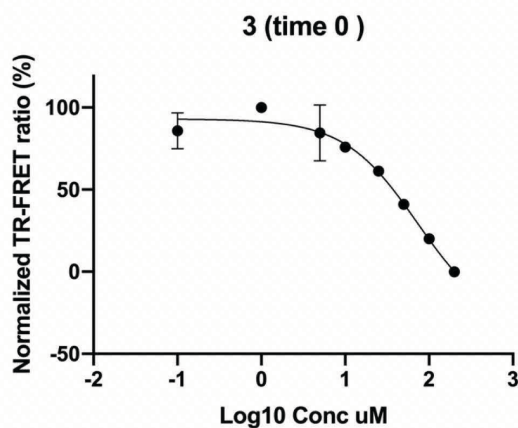
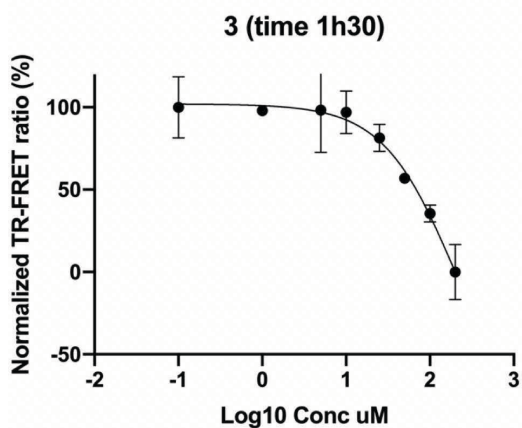
Results Brd4(BD1)

Table 6.1 TR-FRET results

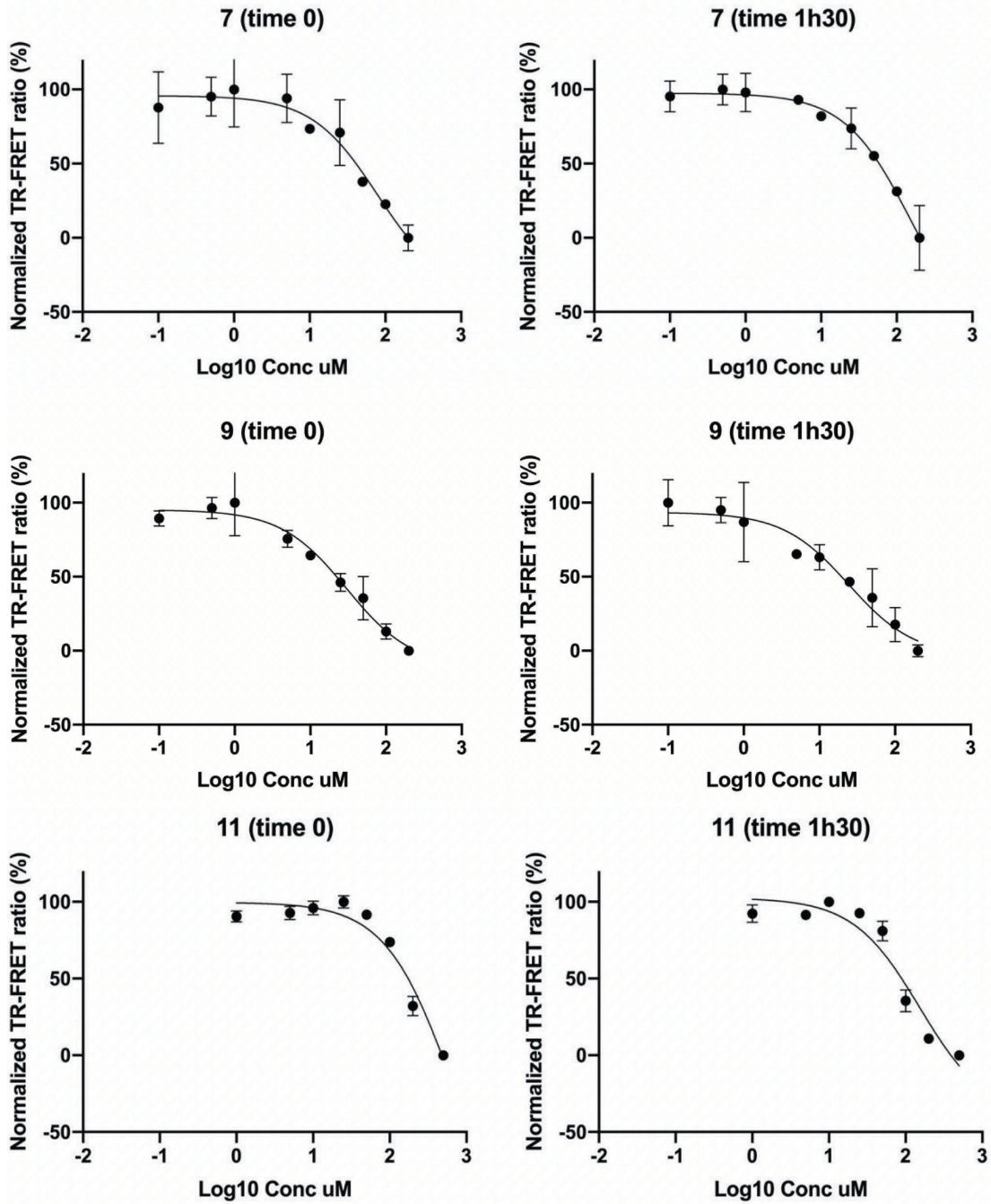
Compound	% inhibition at 50 μ M t=0 h	% inhibition at 50 μ M t=2 h	IC ₅₀ (μ M) t=0 h [CI ₉₅ (μ M)]	IC ₅₀ (μ M) t= 1h30' [CI ₉₅ (μ M)]
2	No inhibition	No inhibition	n.d	n.d
1	No inhibition	No inhibition	n.d	n.d
8	No inhibition	No inhibition	n.d	n.d
16	No inhibition	59.9	IC ₅₀ = 490.6 [CI ₉₅ = 198.8-3419]	IC ₅₀ = 1197 [CI ₉₅ = n.d]
22	No inhibition	No inhibition	n.d	n.d
6	No inhibition	No inhibition	IC ₅₀ = 191.4 [CI ₉₅ = 36.3-2990]	No fitting
17	No inhibition	No inhibition	n.d	n.d
14	No inhibition	26.8	IC ₅₀ = 191.9 [CI ₉₅ = 54.4-1420]	IC ₅₀ = 133.4 [CI ₉₅ = 41.7-601.9]
5	No inhibition	No inhibition	n.d	n.d
19	No inhibition	No inhibition	n.d	n.d
21	No inhibition	No inhibition	n.d	n.d
12	7.5	21.1	No fitting	No fitting
10	4.2	62.9	No fitting	No fitting
11	No inhibition	33.4	IC ₅₀ = 644.4 [CI ₉₅ = n.d]	IC ₅₀ = 152.1 [CI ₉₅ = 48.5-735.9]
20	No inhibition	No inhibition	n.d	n.d
7	44.9	77.5	No fitting	No fitting
4	No inhibition	8.5	No fitting	No fitting
15	No inhibition	No inhibition	n.d	n.d
13	No inhibition	17.3	n.d	n.d
18	18.3	44.1	No fitting	No fitting
SSR4	n.d	n.d	IC ₅₀ = 25.6 [CI ₉₅ = 12.1-55.8]	IC ₅₀ = 11.8 [CI ₉₅ = 5.9-23.9]

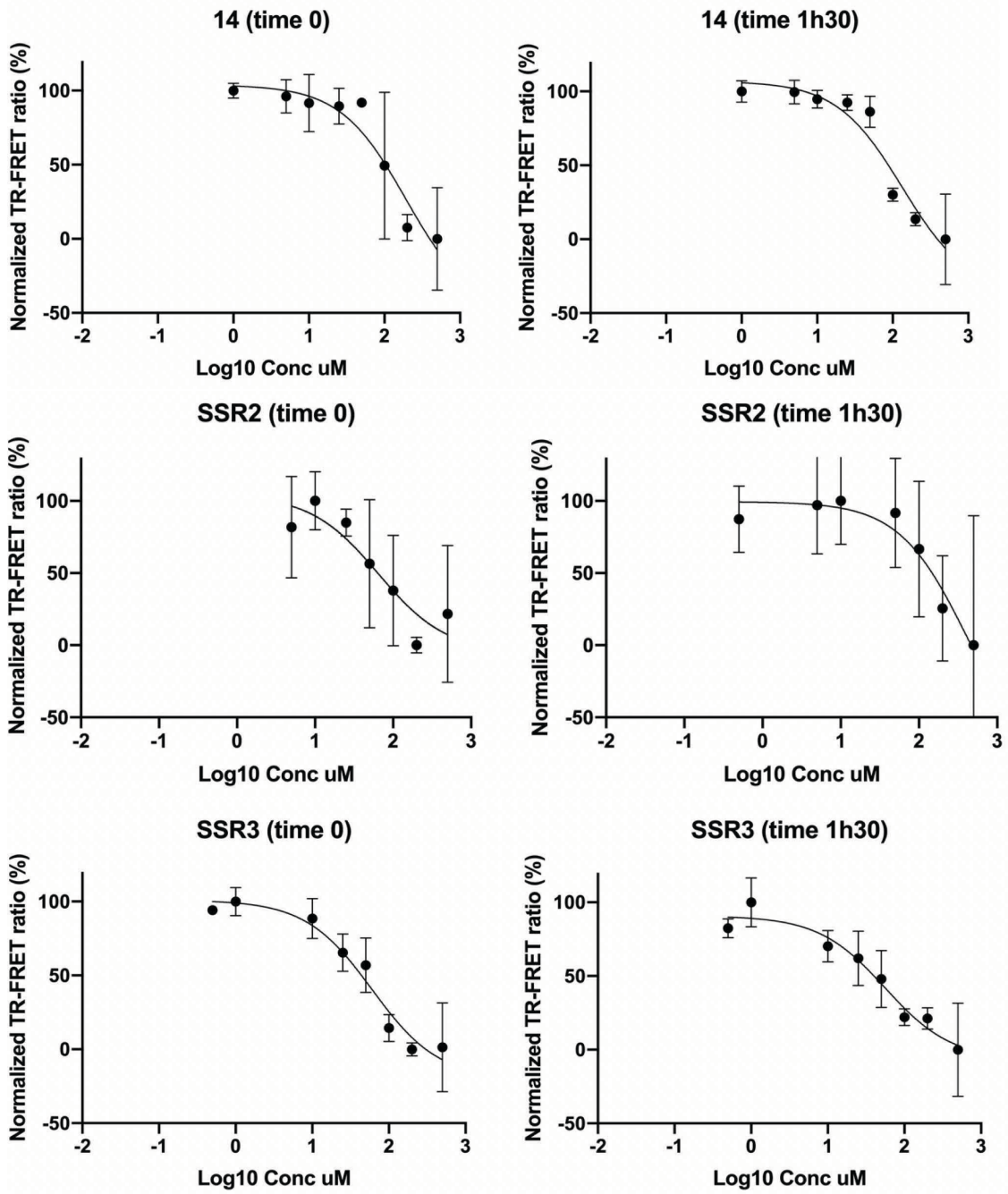
Results Brd4(BD1)

3	n.d	n.d	IC ₅₀ = 72.4 [CI ₉₅ = 35,7-68,2]	IC ₅₀ = 202.4 [CI ₉₅ = 111.7-475.3]
9	n.d	n.d	IC ₅₀ = 30.23 [CI ₉₅ = 15.1-63.8]	IC ₅₀ = 24.4 [CI ₉₅ = 8.3-76.4]
1XA	n.d	n.d	IC ₅₀ = 90.7 CI ₉₅ = n.d	n.d
SSR2	n.d	n.d	IC ₅₀ = 66.8 [CI ₉₅ = 4.8 to 1077]	IC ₅₀ = 471.9 [CI ₉₅ = n.d]
SSR3	n.d	n.d	IC ₅₀ = 60.3 [CI ₉₅ = 26.15-143.6]	IC ₅₀ = 55.9 [CI ₉₅ = 20.4 -163.1]
SSR6	n.d	n.d	IC ₅₀ = 156.6 [CI ₉₅ = 49.3-623.4]	IC ₅₀ = 421,5 [CI ₉₅ = 163.6-2586]
SSR7	n.d	n.d	IC ₅₀ = 84.5 [CI ₉₅ = 36.5-210.5]	IC ₅₀ = 116.9 [CI ₉₅ = 44.6-357]

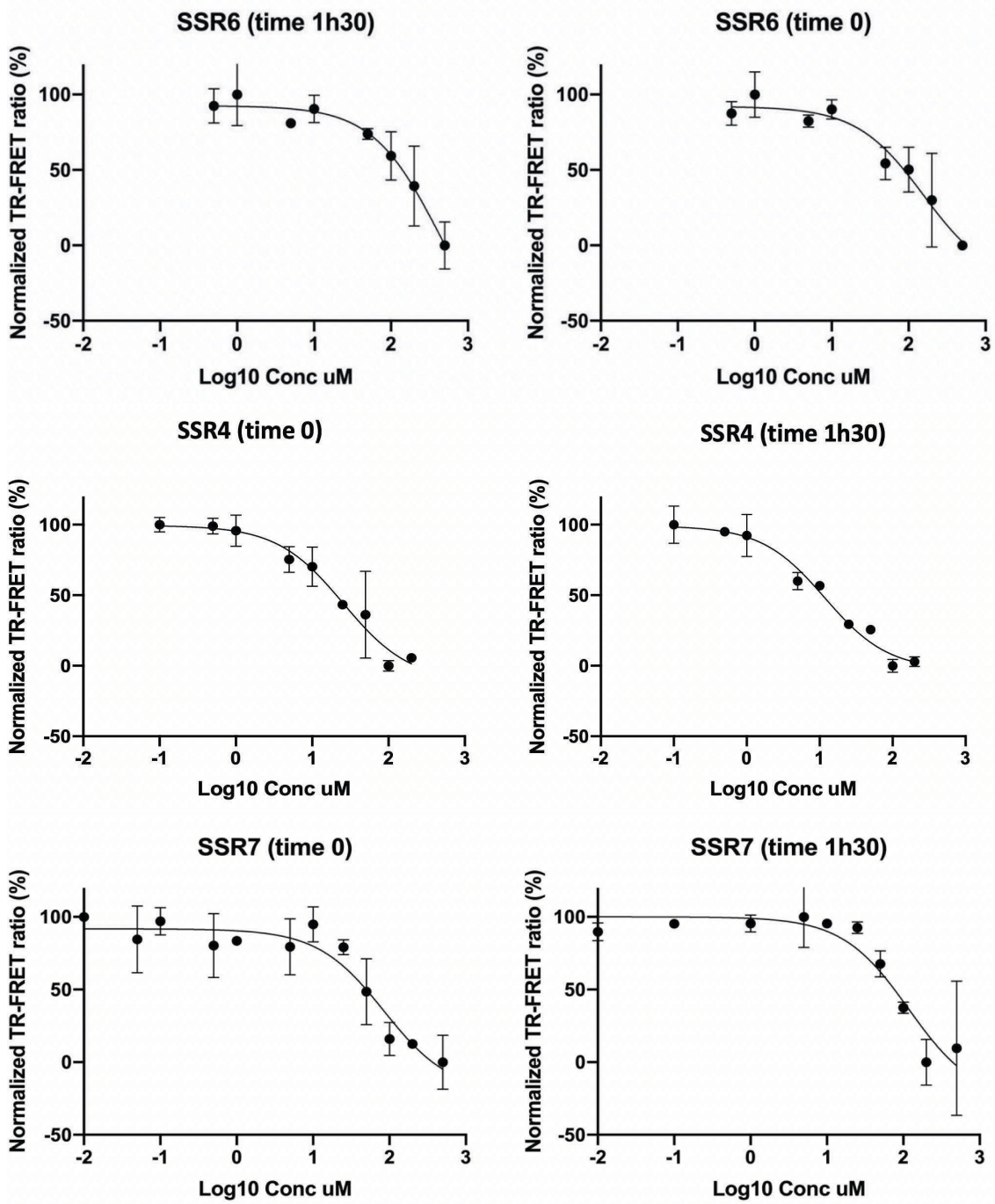


Results Brd4(BD1)





Results Brd4(BD1)



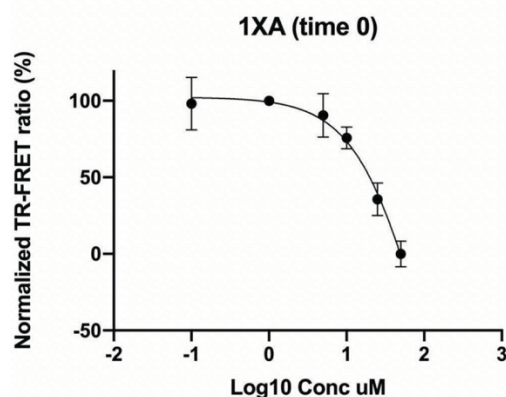


Figure 6.7. Normalized TR-FRET assay showing dose-dependently displacement of a tetra-acetylated H4 peptide from Brd4(BD1) in the presence of each compound at time 0 and after equilibration of 1 h 30 min.

6.6 Summary

As Dr Gisbert Schneider highlights “Automation of science bears the promise of making better decisions faster” [132] reflecting the importance of speeding up the drug discovery pipeline to solve more efficiently and rapidly unmet medical needs. Therefore, the automated platform represents a powerful tool able to give new chemical entities, in a fast manner, preserving the main interaction with the target, with more potency and with new chemical features that otherwise would not be considered. In order to address the reliability of the platform, compounds proposed were screened biophysically. Most of the compounds generated by the platform were commercially available, instead, few of them needed in-house synthesis. With the aim of getting the new chemical structure, SSR1, SSR2, SSR3, SSR4 were synthesized successfully in the lab. After expressing and purifying the protein (Brd4(BD1)) with its His-tag, a TR-FRET experiment, already well established for bromodomains protein, has been set-up in our lab. Out of 29, three compounds (SSR3, 9, SSR4) showed better potency compared to the starting fragment (1XA) capable of displacing the peptide at a concentration $< 50 \mu\text{M}$. Compounds 16, 6, 14, 11, 7, 3, SSR2, SSR6, SSR7

were able to compete with the peptide, but with a lower potency compared to the starting fragment (1XA). In order to know the binding mode of the new compounds, X-ray structures were solved for SSR4, SSR2 and SSR3. The different binding mode observed can somehow explain the differences observed of these compounds in disrupting the interaction between Brd4(BD1) and the peptide. As it appears clear, indeed, the 3 compounds preserve a similarity structure of >90% but SSR2 binds in a different manner. Indeed, while in SSR4 and SSR3 the alkyl group is pointing inside the binding site, in SSR2 the presence of an -OH in the alkyl chain rearrange the binding mode of the structure making the -OH solvent exposed and therefore flipping the structure in a different position, resulting in a loss of activity. Overall, these results showed that combining a computational and a biophysical approach enhance the possibility of finding ligands that bind to a specific protein like Brd4(BD1) or any other target and prompt the possibility of using the automated platform for other targets once known a specific ligand that need to be optimized.

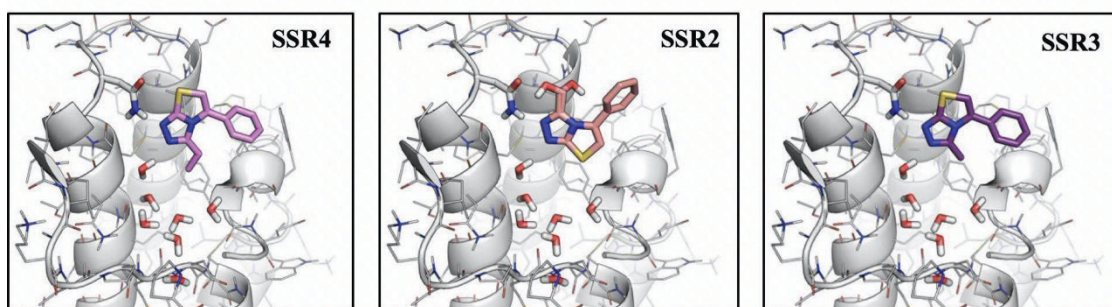


Figure 6.6: X-ray structures of SSR4, SSR2 and SSR3 in complex with Brd4(BD1).

Chapter 7

Discussions

7.1 General discussion

We are now facing a new era of incredible changes in science, that is evolving at breakneck speed. After the important and remarkable event 70 years ago of the discovery of the genetic code with the elucidation of the DNA helix made by Drs Watson and Crick; the understanding of protein structures and their roles in modulating cell functions, the elucidation of the human genome at the beginning of this century and more recently the new advances in gene therapy, with the first drug approved by EMA in 2012 (Glybera®), represent important and big steps in how we understand science and drug discovery. More recently, new genetic approaches i.e., gene editing with CRISPR-Cas9 system (2020 Chemistry Nobel Prize for Drs Jennifer Doudna and Emmanuelle Charpentier) has changed how we understand molecular biology, and more specifically, the target validation process in drug discovery. All breakthrough changes have seen as protagonists dedicate-passionate people that have spent their life and huge efforts in understanding the right combinations of pieces to complete a complex puzzle. Academic-basic science remains the right propeller to start projects that can become the "revolutionary science" to discover novel therapies that will allow in the future to treat difficult, until now incurable, diseases.

With the aim to contribute to the elucidation and validation of novel protein-ligand interactions and in the identification of molecules for until now considered undruggable proteins, in the present PhD thesis the use of different biophysical techniques has been discussed and applied for different therapeutic relevant proteins.

Biophysical techniques have seen an increasing role in recent years in drug discovery programs. In general, when using biophysical techniques in drug discovery, combination of these techniques has to be considered. Of course,

the more information you have, the most accurate and precise will be the outcome of the study, but it is always a good practice to select techniques that fit with your project, but most important that can be easily implemented in your Lab. In the following paragraphs, a deep discussion will be done for each project included in this thesis.

As underlined in chapter 4, FBDD represents a useful approach to assess the *ligandability* of a protein, since fragments are efficient probes to explore the chemical space at protein surfaces [109] and fragments hits can be an effective starting point for a drug discovery project [110]. Given that, we selected FBDD to start the arduous goal to find molecules able to bind to the E3 ligase Fbw7, considered undruggable until now. Using this approach, we could find fragments that show different MoA: i) allosteric inhibition, ii) competitive inhibition, iii) molecular glue effect, iv) or simply silent ligands, that can be developed in the future for PROTAC molecules construction.

After a laborious time spent to optimize the expression and purification protocol of Fbw7-Skp1 complex with the aim to increase the low yield that we were obtaining, we decided to further proceed to screen the fragment library. Aware of the limitation of the protein amount, we decided to apply DSF as the first primary screening, for the ease to implement the assay in the Lab. After testing different small-molecules coming from a different project led by Miriam Martinez and some of the fragments of the library and demotivated by the misleading results of the positive control (the peptide of the natural substrate cyclin-E), we decided to drop-out DSF and select another biophysical technique. We then decided to screen the fragment library using SPR. Indeed, SPR positions itself as the most versatile, low material consumption and exhaustive technique to use in a drug discovery pipeline. Its applicability depends on the purpose of the project and probably

the most important advantage of the technique is the possibility to obtain both thermodynamic and kinetic information. The screening of approximately 700 fragments led to 220 positives that were tested at different concentrations to extrapolate dose-response curves with the purpose to shrink the number of fragments identified. In total 48 fragments gave a dose response curve with a $K_d < 2\text{mM}$ and interestingly 21 fragments had a $K_d < 200\mu\text{M}$. Intrigued by the results obtained, we analysed 13 fragments out of 21 with STD-NMR, selecting the most potent and relevant structures for further optimizations. Nine fragments resulted positive binders. Excited by these results and to have more insights into the mechanism of action we set-up and performed FP assay to help to address this objective. Using FP, none of the fragments tested showed to be a modulator of PPI interaction between Fbw7-Skp1 and the natural substrates. This result made us think that our fragments bind allosterically to Fbw7. However, we have also to consider that some of the fragments could be too weak to compete with peptide binding. Noteworthy, the fragment could influence the ubiquitination pathway and therefore a ubiquitination assay can unveil this MoA. Work is on-going with this purpose. In parallel, to guide the project in a more specific direction, it is quite crucial to elucidate the binding site (and, if possible, the binding mode) of the fragments either by the resolution of a X-ray structure or by NMR.

Unlike other of protein classes, in which a ligand is active if acts as positive or negative modulator of the activity of the protein, E3 ligases open the ground to exploit the possibility to hijack the ubiquitin-proteasome system to destroy a POI. In this scenario, a fragment or an optimized derivate can be, therefore, integrated in a PROTAC molecule to selectively guide the degradation of a specific protein, either to kill its deleterious activity in the cells or to study and probe its function. A long journey has still to be travelled

but for sure great promises arise from the aforementioned work and the best has still to come....Fbw7-based PROTAC molecules are being developed in the Lab!

Millions of years of evolution have tuned the structure of proteins to elicit specific functions in the cell. As a result, proteins orchestrate the dynamics of cellular and extracellular activities that allow life. Each protein is characterized by a defined sequence of aminoacids that all together give shape to a structured entity able to exercise its role in the cell. The conformational stability of the protein is made possible by the harmonic “packaging” of each single aminoacid. Change in one single aminoacid can alter, in a deleterious way, the function of the protein. This is illustrated, for instance, in disease such as sickle cell anaemia in which β -hemoglobin function is disrupted by a single mutation from a valine to a glutamic acid or with the gain of function of RAF protein in cancer, in which the cause must be linked again to the single mutation from a valine to a glutamic acid. Proteins can also present a very high structural identity such as the case of CRBP-I and CRBP-II (56% residue identity and 70% residue homology) but address different specialized physiological functions.

In order to investigate the binding mechanisms in CRBP-I and CRBP-II and to explore their functional implications, a detailed analysis of apo- and holo-forms was performed in collaboration with the Lab of Dr F.J. Luque, combining extended atomistic molecular dynamics (MD) simulations and parallel-tempering metadynamics (PT-metaD). These studies revealed that the difference has to be linked with the Ile/Leu mutation at position 78 on the binding of retinol to rat CRBP-I and -II. As described in Chapter 5 of this thesis, in order to verify what has been found computationally, rat CRBP-I, rat CRBP-II and its Leu78 \rightarrow Ile single-mutated variant were expressed and

SPR was used to characterize the kinetic rate constants for the association (k_{on}) and dissociation (k_{off}) of retinol. Unfortunately, it was not possible to compare directly CRBP-I and CRBP-II since no retinol binding was observed in the case of CRBP-I, probably due to the restraint generated from the covalent linking of the protein to the sensor chip, that may block the entry portal of retinol. To overcome this problem, since in the case of rat CRBP-II it was possible to extrapolate retinol K_d , we decided to express and assess with SPR its Leu78 \rightarrow Ile single-mutated variant that recalls the difference between CRBP-I and CRBP-II in the entry portal of retinol. The results show that the k_{on} remains essentially unaltered for both CRBP-II and the mutated variant. However, the k_{off} of retinol is slowed down by a factor of 2.2 in the mutated protein. The increased residence time originated from the single-point mutation Leu78 \rightarrow Ile agreed with the expected strengthening of the interaction of retinol with the mutated residue in CRBP-I (Ile) relative to CRBP-II (Leu), as deduced from the PT-metaD simulations and the decomposition analysis presented in the paper (Annex). Furthermore, the dissociation constant (K_d) decreased by 2.8-fold in the mutated CRBP-II, which compares with the lower limit of the experimental ratio between CRBP-I and CRPB-II (3.3-fold). To conclude, SPR helped significantly to understand that the main difference between the isoform is related to a methyl isomerism between the two similar residues Leu78 \rightarrow Ile found in the entry portal of the two isoforms.

Once a molecule is found to bind to a protein pocket, a common practice is to optimize the structure, improve its binding strength and its physicochemical properties. In this process medicinal chemists are chased to make the perfect magic modifications that can drastically increase the potency of the molecule, resulting in a lead molecule to bring to the next steps. Often the optimization can be complicated, with unproductive

synthetic pathways. In this scenario, our group decided to implement an automatic fragment evolution platform, not only able to simplify the medicinal chemists work on finding the best substitution to make to increase potency, but also to find different molecules that preserve the main interaction of the starting molecule but with novel scaffolds, that otherwise would not be considered.

In order to address the reliability of the platform, we used a benchmark protein (Brd4(BD1)). A TR-FRET experiment was set-up and affinity assays were carried out. Most of the compounds proposed by the platform were commercially available, instead, few of them needed in-house synthesis. With the aim of getting the new chemical structure, SSR1, SSR2, SSR3, SSR4 were synthesized successfully in the lab by myself.

Out of 29, three compounds showed better potency compared to the starting fragment. Other nine compounds were able to compete with the fluorophore peptide but with a lower potency compared with the starting fragment. X-ray structures of 3 compounds with the same scaffold were obtained, which allowed us to elucidate the difference in potency seen in the TR-FRET assay. Overall these results showed that combining computational and biophysical approaches enhance the possibility of finding ligands that bind to a specific protein, like Brd4(BD1), and highlight the possibility of applying the automated platform for other proteins of interest to get new and diverse chemical structures.

Chapter 8

Conclusions

Conclusions

8.1 Fbw7-Skp1 project (Chapter 4)

- We were able to express and purify the Fbw7-Skp1 multisubunit-protein in low yields.
- After several attempts and conditions assayed, we concluded that DSF was not a reliable screening approach for the Fbw7-Skp1 system.
- Using SPR screening (protein immobilized covalently), out of 700 fragments that contains our in-house fragment library, 22 fragments gave a dose response curve with a $K_d < 200 \mu\text{M}$ by its ability to bind to the Fbw7-Skp1 protein.
- To confirm the positive fragments from the primary SPR screening, a different immobilization of SPR was performed, and positive fragments assayed.
- Additionally, nine of the positive fragments were confirmed by STD-NMR.
- In order to obtain structural information of the binding sites of these fragments able to bind Fbw7, a plethora of approaches are being assayed.
- Until now, no X-ray structure of the fragments bound to Fbw7-Skp1 has been obtained.
- MD simulations gave us some hints about where the fragments bind.
- A challenging 2D-NMR has been started. Crucially for this experiment, we were able to express and purify the labelled ^{15}N Fbw7-Skp1.
- We set-up a FP assay in order to disentangle if the identified fragments disrupt the PPI between Fbw7 and its substrates. We have not been able to observe any disruption of the natural PPI.

- In summary, the combination of several biophysical techniques aided to find fragments that bind a target that have been considered undruggable until now, the E3 ligase Fbw7-Skp1.
- These fragments could be employed to elucidate the best pharmacological strategy to target the Fbw7 E3 ligase.
- These fragments could also be employed as a head-ligands of Fbw7-based PROTACs.

8.2 CRBPs project (Chapter 5)

- We were able to express and purify CRBP-I and CRBP-II proteins.
- We were also able to perform site directed mutagenesis to obtain CRBP-II Leu78→Ile variant plasmid and express and purify the mutant protein.
- Using SPR screening (protein immobilized covalently), we were able to characterize the kinetic rate constants for the association (k_{on}) and dissociation (k_{off}) of retinol to CRBP-II and its Leu78 → Ile single-mutated variant.
- SPR attempts were also done to obtain kinetic information of CRBP-I, but unfortunately, we have not been able to observe any relevant outcome.
- SPR result suggests an unexpected role of the methyl isomerism Leu78 →Ile between CRBP proteins.

8.3 Brd4(BD1) project (Chapter 6)

- We were able to express and purify His-Brd4(BD1) protein.
- We successfully synthesized and characterized SSR1, SSR2, SSR3 and SSR4, new chemical structures coming from the automated fragment platform evolution.

Conclusions

- We set-up a TR-FRET experiment and we tested 29 compounds coming from the automated platform.
- Three compounds showed better inhibitory potency against Brd4(BD1) than the starting fragment, 1XA.
- Nine compounds were able to compete with the peptide but with lower potency than the starting fragment (1XA)
- The results of Chapter 6 demonstrated the capacity of the automated platform to suggest potent and diverse molecules from a given fragment.

Conclusions

Annex

Modulating Ligand Dissociation through Methyl Isomerism in Accessory Sites: Binding of Retinol to Cellular Carriers

Carolina Estarellas^{1,2}, Salvatore Scaffidi³, Giorgio Saladino², Francesca Spyraakis⁴, Lorella Franzoni⁵, Carles Galdeano³, Axel Bidon-Chanal¹, Francesco Luigi Gervasio², and F. Javier Luque¹

Journal of Physical Chemistry Letters **2019**, *10*, 7333-7339

1. Department of Nutrition, Food Science and Gastronomy, Faculty of Pharmacy and Food Sciences, Institute of Biomedicine (IBUB) and Institute of Theoretical and Computational Chemistry (IQTCUB), University of Barcelona, 08921 Santa Coloma de Gramanet, Spain.
2. Department of Chemistry, University College London, London WC1E 6BT, United Kingdom.
3. Department of Pharmacy, Pharmaceutical Technology, and Physical Chemistry, Faculty of Pharmacy and Food Science, and Institute of Biomedicine (IBUB), University of Barcelona, 08028 Barcelona, Spain.
4. Department of Drug Science and Technology, University of Turin, 10125 Turin, Italy.
5. Department of Biomedical, Biotechnological and Translational Sciences, University of Parma, 43121 Parma, Italy

Modulating Ligand Dissociation through Methyl Isomerism in Accessory Sites: Binding of Retinol to Cellular Carriers

Carolina Estarellas,^{†,§} Salvatore Scaffidi,[‡] Giorgio Saladino,[§] Francesca Spyrikis,[⊥] Lorella Franzoni,[¶] Carles Galdeano,[‡] Axel Bidon-Chanal,^{*,†} Francesco Luigi Gervasio,^{*,§} and F. Javier Luque^{*,†}

[†]Department of Nutrition, Food Science and Gastronomy, Faculty of Pharmacy and Food Sciences, Institute of Biomedicine (IBUB) and Institute of Theoretical and Computational Chemistry (IQTUCB), University of Barcelona, 08921 Santa Coloma de Gramenet, Spain

[§]Department of Chemistry, University College London, London WC1E 6BT, United Kingdom

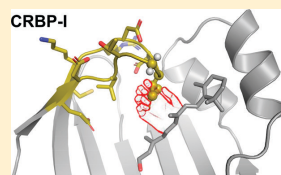
[‡]Department of Pharmacy, Pharmaceutical Technology, and Physical Chemistry, Faculty of Pharmacy and Food Science, and Institute of Biomedicine (IBUB), University of Barcelona, 08028 Barcelona, Spain

[⊥]Department of Drug Science and Technology, University of Turin, 10125 Turin, Italy

[¶]Department of Biomedical, Biotechnological and Translational Sciences, University of Parma, 43121 Parma, Italy

Supporting Information

ABSTRACT: Due to the poor aqueous solubility of retinoids, evolution has tuned their binding to cellular proteins to address specialized physiological roles by modulating uptake, storage, and delivery to specific targets. With the aim to disentangle the structure–function relationships in these proteins and disclose clues for engineering selective carriers, the binding mechanism of the two most abundant retinol-binding isoforms was explored by using enhanced sampling molecular dynamics simulations and surface plasmon resonance. The distinctive dynamics of the entry portal site in the *holo* species was crucial to modulate retinol dissociation. Remarkably, this process is controlled to a large extent by the replacement of Ile by Leu in the two isoforms, thus suggesting that fine control of ligand release can be achieved through a rigorous selection of conservative mutations in accessory sites.



Retinol is essential for many physiological processes like cell growth and differentiation, morphogenesis, and vision.¹ However, the poor aqueous solubility makes the assistance of plasma and cellular binding proteins necessary for the delivery to target tissues and the uptake and transport to specific partners in the cell.^{2–4} In fact, the efficient transport of hydrophobic molecules has been solved by evolution through selection of specialized binding proteins, such as the calycin and SEC14-like superfamilies.^{5–7} Although amino acid homology between the members of this widely distributed protein family is typically low, they share a similar β -barrel fold.^{8,9} Many of these proteins contain only this structural domain and can presumably be involved in transport of hydrophobic compounds, while others may have other domains, reflecting the involvement in a variety of cellular functions, such as signal transduction and regulatory roles. Nevertheless, precise knowledge of the mechanisms of recognition and binding is required to understand the roles in the cell, as illustrated by the ligand exchange mechanism that couples transfer of α -tocopherol and phosphatidylinositol phosphate lipids between the endosome and plasma membranes.^{10,11}

The two most abundant intracellular retinol-binding proteins (CRBP; isoforms I and II) have distinct tissue distribution and binding affinity for retinol, reflecting the

specialized adaptation of CRBP-I as retinol storage in the liver and the uptake of retinol from the intestinal lumen and release into the blood by CRBP-II in epithelial cells.⁴ The impact of residue substitutions selected by evolution in tuning the thermodynamics and kinetics of retinol binding to these isoforms is a conundrum. Hence, understanding the ability of CRBPs to sequester and protect retinol from the cellular milieu and to direct it to dedicated targets is essential for furthering metabolic engineering through selective nanocarriers and for drug discovery in retinoid-related diseases.^{12–14}

The structural fold of CRBP-I and -II consists of a β -barrel formed by two almost orthogonal five-stranded β -sheets (A–E and F–J) and two short helices (α I and α II) inserted between β A and β B strands (Figure 1).¹⁵ The entry portal site is a crucial element formed by helices α I and α II and turns β C– β D and β E– β F that enables retinol to enter into the cavity. Both NMR^{15–17} and X-ray¹⁸ data show that the binding mode of retinol is highly similar not only in human CRBP-I and -II but also in rat CRBPs.^{19–22} Despite the high structural identity between rat CRBP-I and -II (56% residue identity and 70% residue homology), the retinol dissociation constant (K_D) for

Received: September 27, 2019

Accepted: November 12, 2019

Published: November 12, 2019

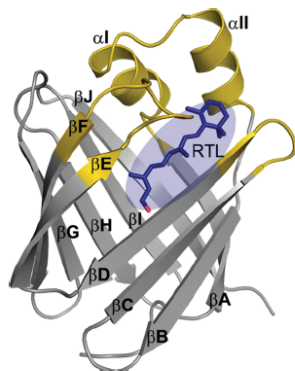


Figure 1. Representation of the retinol-CRBP complex. The structural fold of CRBP consists of 10 antiparallel β -strands (A–J) and two short α helices (I and II). The regions that define the entry portal site are highlighted in yellow. Retinol (RTL) is shown as blue sticks.

CRBP-I is smaller relative to CRBP-II, the ratio between binding affinities varying from ~ 100 -fold¹² according to NMR measurements to 3.3-fold based on fluorimetric assays.²³ At present, it is unclear whether the affinity difference between the two isoforms stems from the few residue substitutions that line the binding pocket in the interior of the β -barrel (SI Figure S1) or alternatively to differences in the dynamics of CRBP-I and -II,¹⁵ which might affect the entry/release of retinol to/from the binding cavity.

To investigate the binding mechanisms in rat CRBP-I and -II and explore their functional implications, a detailed analysis of *apo* and *holo* forms was performed by combining extended atomistic molecular dynamics (MD) simulations and parallel-tempering metadynamics (PT-metaD). We characterized the conformational flexibility of the two isoforms as well as the free energy surfaces for the opening/closing of the portal site in both *apo* and *holo* forms and the formation/breaking of interactions between retinol and protein in the *holo* species. Furthermore, surface plasmon resonance (SPR) was used to examine the thermodynamics and kinetics of retinol binding. Overall, both theoretical and experimental results provide detailed insight into the binding mechanism, disclosing a linkage between retinol binding and the flexibility of the entry portal, particularly regarding the methyl isomerism between Ile and Leu in this accessory site of the two isoforms.

The conformational flexibility of *apo* and *holo* forms of CRBP-I and -II was examined from three independent MD simulations (5, 3, and 3 μ s) performed for each system, covering a total of 44 μ s. The root-mean-square deviation profiles supported the structural stability of the simulated systems along the trajectories (SI Figure S2). In both isoforms, the presence of retinol reduced the structural fluctuations of the protein, as expected from the interactions formed with residues in the binding cavity (Figure 2; see also SI Figures S3–S6). However, the pattern of residue fluctuations differed in the two isoforms. While *apo*-I showed increased fluctuations in loops β E– β F, β G– β H, and to a lesser extent β C– β D, *apo*-II exhibited larger fluctuations in helices α I– α II and to a lesser extent in loops β C– β D and β D– β E (Figure 2A). Remarkably, most of these elements define the entry portal site, suggesting that the two isoforms differ in the dynamics of key structural

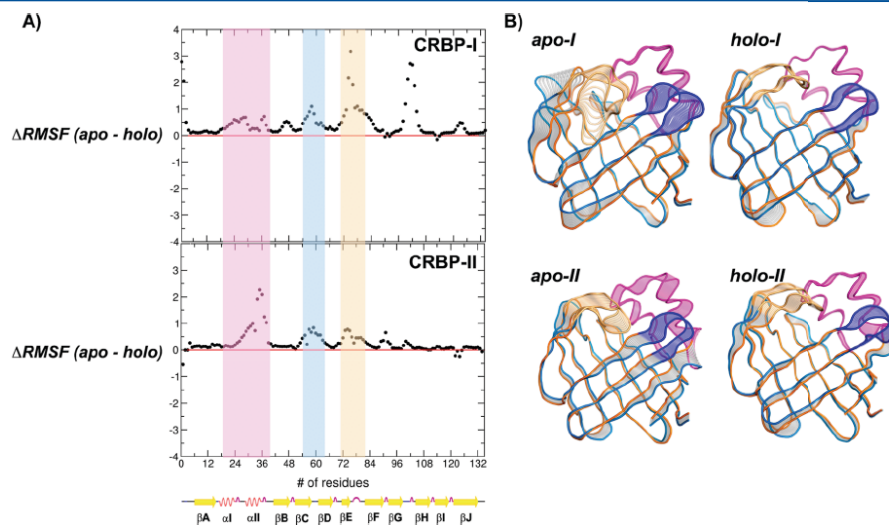


Figure 2. (A) Difference in residue fluctuations (Δ RMSF) observed along 5 μ s MD trajectories between *apo* and *holo* states for (top) CRBP-I and (bottom) CRBP-II. Highlighted regions correspond to structural elements of the entry portal site: helices α I and α II (magenta), loop β C– β D (blue), and loop β E– β F (orange). (B) ED analysis of (top) CRBP-I and (bottom) CRBP-II derived from the 5 μ s MD simulation. Only the first projection of the whole system is shown for *apo* and *holo* forms.

elements implicated in the entry/release of retinol to/from the binding cavity.^{24,25}

Essential dynamics (ED) analysis was employed to gain insight into the distinct flexibility of CRBP-I and CRBP-II. The analysis was performed for the backbone atoms of residues 7–134 to avoid the noise due to the mobile parts at the N- and C-termini. The first essential mode (Figure 2B) accounted for 15–25% of the entire structural variance and generally was 2-fold larger than the contribution explained by the second mode. The *apo* systems exhibited larger structural deformations, especially in the entry portal site, although helices *αI* and *αII* were stiffer in *apo-I* than in *apo-II*. On the other hand, *holo* systems were more rigid than their *apo* forms, as noted in the lower extent of the backbone motions. However, the decrease in conformational flexibility of the protein backbone did not affect similarly the two isoforms. In fact, the rigidification of the entry portal site was more important in *holo-I* than that in *holo-II* (SI Table S1). At first sight, these results seemed to be in contrast with NMR H/F exchange experiments¹⁵ that suggested a larger flexibility in both *apo* and *holo* states of CRBP-II relative to CRBP-I. However, it is worth noting that residues in the $\beta E-\beta F$ loop of *apo-I* could not be assigned, while present results reveal that this structural element has a crucial influence on the dynamics of the portal site. Indeed, upon exclusion of the $\beta E-\beta F$ loop in ED analyses, CRBP-II was slightly more flexible than CRBP-I in both *apo* and *holo* states (SI Table S1 and Figure S7), thus reconciling the experimental findings about the dynamics of the two isoforms and our results from MD simulations.

To estimate the differences in the dynamics of *apo* and *holo* systems, the conformational entropy was evaluated for the whole system as well as separately for the entry portal site and the protein core, formed mainly by the β -barrel, using the procedure by Harris et al. (see the SI for details).²⁶ As expected, the results (Table 1; see also Table S2 and Figure S8) confirmed that *holo* systems were less flexible than *apo* ones and pointed out that the decrease in entropy was larger for CRBP-I (0.56 kcal mol⁻¹ K⁻¹) than that for CRBP-II (0.14 kcal mol⁻¹ K⁻¹). Furthermore, the conformational entropy (S^{con}) obtained for *apo-I* was larger (by 0.35 kcal mol⁻¹ K⁻¹) than that for *apo-II*, whereas the difference between the *holo*

species was reduced to 0.07 kcal mol⁻¹ K⁻¹. Noteworthy, the entropy difference between *apo-I* and *apo-II* and between *holo-I* and *holo-II* was mainly due to the differences in the entry portal site (*apo*: 0.27 kcal mol⁻¹ K⁻¹; *holo*: -0.06 kcal mol⁻¹ K⁻¹). Overall, these results confirm that the distinct patterns of conformational flexibility between the two isoforms primarily arise from the entry portal site.

Because the stiffness of the portal site in *holo-I* was higher than that in *holo-II*, we hypothesized that the difference in binding affinity between CRBP-I and -II might arise from a larger residence time of retinol in the former isoform. To address this question, PT-metaD was used to evaluate the free energy landscape for the opening/closing of the entry portal site in *apo* and *holo* systems and the binding/unbinding of retinol to/from the *holo* systems. In order to take into account the larger structural flexibility of the open state compared to the closed one, the free energy change for the opening/closing of the entry portal site was estimated by averaging the values determined from three separate calculations, each relying on the use of distinct open structures chosen as reference systems (SI Figures S9–S11). The results pointed out that the opening of the portal site in the *apo* state of CRBP-I and -II was very similar and close to 5.5 ± 0.2 kcal mol⁻¹ (Figure 3A). However, the presence of retinol in the β -barrel had a marked influence on the opening of the portal site in *holo-I* as this process was disfavored by 3.8 kcal mol⁻¹ compared to *apo-I*. Remarkably, the presence of retinol led to a modest increase in the cost of opening the portal site in *holo-II* (by only 0.7 kcal mol⁻¹) relative to *apo-II*. These findings agree with the larger decrease in conformational entropy found for CRBP-I relative to CRBP-II upon retinol binding (see above and Table 1).

The analysis of the structures sampled during the opening of the portal site reveals that there is a slight rearrangement of retinol in the binding pocket, although the ligand remains trapped in the interior of the β -barrel after opening of the loop in both CRBP-I and -II (Figure 4). However, whereas the rearrangement of retinol occurs in a fast process during the first 100 ns of the loop opening for CRBP-II, a slower process that involves a gradual rearrangement of retinol is observed for the loop opening in CRBP-I. This suggests the presence of stronger interactions between the ligand with the residues of the portal site in this latter isoform, as will be discussed later.

PT-MetaD simulations were also used to estimate the free energy for retinol binding/unbinding from the open state of the two isoforms. Calculations were started from suitably chosen *holo* structures characterized by the presence of a portal site open enough to enable the release of retinol from the protein cavity without steric clashes. Similar events were involved in the release of retinol from CRBP-I and -II (SI Figure S12). Briefly, retinol unbinding involves first the breakage of the hydrogen bond formed by the terminal hydroxyl group and subsequent release of the polyene tail, which becomes progressively more exposed to water molecules, followed by the breakage of the van der Waals interactions with the β -ionone ring, leading to immersion in the aqueous environment. The free energy surfaces determined for CRBP-I and -II (Figure 3B) pointed out that the energetic cost for retinol unbinding is only 1 kcal mol⁻¹ higher in *holo-II*.

Overall, the combination of the free energy estimates obtained for the opening/closing in *apo* and *holo* states and the binding/unbinding of retinol from *holo* species indicates that the affinity of CRBP-I for retinol is 2.4 kcal mol⁻¹ more favorable relative to that of CRBP-II (Figure 5). Noteworthy,

Table 1. Entropy (S^{con}) and Entropy Difference (ΔS) of the Whole Protein, and Its Core and Entry Portal Site, Determined from the Analysis of the $S\mu\text{s}$ MD Trajectory^a

system	protein ^b	core ^c	portal site ^d
$S^{\text{con}}(\text{apo-I})^e$	2.97	1.76	1.21
$S^{\text{con}}(\text{holo-I})$	2.41	1.59	0.81
$S^{\text{con}}(\text{apo-II})$	2.62	1.68	0.94
$S^{\text{con}}(\text{holo-II})$	2.48	1.61	0.87
$\Delta S(\text{apo-I and holo-I})$	0.56	0.17	0.38
$\Delta S(\text{apo-II and holo-II})$	0.14	0.07	0.07
$\Delta S(\text{apo-I and apo-II})$	0.35	0.08	0.27
$\Delta S(\text{holo-I and holo-II})$	-0.07	-0.02	-0.06

^aValues (kcal mol⁻¹ K⁻¹) were determined considering only the backbone atoms. ^bResidues 7–134. ^cExcluding the structural elements of the entry portal site. ^dHelices αI and αII and loops $\beta C-\beta D$ and $\beta E-\beta F$. ^eThe error of the conformational entropy was estimated from the standard deviation of S^{con} obtained in the fitting at increasing simulation windows, with an upper value of 0.008 kcal mol⁻¹ K⁻¹ for the whole protein and the portal site and 0.005 kcal mol⁻¹ K⁻¹ for the protein core.

7335

DOI: 10.1021/acs.jpclett.9b02861
J. Phys. Chem. Lett. 2019, 10, 7333–7339

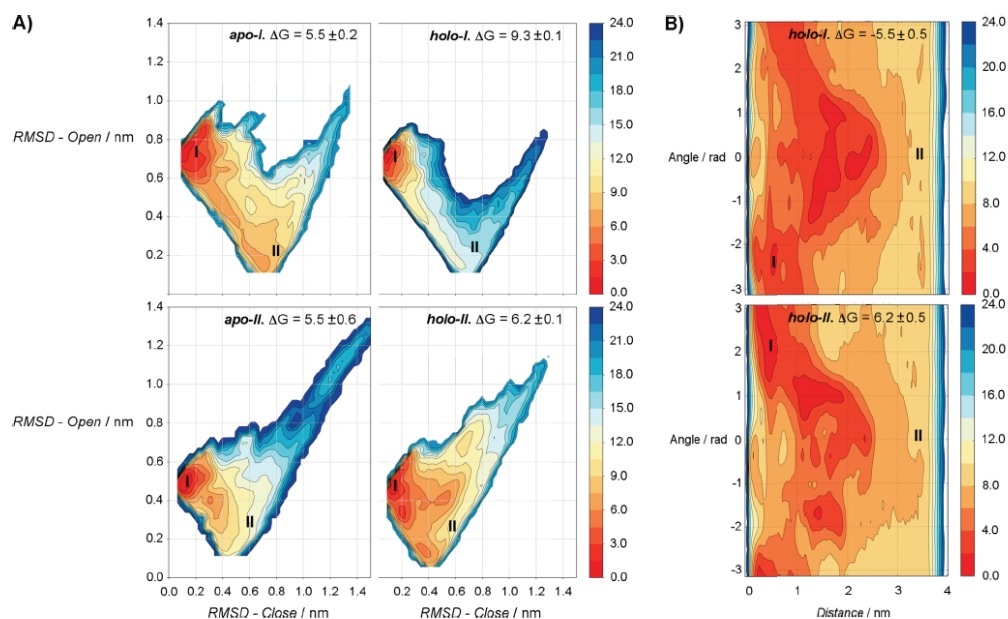


Figure 3. (A) Free energy surface for the opening/closing of the entry portal site in *apo* and *holo* forms of CRBP-I and -II. Closed and open states are indicated by symbols I and II, respectively. Contour lines are drawn every 1.5 kcal mol⁻¹. Values in the plots are the average of three estimates generated by using different reference structures for the open state. (B) Free energy surface for the binding/unbinding of retinol from *holo*-I and *holo*-II. Bound and unbound states are indicated by symbols I and II, respectively. Contour lines are drawn every 2 kcal mol⁻¹.

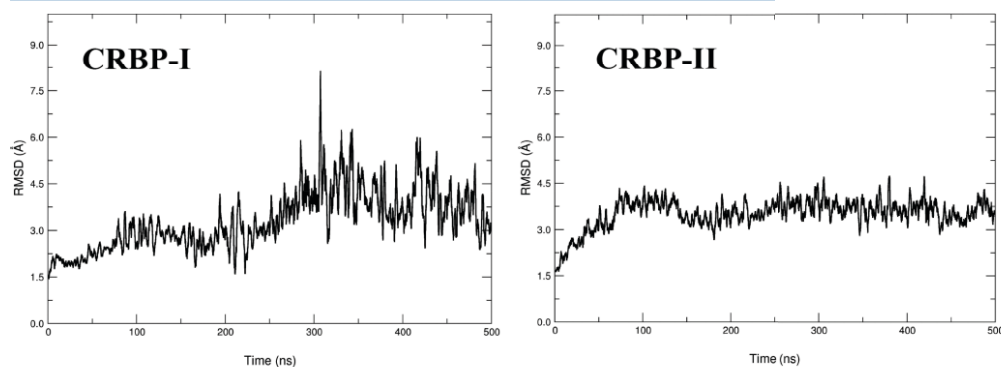


Figure 4. RMSD (Å) of retinol along the opening of the portal site in (left) CRBP-I and (right) -II. The RMSD was determined relative to the arrangement of retinol in the energy-minimized structure of *holo*-I and *holo*-II species, after alignment of the protein core. The snapshots were taken during the 500 ns of the pT-metaD simulation at 300 K.

this agrees with the experimentally observed greater affinity of retinol for CRBP-I as the predicted affinity lies between the range of experimental values, which vary from an upper threshold of $< \sim 100$ -fold¹² to a lower limit of 3.3-fold greater affinity for CRBP-I.²³ Remarkably, our results also revealed that the difference in binding affinity is mainly determined by the opening/closing of the entry portal site in the *holo* state. This suggests that the larger cost of opening the *holo*-I complex

cannot be attributed to the interactions formed by the portal site with the rest of the protein as the free energy changes determined for the opening of the portal site in the *apo* species are highly similar in the two isoforms (Figure 5; see also SI Figure S13). Therefore, it may be speculated that the interactions formed between retinol and the entry portal site in the *holo* species are more favorable in CRBP-I than those in CRBP-II, thus providing a basis to justify the larger decrease in

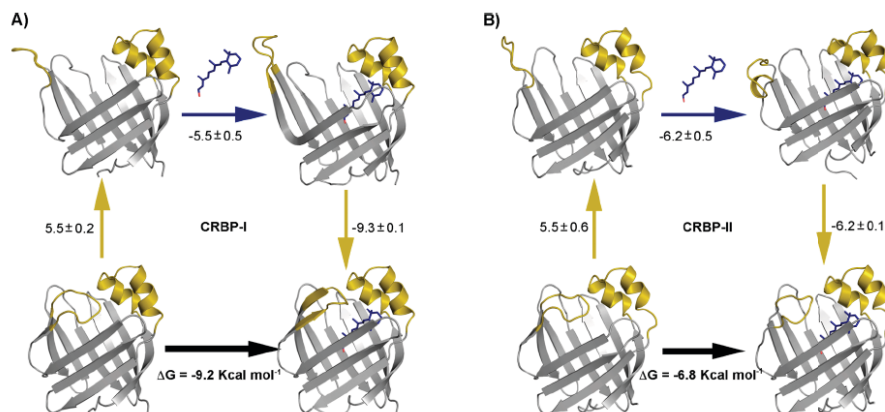


Figure 5. Thermodynamic cycle for the opening of the entry portal site for *apo* states, the binding of retinol, and the closure of the portal site in *holo* systems in (A) CRBP-I and (B) CRBP-II. Values are in kcal mol⁻¹.

conformational flexibility observed upon retinol binding to CRBP-I relative to CRBP-II.

This assumption was confirmed from the analysis of the interaction energies between retinol and the structural elements that define the portal site in *holo-I* and *holo-II* (Table 2). Whereas the interaction energy with helices αI and

Table 2. Interaction Energies (E_{int} ; kcal mol⁻¹) and Its Electrostatic (E_{ele} ; kcal mol⁻¹) and van der Waals (E_{vdw}) Components between Retinol and the Entry Portal Site for *holo-I* and *holo-II*

system ^a	E_{int}	E_{ele}	E_{vdw}
<i>holo-I</i> (αI and αII)	-10.6	0.2	-10.8
<i>holo-II</i> (αI and αII)	-11.0	-0.4	-10.6
<i>holo-I</i> ($\beta C-\beta D$)	-6.4	0.1	-6.5
<i>holo-II</i> ($\beta C-\beta D$)	-6.0	0.3	-6.3
<i>holo-I</i> ($\beta E-\beta F$)	-6.4	-0.2	-6.2
<i>holo-II</i> ($\beta E-\beta F$)	-3.8	-0.4	-3.4

^aCalculations performed for 50 snapshots taken regularly in the last microsecond of the 5 μ s MD simulations. Helices αI and αII comprise residues Glu15–Leu37, and loops $\beta C-\beta D$ and $\beta E-\beta F$ involve residues Ser55–Asn59 and Glu73–Cys/Val83, respectively.

αII and loop $\beta C-\beta D$ was similar in the two *holo* systems, the interaction of retinol with loop $\beta E-\beta F$ was 2.6 kcal mol⁻¹

more stabilizing in *holo-I*. Further decomposition into pairwise ligand–residue contributions revealed that the energy difference was mainly due to the interactions with Gly77 and Ile78 in CRBP-I, which were 2.1 kcal mol⁻¹ more stabilizing than the interactions with Gly77 and Leu78 in CRBP-II (SI Table S3). In contrast, other residue substitutions located in the loop $\beta E-\beta F$ contributed less than 0.2 kcal mol⁻¹, even though this can be justified from either the solvent-exposed arrangement of the side chain of these residues or the large distance from the mutated residue to retinol. In contrast, residues at position 78 (Ile in CRBP-I, Leu in CRBP-II) are located at the top of the loop $\beta E-\beta F$, pointing toward the interior of the β -barrel, and form van der Waals contacts with the β -ionone ring and the unsaturated chain of retinol (Figure 6). Overall, these results point out that the difference in the interaction energy with retinol can be mainly attributed to the conservative mutation of Ile78 in CRBP-I to Leu78 in CRBP-II, disclosing an unexpected effect related to the methyl isomerism between the side chains of these two residues.

To confirm the impact of the Ile/Leu mutation at position 78 on the binding of retinol to CRBP-I and -II, SPR was used to characterize the kinetic rate constants for the association (k_{on}) and dissociation (k_{off}) of retinol to CRBP-II and its Leu78 \rightarrow Ile single-mutated variant (SI Figures S14 and S15 and Table S4). The results show that the k_{on} remains essentially unaltered for both CRBP-II and the mutated

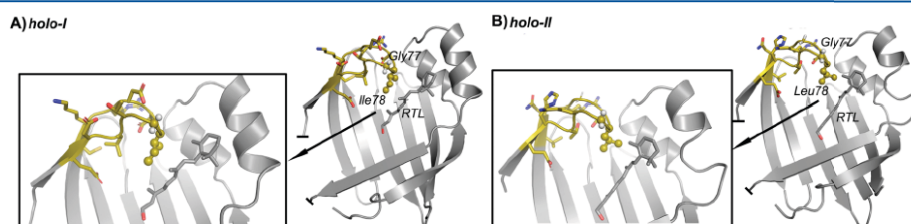


Figure 6. Representation of (A) *holo-I* and (B) *holo-II*. Residues in the $\beta E-\beta F$ loop with higher contribution to the interaction energy with retinol (RTL) are highlighted as spheres, and retinol is shown as sticks (β -sheets C and D are not shown for the sake of clarity).

variant (Table 3). However, the k_{off} of retinol is slowed down by a factor of ~ 2.2 in the mutated protein. The increased

Table 3. Kinetic Rate Constants (k_{on} , $\text{M}^{-1} \text{s}^{-1}$; k_{off} , s^{-1}) and Dissociation Constant (K_{D} ; M) for the Interactions between Retinol and CRBP-II and Its Leu78 \rightarrow Ile Single-Mutated Variant

CRBP-II	k_{on}	k_{off}	K_{D}
wild-type	241.5	17.9×10^{-3}	7.4×10^{-5}
Leu78 \rightarrow Ile	312.4	8.1×10^{-3}	2.6×10^{-5}

residence time originating from the single-point mutation Leu78 \rightarrow Ile agrees with the expected strengthening of the interaction of retinol with the mutated residue in CRBP-I (Ile) relative to CRBP-II (Leu), as deduced from the PT-metaD simulations and the decomposition analysis presented above (Table 2 and SI Table S3). Furthermore, the dissociation constant (K_{D}) decreased by ~ 2.8 -fold in the mutated CRBP-II, which compares with the lower limit of the experimental ratio between CRBP-I and CRBP-II (~ 3.3 -fold).²⁵

The selectivity of different members of cytosolic binding proteins toward distinct retinoid-like compounds has been related to the presence of specific residues in the β -barrel.^{27,28} Furthermore, other studies have identified hydrophobic contacts between ligands and structural elements of the mobile gate in SEC14-like family members.^{29–31} Present results, however, point out that a seemingly minor chemical change related to the methyl isomerism between Ile and Leu at the portal site modulates the binding properties of retinol between closely related CRBP isoforms. The net effect is the enhanced free energy penalty associated with the closed \rightarrow open transition, which would disfavor the release of the ligand and increase the residence time of retinol in the interior of the β -barrel. Noteworthy, the affinity for the two isoforms is finely modulated by the differential interaction of the β -ionone unit of retinol with the residue (Ile/Leu) at the top of the loop $\beta\text{E}-\beta\text{F}$, suggesting an unexpected role of the methyl isomerism between the two similar residues.

From a functional point of view, these results unveil a subtle regulation mechanism that underlies the distinct physiological role of the two isoforms. In enterocytes, CRBP-II plays an important, but not essential, role in assisting the transient exchange of retinol from the intestinal lumen to the lymph, making it necessary to have an efficient delivery system. In contrast, CRBP-I is highly expressed in hepatic cells, where it participates in the storage of retinol and controls its mobilization to ensure a steady supply in the blood plasma. Therefore, the strengthened interaction of retinol with the β -barrel lid may have evolved as a mechanism to self-regulate long-term retinol mobilization subject to specific requirements of active retinoid metabolites to the target cells without being affected by the fluctuations of the dietary intake.

Finally, these findings demonstrate that conservative changes in specific residues at remote sites distinct from the binding pocket, which should not alter the gross structural and physicochemical features of the protein, may result in fine-tuning of the ligand's binding properties. Thus, a thoughtful selection of residue variations may be instrumental for engineering a gradual evolution of structure–function relationships in nanomolecular devices.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcllett.9b02861.

Additional results for the structural, dynamical, and energetic analysis of the retinol-carrier proteins, results from SPR assays, and description of theoretical and experimental methods (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: abidonchanalb@ub.edu (A.B.-C.).

*E-mail: f.l.gervasio@ucl.ac.uk (F.L.G.).

*E-mail: ffluque@ub.edu (F.J.L.).

ORCID

Carolina Estarellas: 0000-0002-0944-9053

Francesca Spyarakis: 0000-0002-4016-227X

Axel Bidon-Chanal: 0000-0002-1666-1490

Francesco Luigi Gervasio: 0000-0003-4831-5039

F. Javier Luque: 0000-0002-8049-3567

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank the Ministerio de Economía y Competitividad (MINECO: SAF2017-88107-R, SAF2015-68749-R; AEI/FEDER UE) and the Generalitat de Catalunya (2017SGR1746) for financial support, the Barcelona Supercomputing Center (BSC; BCV-2013-1-0016 and BCV-2013-2-0016), PRACE (BSC; 2010PA3167), and HecBiosim (EPSRC EP/R029407/1) for computational resources. G.S. and F.L.G. acknowledge EPSRC (EP/M013898/1; EP/P022138/1; EP/P011306/1) for financial support. S.C. acknowledges the University of Barcelona for an APIF fellowship. C.E. thanks MINECO for a postdoctoral fellowship (FPDI-2013-15572) and funding from the EU Horizon 2020 research and innovation programme under Marie Skłodowska-Curie Grant Agreement No. 795116.

■ REFERENCES

- Blomhoff, R.; Blomhoff, H. K. Overview of Retinoid Metabolism and Function. *J. Neurobiol.* 2006, 66, 606–630.
- Olson, C. R.; Mello, C. V. Significance of Vitamin A to Brain Function, Behavior and Learning. *Mol. Nutr. Food Res.* 2010, 54, 489–495.
- Bass, N. M. Cellular Binding Proteins for Fatty Acids and Retinoids: Similar or Specialized Functions? *Mol. Cell. Biochem.* 1993, 123, 191–202.
- Napoli, J. L. Cellular Retinoid Binding-Proteins, CRBP, CRABP, FABP5: Effects on Retinoid Metabolism, Function and Related Diseases. *Pharmacol. Ther.* 2017, 173, 19–33.
- Schaap, F. G.; van der Vusse, G. J.; Glatz, J. F. C. Evolution of the Family of Intracellular Lipid Binding Proteins in Vertebrates. *Mol. Cell. Biochem.* 2002, 239, 69–77.
- Saito, K.; Tautz, L.; Mustelin, T. The Lipid-Binding SEC14 Domain. *Biochim. Biophys. Acta, Mol. Cell Biol. Lipids* 2007, 1771, 719–726.
- Bankaitis, V. A.; Mousley, C. J.; Schaaf, G. The Sec14-Superfamily and Mechanisms for Crosstalk Between Lipid Metabolism and Lipid Signalling. *Trends Biochem. Sci.* 2010, 35, 150–160.
- Sato, Y.; Arai, H.; Miyata, A.; Tokita, S.; Yamamoto, K.; Tanabe, T.; Inoue, K. Primary Structure of alpha-Tocopherol Transfer Protein

- from Rat Liver. Homology with Cellular Retinaldehyde-Binding Protein. *J. Biol. Chem.* 1993, 268, 17705–17710.
- (9) Sha, B.; Phillips, S. E.; Bankaitis, V. A.; Luo, M. Crystal Structure of the *Saccharomyces cerevisiae* Phosphatidylinositol-Transfer Protein. *Nature* 1998, 391, 506–510.
- (10) Kono, N.; Ohto, U.; Hiramatsu, T.; Urabe, M.; Uchida, Y.; Satow, Y.; Arai, H. Impaired α -TTP-PIPs Interactions Underlies Familial Vitamin E Deficiency. *Science* 2013, 340, 1106–1110.
- (11) Lamprakis, C.; Stocker, A.; Cascella, M. Mechanisms of Recognition and Binding of α -TTP to the Plasma Membrane by Multi-Scale Molecular Dynamics Simulations. *Front. Mol. Biosci.* 2015, 2, 36.
- (12) Li, E.; Qian, S. J.; Winter, N. S.; Avignon, A.; Levin, M. S.; Gordon, J. I. Fluorine Nuclear Magnetic Resonance Analysis of the Ligand Binding Properties of Two Homologous Rat Cellular Retinol-Binding Proteins Expressed in *Escherichia coli*. *J. Biol. Chem.* 1991, 266, 3622–3629.
- (13) Eriksson, U.; Das, K.; Busch, C.; Nordlinder, H.; Rask, L.; Sundelin, J.; Sällström, J.; Peterson, P. A. Cellular Retinol-Binding Protein. Quantitation and Distribution. *J. Biol. Chem.* 1984, 259, 13464–13470.
- (14) Bushue, N.; Wan, Y. J. Retinoid Pathway and Cancer Therapeutics. *Adv. Drug Delivery Rev.* 2010, 62, 1285–1298.
- (15) Franzoni, L.; Lücke, C.; Pérez, C.; Cavazzini, D.; Rademacher, M.; Ludwig, C.; Spisni, A.; Rossi, G. L.; Rüterjans, H. Structure and Backbone Dynamics of Apo- and Holo-Cellular Retinol-Binding Protein in Solution. *J. Biol. Chem.* 2002, 277, 21983–21997.
- (16) Lu, J.; Lin, C. L.; Tang, C.; Ponder, J. W.; Kao, J. L.; Cistola, D. P.; Li, E. Binding of Retinol Induces Changes in Rat Cellular Retinol-Binding Protein II Conformation and Backbone Dynamics. *J. Mol. Biol.* 2000, 300, 619–632.
- (17) Lu, J.; Cistola, D. P.; Li, E. Two Homologous Rat Cellular Retinol-Binding Proteins Differ in Local Conformational Flexibility. *J. Mol. Biol.* 2003, 330, 799–812.
- (18) Winter, N. S.; Bratt, J. M.; Banaszak, L. J. Crystal Structures of Holo and Apo-Cellular Retinol-Binding Protein II. *J. Mol. Biol.* 1993, 230, 1247–1259.
- (19) Silvaroli, J. A.; Arne, J. M.; Chelstowska, S.; Kiser, P. D.; Banerjee, S.; Golczak, M. Ligand Binding Induces Conformational Changes in Human Cellular Retinol-Binding Protein I (CRBP1) Revealed by Atomic Resolution Crystal Structures. *J. Biol. Chem.* 2016, 291, 8528–8540.
- (20) Menozzi, I.; Vallese, F.; Polverini, E.; Folli, C.; Berni, R.; Zanotti, G. Structural and Molecular Determinants Affecting the Interaction of Retinol with Human CRBP1. *J. Struct. Biol.* 2017, 197, 330–339.
- (21) Nossoni, Z.; Assar, Z.; Yapiçi, I.; Nosrati, M.; Wang, W.; Berbasova, T.; Vasileiou, C.; Borhan, B.; Geiger, J. Structures of Holo Wild-Type Human Cellular Retinol-Binding Protein II (hCRBP1) Bound to Retinol and Retinal. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 2014, 70, 3226–3232.
- (22) Tarter, M.; Capaldi, S.; Carrizo, M. E.; Ambrosi, E.; Perduca, M.; Monaco, H. L. Crystal Structure of Human Cellular Retinol-Binding Protein II to 1.2 Å Resolution. *Proteins: Struct., Funct., Genet.* 2008, 70, 1626–1630.
- (23) Kane, M. A.; Bright, F. V.; Napoli, J. L. Binding Affinities of CRBP1 and CRBP2 for 9-*cis*-Retinoids. *Biochim. Biophys. Acta, Gen. Subj.* 2011, 1810, 514–518.
- (24) Franzoni, L.; Cavazzini, D.; Rossi, G. L.; Lücke, C. New Insights on the Protein-Ligand Interaction Differences Between the Two Primary Cellular Retinol Carriers. *J. Lipid Res.* 2010, 51, 1332–1343.
- (25) Mittag, T.; Franzoni, L.; Cavazzini, D.; Schaffhausen, B.; Rossi, G. L.; Günther, U. L. Retinol Modulates Site-Specific Mobility of Apo-Cellular Retinol-Binding Protein to Promote Ligand Binding. *J. Am. Chem. Soc.* 2006, 128, 9844–9848.
- (26) Harris, S. A.; Gavathiotis, E.; Searle, M. S.; Orozco, M.; Laughton, C. A. Cooperativity in Drug-DNA Recognition: A Molecular Dynamics Study. *J. Am. Chem. Soc.* 2001, 123, 12658–12663.
- (27) Helbling, R. E.; Bolze, C. S.; Golczak, M.; Palczewski, K.; Stocker, A.; Cascella, M. Cellular Retinaldehyde Binding Protein—Different Binding Modes and Micro-Solvation Patterns for High-Affinity 9-*cis*- and 11-*cis*-Retinal Substrates. *J. Phys. Chem. B* 2013, 117, 10719–10729.
- (28) Bolze, C. S.; Helbling, R. E.; Owen, R. L.; Pearson, A. R.; Pompidor, G.; Dworkowski, F.; Fuchs, M. R.; Furrer, J.; Golczak, M.; Palczewski, K.; Cascella, M.; Stocker, A. Human Cellular Retinaldehyde-Binding Protein Has Secondary Thermal 9-*cis*-Retinal Isomerase Activity. *J. Am. Chem. Soc.* 2014, 136, 137–146.
- (29) Meier, R.; Tomizaki, T.; Schulze-Briese, C.; Baumann, U.; Stocker, A. The Molecular Basis of Vitamin E Retention: Structure of Human α -Tocopherol Transfer Protein. *J. Mol. Biol.* 2003, 331, 725–734.
- (30) He, X.; Lobsiger, J.; Stocker, A. Bothnia Distrophy is Caused by Domino-Like Rearrangements in Cellular Retinaldehyde-Binding Protein Mutant R234W. *Proc. Natl. Acad. Sci. U. S. A.* 2009, 106, 18545–18550.
- (31) Christen, M.; Marcaida, M. J.; Lamprakis, C.; Aeschmann, W.; Vaithilingam, J.; Schneider, P.; Hilbert, M.; Schneider, G.; Cascella, M.; Stocker, A. Structural Insights on Cholesterol Endosynthesis: Binding of Squalene and 2,3-Oxidosqualene to Supernatant Protein Factor. *J. Struct. Biol.* 2015, 190, 261–270.

Bibliography

- (1) Drewes, G.; Knapp, S. Chemoproteomics and Chemical Probes for Target Discovery. *Trends Biotechnol.* **2018**, *36* (12), 1275–1286.
- (2) Bunnage, M. E.; Chekler, E. L. P.; Jones, L. H. Target Validation Using Chemical Probes. *Nat. Chem. Biol.* **2013**, *9* (4), 195–199.
- (3) MacArron, R.; Banks, M. N.; Bojanic, D.; Burns, D. J.; Cirovic, D. A.; Garyantes, T.; Green, D. V. S.; Hertzberg, R. P.; Janzen, W. P.; Paslay, J. W.; et al. Impact of High-Throughput Screening in Biomedical Research. *Nat. Rev. Drug Discov.* **2011**, *10* (3), 188–195.
- (4) Ligands, S. *Protein-Ligand Interactions*; 2013; Vol. 1008.
- (5) Van Montfort, R. L. M.; Workman, P. Structure-Based Drug Design: Aiming for a Perfect Fit. *Essays Biochem.* **2017**, *61* (5), 431–437.
- (6) Jaskolski, M.; Dauter, Z.; Wlodawer, A. A Brief History of Macromolecular Crystallography, Illustrated by a Family Tree and Its Nobel Fruits. *FEBS J.* **2014**, *281* (18), 3985–4009.
- (7) Navia, M. A.; Fitzgerald, P. M. D.; McKeever, B. M.; Leu, C.-T.; Heimbach, J. C.; Herber, W. K.; Sigal, I. S.; Darke, P. L.; Springer, J. P. Three-Dimensional Structure of Aspartyl Protease from Human Immunodeficiency Virus HIV-1. *Nature* **1989**, *337* (6208), 615–620.
- (8) Washburn, S. A Second Front against AIDS Feedback Cooling of Currents. 6–7.
- (9) Hubbard, R. E. 3D Structure and the Drug Discovery Process. *Struct. Drug Discov.* **2007**, 1–31.
- (10) Shuker, S. B.; Hajduk, P. J.; Meadows, R. P.; Fesik, S. W. Discovering High-Affinity Ligands for Proteins: SAR by NMR. *Science* (80-.). **1996**, *274* (5292), 1531–1534.
- (11) Oltersdorf, T.; Elmore, S. W.; Shoemaker, A. R.; Armstrong, R. C.; Augeri, D. J.; Belli, B. A.; Bruncko, M.; Deckwerth, T. L.; Dinges, J.; Hajduk, P. J.; et al. An Inhibitor of Bcl-2 Family Proteins Induces Regression of Solid Tumours. *Nature* **2005**, *435* (7042), 677–681.
- (12) Renaud, J. P.; Chung, C. W.; Danielson, U. H.; Egner, U.; Hennig, M.; Hubbard, R. E.; Nar, H. Biophysics in Drug Discovery: Impact, Challenges and Opportunities. *Nat. Rev. Drug Discov.* **2016**, *15* (10), 679–698.
- (13) Lamoree, B.; Hubbard, R. E. Current Perspectives in Fragment-Based Lead Discovery (FBLD). *Essays Biochem.* **2017**, *61* (5), 453–464.
- (14) Moreno-Chicano, T.; Ebrahim, A.; Axford, D.; Appleby, M. V.; Beale, J. H.; Chaplin, A. K.; Duyvesteyn, H. M. E.; Ghiladi, R. A.; Owada, S.; Sherrell, D. A.; et al. High-Throughput Structures of Protein{-}ligand Complexes at Room Temperature Using Serial Femtosecond Crystallography. *IUCrJ* **2019**, *6* (6), 1074–1085.
- (15) Da Fonseca, P. C. A.; Morris, E. P. Cryo-EM Reveals the Conformation of a Substrate Analogue in the Human 20S Proteasome Core. *Nat. Commun.* **2015**, *6* (May), 2–7.
- (16) Martin-Garcia, J. M.; Conrad, C. E.; Coe, J.; Roy-Chowdhury, S.; Fromme, P. Serial Femtosecond Crystallography: A Revolution in Structural Biology. *Arch. Biochem. Biophys.* **2016**, *602*, 32–47.
- (17) Johansson, L. C.; Stauch, B.; Ishchenko, A.; Cherezov, V. A Bright Future for Serial Femtosecond Crystallography with XFELs. *Trends Biochem. Sci.* **2017**, *42*

Bibliography

- (9), 749–762.
- (18) Bai, X. chen; McMullan, G.; Scheres, S. H. W. How Cryo-EM Is Revolutionizing Structural Biology. *Trends Biochem. Sci.* **2015**, *40* (1), 49–57.
- (19) Liao, M.; Cao, E.; Julius, D.; Cheng, Y. Structure of the TRPV1 Ion Channel Determined by Electron Cryo-Microscopy. *Nature* **2013**, *504* (7478), 107–112.
- (20) Bai, X. C.; Yan, C.; Yang, G.; Lu, P.; Ma, D.; Sun, L.; Zhou, R.; Scheres, S. H. W.; Shi, Y. An Atomic Structure of Human γ -Secretase. *Nature* **2015**, *525* (7568), 212–217.
- (21) Ballone, A.; Centorrino, F.; Ottmann, C. 14-3-3: A Case Study in Ppi Modulation. *Molecules* **2028**, *23* (6), 1–14.
- (22) Saur, M.; Hartshorn, M. J.; Dong, J.; Reeks, J.; Bunkoczi, G.; Jhoti, H.; Williams, P. A. Fragment-Based Drug Discovery Using Cryo-EM. *Drug Discov. Today* **2020**, *25* (3), 485–490.
- (23) Winegarden, W. The Economics of Pharmaceutical Pricing. *Econ. Pharm. Pricing* **2014**, No. June, 1–29.
- (24) Wu, T.; Yu, J.; Gale-day, Z.; Woo, A.; Suresh, A.; Hornsby, M.; Gestwicki, J. E. Three Essential Resources to Improve Differential Scanning Fluorimetry (DSF) Experiments. *bioRxiv* **2020**.
- (25) Simeonov, A. Recent Developments in the Use of Differential Scanning Fluorometry in Protein and Small Molecule Discovery and Characterization. *Expert Opin. Drug Discov.* **2013**, *8* (9), 1071–1082.
- (26) Molina, D. M.; Jafari, R.; Ignatushchenko, M.; Seki, T.; Larsson, E. A.; Dan, C.; Sreekumar, L.; Cao, Y.; Nordlund, P. Monitoring Drug Target Engagement in Cells and Tissues Using the Cellular Thermal Shift Assay. *Science (80-.)*. **2013**, *341* (6141), 84–87.
- (27) Olaru, A.; Bala, C.; Jaffrezic-Renault, N.; Aboul-Enein, H. Y. Surface Plasmon Resonance (SPR) Biosensors in Pharmaceutical Analysis. *Crit. Rev. Anal. Chem.* **2015**, *45* (2), 97–105.
- (28) Patching, S. G. Surface Plasmon Resonance Spectroscopy for Characterisation of Membrane Protein-Ligand Interactions and Its Potential for Drug Discovery. *Biochim. Biophys. Acta - Biomembr.* **2014**, *1838* (1 PARTA), 43–55.
- (29) Copeland, R. A.; Pompliano, D. L.; Meek, T. D. Drug–Target Residence Time and Its Implications for Lead Optimization. *Nat. Rev. Drug Discov.* **2006**, *5* (9), 730–739.
- (30) GE Healthcare. Biacore Systems in Small Molecule Drug Discovery. *Appl. Note* **2016**, *2* (28997553 AC).
- (31) Lea, W. A.; Simeonov, A. Fluorescence Polarization Assay. *Expert Opin Drug Discov.* **2012**, *6* (1), 17–32.
- (32) Rossi, A. M.; Taylor, C. W. Analysis of Protein-Ligand Interactions by Fluorescence Polarization. *Nat. Protoc.* **2011**, *6* (3), 365–387.
- (33) Goldflam, M.; Tarragó, T.; Gairí, M.; Giralt, E. NMR Studies of Protein–Ligand Interactions. In *Protein NMR Techniques*; Shekhtman, A., Burz, D. S., Eds.; Humana Press: Totowa, NJ, **2012**; pp 233–259.
- (34) Becker, W.; Bhattiprolu, K. C.; Gubensäk, N.; Zangger, K. Investigating Protein–Ligand Interactions by Solution Nuclear Magnetic Resonance Spectroscopy. *ChemPhysChem* **2018**, *19* (8), 895–906.
- (35) Kleinpeter, A. B.; Jureka, A. S.; Falahat, S. M.; Green, T. J.; Petit, C. M. Structural Analyses Reveal the Mechanism of Inhibition of Influenza Virus NS1 by Two Antiviral Compounds. *J. Biol. Chem.* **2018**, *293* (38), 14659–14668.

Bibliography

- (36) Pellecchia, M.; Sem, D. S.; Wüthrich, K. NMR in Drug Discovery. *Nat. Rev. Drug Discov* **2002**, *1* (3), 211–219.
- (37) Broussard, J. A.; Green, K. J. Research Techniques Made Simple: Methodology and Applications of Förster Resonance Energy Transfer (FRET) Microscopy. *J. Invest. Dermatol.* **2017**, *137* (11), e185–e191.
- (38) Stryer, L.; Haugland, R. P. Energy Transfer: A Spectroscopic Ruler. *Proc. Natl. Acad. Sci. U. S. A.* **1967**, *58* (2), 719–726.
- (39) Algar, W. R.; Hildebrandt, N.; Vogel, S. S.; Medintz, I. L. FRET as a Biomolecular Research Tool — Understanding Its Potential While Avoiding Pitfalls. *Nat. Methods* **2019**, *16* (9), 815–829.
- (40) Ghisaidoobe, A.; Chung, S. Intrinsic Tryptophan Fluorescence in the Detection and Analysis of Proteins: A Focus on Förster Resonance Energy Transfer Techniques. *Int. J. Mol. Sci.* **2014**, *15* (12), 22518–22538.
- (41) Comley, J. TR-FRET Based Assays - Getting Better with Age. **2006**, *7*, 22–37.
- (42) Baker, M. Fragment-Based Phenotypic Screening Is a Hit. *Nat. Rev. Drug Discov.* **2017**, *16* (4), 225–226.
- (43) Parker, C. G.; Galmozzi, A.; Wang, Y.; Correia, B. E.; Sasaki, K.; Joslyn, C. M.; Kim, A. S.; Cavallaro, C. L.; Lawrence, R. M.; Johnson, S. R.; et al. Ligand and Target Discovery by Fragment-Based Screening in Human Cells. *Cell* **2017**, *168* (3), 527–541.e29.
- (44) Bohacek, R. S.; McMartin, C.; Guida, W. C. The Art and Practice of Structure-Based Drug Design: A Molecular Modeling Perspective. *Med. Res. Rev.* **1996**, *16* (1), 3–50.
- (45) Erlanson, D. A.; Fesik, S. W.; Hubbard, R. E.; Jahnke, W.; Jhoti, H. Twenty Years on: The Impact of Fragments on Drug Discovery. *Nat. Rev. Drug Discov.* **2016**, *15* (9), 605–619.
- (46) Hopkins, A. L.; Groom, C. R.; Alex, A. Ligand Efficiency: A Useful Metric for Lead Selection. *Drug Discov. Today* **2004**, *9* (10), 430–431.
- (47) Coyne, A. G.; Scott, D. E.; Abell, C. Drugging Challenging Targets Using Fragment-Based Approaches. *Curr. Opin. Chem. Biol.* **2010**, *14* (3), 299–307.
- (48) Morreale, F. E.; Testa, A.; Chaugule, V. K.; Bortoluzzi, A.; Ciulli, A.; Walden, H. Mind the Metal: A Fragment Library-Derived Zinc Impurity Binds the E2 Ubiquitin-Conjugating Enzyme Ube2T and Induces Structural Rearrangements. *J. Med. Chem.* **2017**, *60* (19), 8183–8191.
- (49) Baell, J.; Walters, M. A. Introducing the Pains, Nature 2014. *Nature* **2014**, *513*, 8–10.
- (50) Baker, L. M.; Aimon, A.; Murray, J. B.; Surgenor, A. E.; Matassova, N.; Roughley, S. D.; Collins, P. M.; Krojer, T.; von Delft, F.; Hubbard, R. E. Rapid Optimisation of Fragments and Hits to Lead Compounds from Screening of Crude Reaction Mixtures. *Commun. Chem.* **2020**, *3* (1).
- (51) Hajduk, P. J.; Greer, J. A Decade of Fragment-Based Drug Design: Strategic Advances and Lessons Learned. *Nat. Rev. Drug Discov.* **2007**, *6* (3), 211–219.
- (52) Schuffenhauer, A.; Ruedisser, S.; Marzinzik, A.; Jahnke, W.; Selzer, P.; Jacoby, E. Library Design for Fragment Based Screening. *Curr. Top. Med. Chem.* **2005**, *5* (8), 751–762.
- (53) Mashalidis, E. H.; Śledź, P.; Lang, S.; Abell, C. A Three-Stage Biophysical Screening Cascade for Fragment-Based Drug Discovery. *Nat. Protoc.* **2013**, *8* (11), 2309–2324.
- (54) Robson-Tull, J. Biophysical Screening in Fragment-Based Drug Design: A Brief

Bibliography

- Overview. *Biosci. Horizons* **2018**, *11*, 1–12.
- (55) Souers, A. J.; Levenson, J. D.; Boghaert, E. R.; Ackler, S. L.; Catron, N. D.; Chen, J.; Dayton, B. D.; Ding, H.; Enschede, S. H.; Fairbrother, W. J.; et al. ABT-199, a Potent and Selective BCL-2 Inhibitor, Achieves Antitumor Activity While Sparing Platelets. *Nat. Med.* **2013**, *19* (2), 202–208.
- (56) Bollag, G.; Tsai, J.; Zhang, J.; Zhang, C.; Ibrahim, P.; Nolop, K.; Hirth, P. Vemurafenib: The First Drug Approved for BRAF-Mutant Cancer. *Nat. Rev. Drug Discov* **2012**, *11* (11), 873–886.
- (57) Chevillard, F.; Rimmer, H.; Betti, C.; Pardon, E.; Ballet, S.; Van Hilten, N.; Steyaert, J.; Diederich, W. E.; Kolb, P. Binding-Site Compatible Fragment Growing Applied to the Design of β 2 -Adrenergic Receptor Ligands. *J. Med. Chem.* **2018**, *61* (3), 1118–1129.
- (58) Sabbah, M.; Mendes, V.; Vistal, R. G.; Dias, D. M. G.; Záhorská, M.; Mikušová, K.; Korduláková, J.; Coyne, A. G.; Blundell, T. L.; Abell, C. Fragment-Based Design of Mycobacterium Tuberculosis Inha Inhibitors. *J. Med. Chem.* **2020**, *63* (9), 4749–4761.
- (59) Hoffer, L.; Voitovich, Y. V.; Raux, B.; Carrasco, K.; Muller, C.; Fedorov, A. Y.; Derviaux, C.; Amouric, A.; Betzi, S.; Horvath, D.; et al. Integrated Strategy for Lead Optimization Based on Fragment Growing: The Diversity-Oriented-Target-Focused-Synthesis Approach. *J. Med. Chem.* **2018**, *61* (13), 5719–5732.
- (60) Elledge, S. J.; Jeffrey, P. D.; Kinnucan, E. R. E.; Bowen, Z.; Carrano, A. C.; Schulman, B. A.; Finnin, M. S.; Harper, J. W.; Pagano, M.; Pavletich, N. P. Insights into SCF Ubiquitin Ligases from the Structure of the Skp1–Skp2 Complex. *Nature* **2002**, *408* (6810), 381–386.
- (61) Peet, N. P.; Sunder, S.; Barbuch, R. J. Synthesis of Thiazolotriazepinones. *J. Heterocycl. Chem.* **1982**, *19* (4), 747–752.
- (62) Li, W.; Bengtson, M. H.; Ulbrich, A.; Matsuda, A.; Reddy, V. A.; Orth, A.; Chanda, S. K.; Batalov, S.; Joazeiro, C. A. P. Genome-Wide and Functional Annotation of Human E3 Ubiquitin Ligases Identifies MULAN, a Mitochondrial E3 That Regulates the Organelle’s Dynamics and Signaling. *PLoS One* **2008**, *3* (1), 1–14.
- (63) Ciechanover, A. Proteolysis: From the Lysosome to Ubiquitin and the Proteasome. *Nat. Rev. Mol. Cell Biol.* **2005**, *6* (1), 79–87.
- (64) Varshavsky, A. The Early History of the Ubiquitin Field. *Protein Sci.* **2006**, *15* (3), 647–654.
- (65) Tundo, G. R.; Sbardella, D.; Santoro, A. M.; Coletta, A.; Oddone, F.; Grasso, G.; Milardi, D.; Lacal, P. M.; Marini, S.; Purrello, R.; et al. The Proteasome as a Druggable Target with Multiple Therapeutic Potentialities: Cutting and Non-Cutting Edges. *Pharmacol. Ther.* **2020**, *213*, 107579.
- (66) Galdeano, C. Drugging the Undruggable: Targeting Challenging E3 Ligases for Personalized Medicine. *Future Med. Chem.* **2017**, *9* (4), 347–350.
- (67) Lopez-Girona, A.; Mendy, D.; Ito, T.; Miller, K.; Gandhi, A. K.; Kang, J.; Karasawa, S.; Carmel, G.; Jackson, P.; Abbasian, M.; et al. Cereblon Is a Direct Protein Target for Immunomodulatory and Antiproliferative Activities of Lenalidomide and Pomalidomide. *Leukemia* **2012**, *26* (11), 2326–2335.
- (68) Wells, J. A.; McClendon, C. L. Reaching for High-Hanging Fruit in Drug Discovery at Protein-Protein Interfaces. *Nature* **2007**, *450* (7172), 1001–1009.
- (69) Galdeano, C.; Gadd, M. S.; Soares, P.; Scaffidi, S.; Van Molle, I.; Birced, I.; Hewitt, S.; Dias, D. M.; Ciulli, A. Structure-Guided Design and Optimization of Small Molecules Targeting the Protein-Protein Interaction between the von Hippel-

Bibliography

- Lindau (VHL) E3 Ubiquitin Ligase and the Hypoxia Inducible Factor (HIF) Alpha Subunit with in Vitro Nanomolar Affinities. *J. Med. Chem.* **2014**, *57* (20).
- (70) Chan, C. H.; Morrow, J. K.; Li, C. F.; Gao, Y.; Jin, G.; Moten, A.; Stagg, L. J.; Ladbury, J. E.; Cai, Z.; Xu, D.; et al. Pharmacological Inactivation of Skp2 SCF Ubiquitin Ligase Restricts Cancer Stem Cell Traits and Cancer Progression. *Cell* **2013**, *154* (3), 556–568.
- (71) Davies, T. G.; Wixted, W. E.; Coyle, J. E.; Griffiths-Jones, C.; Hearn, K.; McMenamin, R.; Norton, D.; Rich, S. J.; Richardson, C.; Saxty, G.; et al. Monoacidic Inhibitors of the Kelch-like ECH-Associated Protein 1: Nuclear Factor Erythroid 2-Related Factor 2 (KEAP1:NRF2) Protein-Protein Interaction with High Cell Potency Identified by Fragment-Based Discovery. *J. Med. Chem.* **2016**, *59* (8), 3991–4006.
- (72) Ito, T.; Ando, H.; Suzuki, T.; Ogura, T.; Hotta, K.; Imamura, Y.; Yamaguchi, Y.; Handa, H. Identification of a Primary Target of Thalidomide Teratogenicity. *Science* (80-.). **2010**, *327* (5971), 1345–1350.
- (73) Spradlin, J. N.; Hu, X.; Ward, C. C.; Brittain, S. M.; Jones, M. D.; Ou, L.; To, M.; Proudfoot, A.; Ornelas, E.; Woldegiorgis, M.; et al. Harnessing the Anti-Cancer Natural Product Nimbolide for Targeted Protein Degradation. *Nat. Chem. Biol.* **2019**, *15* (7), 747–755.
- (74) Gadd, M. S.; Testa, A.; Lucas, X.; Chan, K. H.; Chen, W.; Lamont, D. J.; Zengerle, M.; Ciulli, A. Structural Basis of PROTAC Cooperative Recognition for Selective Protein Degradation. *Nat. Chem. Biol.* **2017**, *13* (5), 514–521.
- (75) Konstantinidou, M.; Li, J.; Zhang, B.; Wang, Z.; Shaabani, S.; Brake, F. Ter; Essa, K.; Dömling, A. PROTACs— a Game-Changing Technology. *Expert Opin. Drug Discov.* **2019**, *14* (12), 1255–1268.
- (76) Lai, A. C.; Crews, C. M. Induced Protein Degradation: An Emerging Drug Discovery Paradigm. *Nat. Rev. Drug Discov.* **2017**, *16* (2), 101–114.
- (77) Sakamoto, K. M.; Kim, K. B.; Kumagai, A.; Mercurio, F.; Crews, C. M.; Deshaies, R. J. Protacs: Chimeric Molecules That Target Proteins to the Skp1-Cullin-F Box Complex for Ubiquitination and Degradation. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98* (15), 8554–8559.
- (78) Schneekloth, J. S.; Fonseca, F. N.; Koldobskiy, M.; Mandal, A.; Deshaies, R.; Sakamoto, K.; Crews, C. M. Chemical Genetic Control of Protein Levels: Selective in Vivo Targeted Degradation. *J. Am. Chem. Soc.* **2004**, *126* (12), 3748–3754.
- (79) Lu, J.; Qian, Y.; Altieri, M.; Dong, H.; Wang, J.; Raina, K.; Hines, J.; Winkler, J. D.; Crew, A. P.; Coleman, K.; et al. Hijacking the E3 Ubiquitin Ligase Cereblon to Efficiently Target BRD4. *Chem. Biol.* **2015**, *22* (6), 755–763.
- (80) Buckley, D. L.; Raina, K.; Darricarrere, N.; Hines, J.; Gustafson, J. L.; Smith, I. E.; Miah, A. H.; Harling, J. D.; Crews, C. M. HaloPROTACS: Use of Small Molecule PROTACs to Induce Degradation of HaloTag Fusion Proteins. *ACS Chem. Biol.* **2015**, *10* (8), 1831–1837.
- (81) Winter, G. E.; Buckley, D. L.; Paulk, J.; Roberts, J. M.; Souza, A.; Dhe-Paganon, S.; Bradner, J. E. Phthalimide Conjugation as a Strategy for in Vivo Target Protein Degradation. *Science* (80-.). **2015**, *348* (6241), 1376–1381.
- (82) Bondeson, D. P.; Mares, A.; Smith, I. E. D.; Ko, E.; Campos, S.; Miah, A. H.; Mulholland, K. E.; Routly, N.; Buckley, D. L.; Gustafson, J. L.; et al. Catalytic in Vivo Protein Knockdown by Small-Molecule PROTACs. *Nat. Chem. Biol.* **2015**, *11* (8), 611–617.

Bibliography

- (83) Zengerle, M.; Chan, K. H.; Ciulli, A. Selective Small Molecule Induced Degradation of the BET Bromodomain Protein BRD4. *ACS Chem. Biol.* **2015**, *10* (8), 1770–1777.
- (84) McCoull, W.; Cheung, T.; Anderson, E.; Barton, P.; Burgess, J.; Byth, K.; Cao, Q.; Castaldi, M. P.; Chen, H.; Chiarparin, E.; et al. Development of a Novel B-Cell Lymphoma 6 (BCL6) PROTAC To Provide Insight into Small Molecule Targeting of BCL6. *ACS Chem. Biol.* **2018**, *13* (11), 3131–3141.
- (85) Bassi, Z. I.; Fillmore, M. C.; Miah, A. H.; Chapman, T. D.; Maller, C.; Roberts, E. J.; Davis, L. C.; Lewis, D. E.; Galwey, N. W.; Waddington, K. E.; et al. Modulating PCAF/GCN5 Immune Cell Function through a PROTAC Approach. *ACS Chem. Biol.* **2018**, *13* (10), 2862–2867.
- (86) Zorba, A.; Nguyen, C.; Xu, Y.; Starr, J.; Borzilleri, K.; Smith, J.; Zhu, H.; Farley, K. A.; Ding, W. D.; Schiemer, J.; et al. Delineating the Role of Cooperativity in the Design of Potent PROTACs for BTK. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115* (31), E7285–E7292.
- (87) Popow, J.; Arnhof, H.; Bader, G.; Berger, H.; Ciulli, A.; Covini, D.; Dank, C.; Gmaschitz, T.; Greb, P.; Karolyi-Özguer, J.; et al. Highly Selective PTK2 Proteolysis Targeting Chimeras to Probe Focal Adhesion Kinase Scaffolding Functions. *J. Med. Chem.* **2019**, *62* (5), 2508–2520.
- (88) Nunes, J.; McGonagle, G. A.; Eden, J.; Kiritharan, G.; Touzet, M.; Lewell, X.; Emery, J.; Eidam, H.; Harling, J. D.; Anderson, N. A. Targeting IRAK4 for Degradation with PROTACs. *ACS Med. Chem. Lett.* **2019**, *10* (7), 1081–1085.
- (89) Scheepstra, M.; Hekking, K. F. W.; van Hijfte, L.; Folmer, R. H. A. Bivalent Ligands for Protein Degradation in Drug Discovery. *Comput. Struct. Biotechnol. J.* **2019**, *17*, 160–176.
- (90) Mayor-Ruiz, C.; Bauer, S.; Brand, M.; Kozička, Z.; Siklos, M.; Imrichova, H.; Kaltheuner, I. H.; Hahn, E.; Seiler, K.; Koren, A.; et al. Rational Discovery of Molecular Glue Degradors via Scalable Chemical Profiling. *Nat. Chem. Biol.* **2020**, *16* (11), 1199–1207.
- (91) Arvinas. Arvinas Releases Updated Dose Escalation Data from Clinical Trial of PROTAC® Protein Degradator ARV-110 in Patients with Metastatic Castration-Resistant Prostate Cancer. *Press Release* **2020**, No. May, 1–6.
- (92) Skaar, J. R.; Pagan, J. K.; Pagano, M. Mechanisms and Function of Substrate Recruitment by F-Box Proteins. *Nat. Rev. Mol. Cell Biol.* **2013**, *14* (6), 369–381.
- (93) Welcker, M.; Clurman, B. E. FBW7 Ubiquitin Ligase: A Tumour Suppressor at the Crossroads of Cell Division, Growth and Differentiation. *Nat. Rev. Cancer* **2008**, *8* (2), 83–93.
- (94) King, B.; Trimarchi, T.; Reavie, L.; Xu, L.; Mullenders, J.; Ntziachristos, P.; Aranda-Orgilles, B.; Perez-Garcia, A.; Shi, J.; Vakoc, C.; et al. XThe Ubiquitin Ligase FBXW7 Modulates Leukemia-Initiating Cell Activity by Regulating MYC Stability. *Cell* **2013**, *153* (7), 1552–1566.
- (95) Yalla, K.; Elliott, C.; Day, J. P.; Findlay, J.; Barratt, S.; Hughes, Z. A.; Wilson, L.; Whiteley, E.; Popiolek, M.; Li, Y.; et al. FBXW7 Regulates DISC1 Stability via the Ubiquitin-Proteasome System. *Mol. Psychiatry* **2018**, *23* (5), 1278–1286.
- (96) Hao, B.; Oehlmann, S.; Sowa, M. E.; Harper, J. W.; Pavletich, N. P. Structure of a Fbw7-Skp1-Cyclin E Complex: Multisite-Phosphorylated Substrate Recognition by SCF Ubiquitin Ligases. *Mol. Cell* **2007**, *26* (1), 131–143.
- (97) Orlicky, S.; Tang, X.; Willems, A.; Tyers, M.; Sicheri, F. Structural Basis for Phosphodependent Substrate Selection and Orientation by the SCFCdc4 Ubiquitin

Bibliography

- Ligase. *Cell* **2003**, *112* (2), 243–256.
- (98) Shimizu, K.; Nihira, N. T.; Inuzuka, H.; Wei, W. Physiological Functions of FBW7 in Cancer and Metabolism. *Cell. Signal.* **2018**, *46* (November 2017), 15–22.
- (99) Yumimoto, K.; Nakayama, K. I. Recent Insight into the Role of FBXW7 as a Tumor Suppressor. *Semin. Cancer Biol.* **2020**, No. November 2019, 0–1.
- (100) Yeh, C. H.; Bellon, M.; Nicot, C. FBXW7: A Critical Tumor Suppressor of Human Cancers. *Mol. Cancer* **2018**, *17* (1), 1–19.
- (101) Vom, A.; Headey, C. S.; Wang, A. G.; Capuano, A. Ben; Yuriev, E.; Scanlon, A. M. J.; A, J. S. S. Detection and Prevention of Aggregation-Based False Positives in STD-NMR-Based Fragment Screening. *Aust. J. Chem* **2013**, *66*, 1518–1524.
- (102) Laplante, S. R.; Carson, R.; Gillard, J.; Aubry, N.; Coulombe, R.; Bordeleau, S.; Bonneau, P.; Little, M.; O'Meara, J.; Beaulieu, P. L. Compound Aggregation in Drug Discovery: Implementing a Practical NMR Assay for Medicinal Chemists. *J. Med. Chem.* **2013**, *56* (12), 5142–5150.
- (103) Le Guilloux, V.; Schmidtke, P.; Tuffery, P. Fpocket: An Open Source Platform for Ligand Pocket Detection. *BMC Bioinformatics* **2009**, *10*, 1–11.
- (104) Schmidtke, P.; Bidon-chanal, A.; Luque, F. J.; Barril, X. MDpocket: Open-Source Cavity Detection and Characterization on Molecular Dynamics Trajectories. *Bioinformatics* **2011**, *27* (23), 3276–3285.
- (105) Chemical Computing Group Inc. Molecular Operating Environment (MOE), 2014.09. 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7 **2015**.
- (106) Ruiz-Carmona, S.; Schmidtke, P.; Luque, F. J.; Baker, L.; Matassova, N.; Davis, B.; Roughley, S.; Murray, J.; Hubbard, R.; Barril, X. Dynamic Undocking and the Quasi-Bound State as Tools for Drug Discovery. *Nat. Chem.* **2017**, *9* (3), 201–206.
- (107) Jakalian, A.; Jack, D. B.; Bayly, C. I. Fast, Efficient Generation of High-Quality Atomic Charges. AM1-BCC Model: II. Parameterization and Validation. *J. Comput. Chem.* **2002**, *23* (16), 1623–1641.
- (108) Bayly, C. I.; McKay, D.; Truchon, J. F. An Informal AMBER Small Molecule Force Field: Parm@Frosst. Merck & Co. Internal Development Release **2011**.
- (109) Alvarez-Garcia, D.; Barril, X. Molecular Simulations with Solvent Competition Quantify Water Displaceability and Provide Accurate Interaction Maps of Protein Binding Sites. *J. Med. Chem.* **2014**, *57* (20), 8530–8539.
- (110) Tjernberg, A.; Markova, N.; Griffiths, W. J.; Hallén, D. DMSO-Related Effects in Protein Characterization. *J. Biomol. Screen.* **2006**, *11* (2), 131–137.
- (111) Roughley, S. D.; Hubbard, R. E. How Well Can Fragments Explore Accessed Chemical Space? A Case Study from Heat Shock Protein 90. *J. Med. Chem.* **2011**, *54* (12), 3989–4005.
- (112) Giordanetto, F.; Jin, C.; Willmore, L.; Feher, M.; Shaw, D. E. Fragment Hits: What Do They Look Like and How Do They Bind? *J. Med. Chem.* **2019**, *62* (7), 3381–3394.
- (113) Gao, K.; Oerlemans, R.; Groves, M. R. Theory and Applications of Differential Scanning Fluorimetry in Early-Stage Drug Discovery. *Biophys. Rev.* **2020**, *12* (1), 85–104.
- (114) Bulfer, S. L.; Jean-Francois, F. L.; Arkin, M. R. Making FBDD Work in Academia. In *Fragment-based Drug Discovery Lessons and Outlook*; John Wiley & Sons, Ltd, **2016**; pp 223–246.
- (115) Simonetta, K. R.; Taygerly, J.; Boyle, K.; Basham, S. E.; Padovani, C.; Lou, Y.; Cummins, T. J.; Yung, S. L.; von Soly, S. K.; Kayser, F.; et al. Prospective

Bibliography

- Discovery of Small Molecule Enhancers of an E3 Ligase-Substrate Interaction. *Nat. Commun.* **2019**, *10* (1).
- (116) Bondeson, D. P.; Smith, B. E.; Burslem, G. M.; Buhimschi, A. D.; Hines, J.; Jaime-Figueroa, S.; Wang, J.; Hamman, B. D.; Ishchenko, A.; Crews, C. M. Lessons in PROTAC Design from Selective Degradation with a Promiscuous Warhead. *Cell Chem. Biol.* **2018**, *25* (1), 78-87.e5.
- (117) Kane, M. A.; Bright, F. V.; Napoli, J. L. Binding Affinities of CRBPI and CRBP2 for 9-Cis-Retinoids. *Biochim. Biophys. Acta - Gen. Subj.* **2011**, *1810* (5), 514–518.
- (118) Estarellas, C.; Scaffidi, S.; Saladino, G.; Spyralis, F.; Franzoni, L.; Galdeano, C.; Bidon-Chanal, A.; Gervasio, F. L.; Luque, F. J. Modulating Ligand Dissociation through Methyl Isomerism in Accessory Sites: Binding of Retinol to Cellular Carriers. *J. Phys. Chem. Lett.* **2019**, 7333–7339.
- (119) Filippakopoulos, P.; Knapp, S. Targeting Bromodomains: Epigenetic Readers of Lysine Acetylation. *Nat. Rev. Drug Discov.* **2014**, *13* (5), 337–356.
- (120) Filippakopoulos, P.; Picaud, S.; Mangos, M.; Keates, T.; Lambert, J. P.; Barsyte-Lovejoy, D.; Felletar, I.; Volkmer, R.; Müller, S.; Pawson, T.; et al. Histone Recognition and Large-Scale Structural Analysis of the Human Bromodomain Family. *Cell* **2012**, *149* (1), 214–231.
- (121) Pérez-Salvia, M.; Esteller, M. Bromodomain Inhibitors and Cancer Therapy: From Structures to Applications. *Epigenetics* **2017**, *12* (5), 323–339.
- (122) Galdeano, C.; Ciulli, A. Selectivity On-Target of Bromodomain Chemical Probes by Structure-Guided Medicinal Chemistry and Chemical Biology. *Future Med. Chem.* **2016**, *8* (13), 1655–1680.
- (123) Chung, C. W.; Dean, A. W.; Woolven, J. M.; Bamborough, P. Fragment-Based Discovery of Bromodomain Inhibitors Part 1: Inhibitor Binding Modes and Implications for Lead Discovery. *J. Med. Chem.* **2012**, *55* (2), 576–586.
- (124) Shi, J.; Vakoc, C. R. The Mechanisms behind the Therapeutic Activity of BET Bromodomain Inhibition. *Mol. Cell* **2014**, *54* (5), 728–736.
- (125) Cochran, A. G.; Conery, A. R.; Sims, R. J. Bromodomains: A New Target Class for Drug Development. *Nat. Rev. Drug Discov.* **2019**, *18* (8), 609–628.
- (126) Zuber, J.; Shi, J.; Wang, E.; Rappaport, A. R.; Herrmann, H.; Sison, E. A.; Magoon, D.; Qi, J.; Blatt, K.; Wunderlich, M.; et al. RNAi Screen Identifies Brd4 as a Therapeutic Target in Acute Myeloid Leukaemia. *Nature* **2011**, *478* (7370), 524–528.
- (127) Baratta, M. G.; Schinzel, A. C.; Zwang, Y.; Bandopadhyay, P.; Bowman-Colin, C.; Kutt, J.; Curtis, J.; Piao, H.; Wong, L. C.; Kung, A. L.; et al. An In-Tumor Genetic Screen Reveals That the BET Bromodomain Protein, BRD4, Is a Potential Therapeutic Target in Ovarian Carcinoma. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112* (1), 232–237.
- (128) Khan, Y. M.; Kirkham, P.; Barnes, P. J.; Adcock, I. M. Brd4 Is Essential for IL-1 β -Induced Inflammation in Human Airway Epithelial Cells. *PLoS One* **2014**, *9* (4), 1–17.
- (129) Chung, C. W.; Coste, H.; White, J. H.; Mirguet, O.; Wilde, J.; Gosmini, R. L.; Delves, C.; Magny, S. M.; Woodward, R.; Hughes, S. A. Discovery and Characterization of Small Molecule Inhibitors of the BET Family Bromodomains. *J. Med. Chem.* **2011**, *54* (11), 3827–3838.
- (130) Gehling, V. S.; Hewitt, M. C.; Vaswani, R. G.; Leblanc, Y.; Coité, A.; Nasveschuk, C. G.; Taylor, A. M.; Harmange, J. C.; Audia, J. E.; Pardo, E.; et al. Discovery, Design, and Optimization of Isoxazole Azepine BET Inhibitors. *ACS Med. Chem.*

Bibliography

- Lett.* **2013**, *4* (9), 835–840.
- (131) Chung, C. W.; Coste, H.; White, J. H.; Mirguet, O.; Wilde, J.; Gosmini, R. L.; Delves, C.; Magny, S. M.; Woodward, R.; Hughes, S. A.; et al. Discovery and Characterization of Small Molecule Inhibitors of the BET Family Bromodomains. *J. Med. Chem.* **2011**, *54* (11), 3827–3838.
- (132) Schneider, G. Automating Drug Discovery. *Nat. Rev. Drug Discov.* **2018**, *17* (2), 97–113.