

*Digital Comprehensive Summaries of Uppsala Dissertations
from the Faculty of Science and Technology 2662*

Decoding the ubiquitin system: charting and characterizing motif-based interactions of E3 ligases

MAXIMILIAN VIELER



ACTA UNIVERSITATIS
UPSALIENSIS
2026



UPPSALA
UNIVERSITET

Dissertation presented at Uppsala University to be publicly examined in Biomedicinskt centrum (BMC) A1:111a, Husargatan 3, 752 37 Uppsala, Thursday, 4 June 2026 at 13:15 for the degree of Doctor of Philosophy. The examination will be conducted in English. Faculty examiner: Associate Professor Magnus Kjærgaard (Department of Molecular Biology and Genetics Neurobiology, Aarhus University, Aarhus, Denmark).

Abstract

Vieler, M. 2026. Decoding the ubiquitin system: charting and characterizing motif-based interactions of E3 ligases. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology* 2662. 95 pp. Uppsala: Acta Universitatis Upsaliensis. ISBN 978-91-513-2808-9.

E3 ligases are central players in protein homeostasis, deciding which proteins are degraded by conjugating ubiquitin onto their target proteins. Beyond degradation, ubiquitination regulates a myriad of further cellular signaling pathways. E3 ligases have developed different strategies to identify their ligands, one of which is the recognition of an autonomous interaction-mediating compact binding sites called Short Linear Motif (SLiMs). SLiMs are typically 3-10-residue stretches embedded in intrinsically disordered regions, and commonly serve as binding interfaces. SLiMs which lead to ubiquitination and proteasomal degradation upon binding to their target E3 ligases are called a degrons. SLiM-based protein-protein interactions typically rely on 3-4 conserved key residues, with less-conserved flanking regions tuning binding affinity and specificity. Elucidating the substrate determinants of the ca. 900 E3 ligases has been one of the central challenges of the E3 ligase field.

This thesis presents the investigation of the SLiM-binding of more than 140 domains from E3 ligases and associated auxiliary proteins, distributed over four studies. SLiM discovery was based on proteomic peptide-phage display (ProP-PD) screening. ProP-PD is a high-throughput method to identify peptide ligands from the proteome, from which consensus motifs can be derived. We disclose 11,460 potential SLiM instances of the E3 ligase machinery. We describe 40 distinct SLiMs, of which 14 are either entirely novel or redefine known binding determinants. The defined SLiMs allow us to predict additional putative proteomic binding sites. Interactions were validated using peptide SPOT arrays, and characterized using fluorescence polarization-based affinity measurements. Motif key residues were further validated using alanine-scanning footprinting. Using artificial intelligence-based modelling, we localized binding sites on the folded domains and validate them through site-directed mutagenesis. Among the results, we highlight the establishment of internal peptide recognition by the TPR domain of STUB1, the definition of the peptide-binding capacity of four MIB-HERC2 domains, and the establishment of a RNF41 binding motif in USP8.

Collectively, the research provides a substantially expanded SLiM-based interaction space of E3 ligases, both in terms of interactions and motifs. By systematically mapping E3 binding motifs and binding sites at unprecedented scale, this thesis provides a resource that advances the biochemical understanding of the protein-protein interactions of E3 ligases. The insights into these interactions may contribute to the decoding of specificity determinants of ubiquitin-mediated signaling, and may lay the groundwork for future efforts to predict, manipulate, and therapeutically target ubiquitin-dependent regulatory networks.

Keywords: ProP-PD, protein-protein interactions, short linear motifs, E3 ligases, binding interface analysis, mutational analysis

Maximilian Vieler, Department of Chemistry for Life Sciences, Biochemistry, Box 576, Uppsala University, SE-75123 Uppsala, Sweden.

© Maximilian Vieler 2026

ISSN 1651-6214

ISBN 978-91-513-2808-9

URN urn:nbn:se:uu:diva-583655 (<http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-583655>)

*Our imagination is struck only by what is great;
but it belongs to the philosophy of nature,
to pause at what is little.*

- Alexander von Humboldt
Personal narrative, Vol. 1 (1814)

List of Papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.

- I. Madhu, P., Benz, C., Simonetti, L., Winters, M. J., Subbanna, M. S., Kliche, J., Gomez Lucas, L., Kotb, H., Krystkowiak, I., Segal, D., Darling, W. T. P., Larsen-Ledet, S., Vieler, M., Konstantinou, A., Zupancic, A., Varga, J. K., Görlitz, K., Miahlic, F., Ellingboe, L., Kraemer, A., Luchow, S., Xiong, R., Schueler-Furman, O., Halabelian, L., Santhakumar, S., Arrowsmith, C., Knapp, S., Erdelyi, M., Stein, A., Vincentelli, R., Pryciak, P., Davey, N. E., Ivarsson, Y. An Atlas of Short Linear Motif Mediated Human Protein-Protein Interactions. *Manuscript*.
- II. **Vieler, M.**, Varga J.K, Gomes Lucas, L., Simonetti, L., Winters, M.J., Subbanna, M.S., Konstantinou, A., Fang, B., Saxena, K., Farges, F., Lenz, C., Schwalm, M., Lima Righetto, G., Kimani, S.W., Zeng, H., Santhakumar, V., Park, S.L., Peter, M., Halabelian, L., Santhakumar, S., Arrowsmith, C., Knapp, S., Schueler-Furman, O., Pryciak, P., Davey, N.E., Ivarsson, Y. Systematic discovery of short linear motif binding across human E3 ubiquitin ligases. *Manuscript*.
- III. Chen, X., Chen, Z., **Vieler, M.**, Benz, C., Madhu, P., Richardson, W., Simonetti, L., Diallo, M., Montes, B.R., Lau, R., Mogwera, K.S.P., Manning, C.E., Adamson, R.J, Chi, G., Nilsson, J., Ivarsson, Y., Bullock A.N. Proteome-scale identification of short linear binding motifs for the BTB-KELCH E3 ligase family. *Manuscript*.
- IV. Konstantinou, A., Córdova-Pérez, A., Varga J.K., Madhu, P., Simonetti, L., **Vieler, M.**, Ishimura, R., Lamoliatte, F., Schueler-Furman, O., Davey, N.E., Kulathu. Y., Ivarsson, Y. Systematic Discovery of Motif-based Interactions of the Auxiliary Domains of USP Family Deubiquitinases. *Nature Communications, accepted*.

Reprint was made with permission from the respective publishers.

Contribution Report

All studies presented in the papers were collaborative. Below I outline my contributions:

Paper I (co-author): Protein overexpress and purification, phage display selections, fluorescence polarization (FP)-based affinity measurements, peptide SPOT array experiments, data analysis, data visualization, editing and reviewing the manuscript

Paper II (first author):

Study conceptualization, construct design, protein overexpress and purification, phage display selections, FP-based affinity measurements, peptide SPOT array experiments, data analysis, data visualization, writing original manuscript, editing and reviewing the manuscript

Paper III (co-author): Phage display selections, data analysis, data visualization.

Paper IV (co-author): Protein expression and purification, FP-based affinity measurements, peptide SPOT array experiments, data analysis

Contents

List of Papers	5
1. Introduction	13
1.1. Modularity of proteins, intrinsically disordered regions and motif-based interaction.....	13
Proteins and their modularity.....	13
Intrinsically disordered regions (IDRs)	14
Interaction units in the intrinsically disordered regions	14
Different types of SLiM-based interactions	16
Challenges in computationally identifying bona fide SLiM ligands	18
1.2. The E3 ligase machinery	19
The ubiquitination machinery.....	19
The diversity of the ubiquitin code.....	22
E3 ligase substrate recognition	23
Regulation and affinity of E3 ligase substrate recognition.....	24
Ubiquitination in health and disease.....	26
1.3. High-throughput methods for interaction and substrate discovery	27
Phage display	27
Validation of ProP–PD binders	30
1.4. Current challenges in E3 ligase research	32
2. Aim of the thesis.....	33
3. Outline of the thesis	34
4. Findings	36
4.1. Proteome-wide discovery of ligands.....	36
Proteome-wide discovery of SLiM-mediated interactions in the E3ligase machinery	36
Expanding the SLiM-based interactomes of SH3 and WW domains.....	36
Expanding the SLiM-based interactomes of KELCH domains.....	39
Exploration of peptide-binding for E3 ligase auxiliary proteins	39
4.2. Description of novel SLiMs of SPRY domains and MIB-HERC2 domains	42
Shared and domain-specific binding preferences of B30.2/SPRY domains ...	42

The MIB-HERC2 domains also bind asparagine-rich sequences.....	45
4.3. Redefining binding motifs.....	48
BRCA1 is able to bind non-phosphorylated ligands and its hetero-dimerization partner BARD1 binds peptides too.....	48
UHRF1, a methyl-lysine reader and E3 ligase, recognizes a non-methylated SLiM.....	51
The SLiM-binder TRAF4 recognizes an undiscovered SLiM.....	53
Beyond its HSP-client nature, STUB1 binds also to internal SLiM.....	54
4.4. Interplay between SLiMs binding to E3 ligases and DUBs.....	56
Defining a RNF41 binding motif and uncovering competition between the E3 ligase and the DUB USP8.....	56
DUBs also use SLiMs for ligand recognition.....	58
4.5. Additional analyses.....	59
A global protein stability profiling approach did not explain the degron nature of our ligands.....	59
Determining SLiMs enables ligand prediction.....	59
SLiM scanning of identified ligands enables us to explore multivalency that may increase the binding potential.....	60
5. Summary of findings.....	63
6. Conclusions and future perspective.....	65
7. Populärvetenskaplig sammanfattning.....	67
8. Populärwissenschaftliche Zusammenfassung.....	70
9. Reflections and Acknowledgments.....	73
10. References.....	76

This thesis was written by the author alone and without content generation using generative artificial intelligence (AI) tools. The Swedish popular science summary was translated from German using the expanded Opus 4.7 algorithm by Claude (Anthropic).

Abbreviations

AlphaScreen	amplified luminescent proximity homogeneous assay screen
ANK domain	ankyrin repeat domain
AP-MS	affinity purification–mass spectrometry
APC/C	anaphase-promoting complex/cyclosome
APEX	engineered ascorbate peroxidase
BARD1	BRCA1-associated RING domain protein 1
BioID	proximity-dependent biotin identification
BIRC6	baculoviral IAP repeat-containing protein 6
BRCA1	breast cancer type 1 susceptibility protein
BRCT	BRCA1 C-Terminus (BRCT) domain
BTB-KELCH	bric-a-brac, tramtrack, broad–KELCH
CBX4	E3 SUMO-protein ligase CBX4
CDK	cyclin-dependent kinase
CDR	complementarity determining regions
CLEC9A	C-type lectin domain family 9 member A
CRL3	cullin-RING E3 ligases complex 3
CRBN	cereblon
CTD	C-terminal domain
CTLH	C-terminal to LisH
DAZAP2	dAZ-associated protein 2
DUB	deubiquitinase
ELM	eukaryotic Linear Motif
ERBB3	receptor tyrosine-protein kinase erbB-3
ESCRT	endosomal sorting complexes required for transport 0
FACS	fluorescence-activated Cell Sorting
FBXL	F-box only like
FBXO	F-box only
FP	fluorescence polarization
GLI2	zinc finger protein GLI2
GPS profiling	global protein stability profiling

HD2 library	second-generation human disorderome library
HECT	homologous to the E6-AP carboxyl terminus
HECTD1	E3 ubiquitin-protein ligase HECTD1
HERC	HECT and RCC1-like
HERC1	probable E3 ubiquitin-protein ligase HERC1
HERC2	E3 ubiquitin-protein ligase HERC2
HHEX	hematopoietically-expressed homeobox protein HHEX
IDR	intrinsically disordered region
ISG15	interferon-stimulated gene 15
ITC	isothermal titration calorimetry
K _D	equilibrium dissociation constant
KLHL	Kelch-like protein
LIG1	DNA ligase 1
LOCI	luminescent oxygen channelling immunoassay
LUBAC	linear ubiquitin assembly complex
MAGE domain	melanoma-associated antigen domain
MAGEA4	melanoma-associated antigen A4
MAPK	mitogen-activated protein kinase
MATH domain	meprin and TRAF Homology domain
MDM2	mouse double minute 2
MEX3C	RNA-binding E3 ubiquitin-protein ligase MEX3C
MIB-HERC2 domain	mind bomb-HECT and RCC1-like 2 domain
MoMaP	motif map of the proteome
MoRF	molecular recognition feature
MST	micro scale thermophoresis
NEDD4	Neural Precursor Cell-Expressed Developmentally Downregulated Protein 4
NEDD8	Neural Precursor Cell-Expressed Developmentally Downregulated Protein 8
NER	nucleotide excision repair
OTX2	homeobox protein OTX2
PH domain	plant homeodomain
PI3K/Akt	phosphatidylinositol 3-kinase/Akt
PLC γ	phospholipase C gamma
PPI	protein-protein interactions
ProP-PD	proteomic peptide-phage display
PROTAC	proteolysis targeting chimera
PRR7	proline-rich protein 7

PSMD4	26S proteasome non-ATPase regulatory subunit 4
PSSM	position-specific scoring matrix
PTM	post-translational modification
RAD23B	lysine-specific demethylase RAD23B
RANBP9	Ran-binding protein 9
RANBP10	Ran-binding protein 10
RBR	RING between RING
RiboVD library	Riboviria viral disorderome library
RING	really interesting new gene
RNF41	E3 ubiquitin-protein ligase NRDP1
SAGA complex	Spt-Ada-Gcn5 acetyltransferase complex
SDM	site directed mutagenesis
SH3	Src homology 3
SH3RF1	E3 ubiquitin-protein ligase SH3RF1
SH3RF3	E3 ubiquitin-protein ligase SH3RF3
SIAH	seven in absentia homolog
SIMBA	systematic intracellular motif binding analysis
SLiM	short linear motif
SNP	single nucleotide polymorphism
SPR	surface plasmon resonance
SPRY domain	SP1a and ryanodine receptor domain
SPSB	SPRY domain-containing SOCS box proteins
STUB1	E3 ubiquitin-protein ligase CHIP
SUMO	small ubiquitin-like modifier
TCF3	transcription factor E2 α
TCF4	transcription factor 4
TGF- β	transforming growth factor- β
TPR domain	tetratricopeptide repeat domain
TRAF domain	TNF receptor-associated factor domain
TRIM	tripartite motif-containing protein
TT domain	tandem tudor domain
UBL domain	ubiquitin like domain
UHRF1	E3 ubiquitin-protein ligase UHRF1
USP8	ubiquitin carboxyl-terminal hydrolase
Y2H	yeast two-hybrid
zf-UBP	zinc finger-ubiquitin binding protein

1. Introduction

1.1. Modularity of proteins, intrinsically disordered regions and motif-based interaction

Proteins and their modularity

Proteins constitute the fundamental functional units, or players, of the cell. Confined by the cell membrane, proteins operate in the intracellular environment is enclosed by a cell membrane. They act together with other biomolecules, such as lipids, carbohydrates, and nucleic acids in diverse combinations. The contacts that these biomolecules engage in, or avoid, are collectively termed molecular interactions. Here, we focus on protein-protein interactions (PPI).

PPIs occur on a spectrum of interface size, binding affinity, and frequency of occurrence. At one end of this spectrum are large, stable protein complexes characterized by extensive interaction surfaces and low turnover rates. Toward the other end lie transient, dynamic interactions in which shorter sequence elements – molecular recognition features (MoRFs), short binding motifs (SLiMs), or post-translationally modified residues – are recognized by modular protein domains^{1,2}. Transient protein interactions involve smaller binding interfaces and are much more numerous than the stable interactions¹, and they are critical for rapid responses related to cell signaling¹. The modular nature of proteins uses this spectrum of discrete structural units to mediate specific functions and interactions³.

This thesis will focus on the modular components that contribute to the interactions of human ubiquitin E3 ligases, which are enzymes with a central role in maintaining cellular proteostasis. Before examining E3 ligase interactions in detail, the following sections introduce the modular components and the mechanism that drive molecular recognition.

Intrinsically disordered regions (IDRs)

A substantial fraction of our proteome is intrinsically disordered, meaning that they do not fold into stable three-dimensional structures under physiological conditions. Indeed, around 30% of the eukaryotic proteome contains intrinsically disordered regions (IDRs) that are longer than 30 residues⁴. The inability of IDRs to take well-folded three-dimensional structures under native conditions comes from a sequence composition that is rich in charged and polar residues and depleted of hydrophobic, core-forming residues⁵. The inability to form a structure can be represented by a folding energy landscape with multiple energy minima⁶, leading to a heterogeneous ensemble of structures in which structural state are interconverted a rapid time scale⁵. Among the conformations sampled, transient local secondary structures or transient long-range contacts can occur, depending on the sequence composition of the IDRs.

The balance between positively and negatively charged, and polar amino acids determines the compaction of IDRs⁷. IDRs dominated by like-charged amino acids - polyelectrolytes - adopt large extended conformations (named swollen coils) due to electrostatic repulsion⁸. Conversely, IDRs devoid of charged residues form compact disordered globules⁹. Finally, the conformational ensemble of IDRs is also governed by (i) physicochemical parameters such as temperature, ionic activity, or pH, (ii) macromolecular crowding (iii), solvent exposure or desiccation, and (iv) post-translational modifications (PTMs) such as phosphorylation or acetylation^{10,11}. These parameters and PTMs add regulatory complexity to IDRs and their PPIs¹².

Interaction units in the intrinsically disordered regions

The disordered nature of IDRs enables them to engage with their binding partners through their distinct sequence features. Among these, MoRFs and SLiMs have been defined^{1,2}. MoRFs are 10 to 70 residue segments within IDRs that transition from disorder to order upon interactions with their binding partner¹³. SLiMs on the other hand are typically 3 to 10 amino acid long stretches in IDRs serving as binding interfaces for their binding partners¹⁴. While MoRFs span whole regions, SLiMs are autonomous compact binding sites that independently mediate interaction. Their small size — only a handful of mediate interaction in a barcode style — allows SLiM to occur frequently within IDRs. The SLiM-based interactions are central elements of the wiring of signaling networks, and they are at the core of this thesis.

SLiMs are commonly recognized by folded modular domains. Well-characterized SLiM binding domains are for instance PDZ domains, Src homology 3 (SH3) domains, and WW domains. These domains occur in different combinations in multidomain proteins, of which some are E3 ligases¹⁵⁻¹⁸. PDZ domains are best known for binding to C-terminal motifs¹⁹. In total, there are over 250 PDZ domains in over 150 different proteins²⁰. PDZ domains are for example found in the E3 ligase LNX^{21,22}. Similarly, around 300 SH3 domains are present in roughly 220 proteins²³. The SH3 domains recognize proline-rich motifs²⁴. Lastly, ca. 50 WW domain containing proteins recognize PPxY motifs, of which there are about 1,500 motif instances in over 1,000 different proteins^{24,25}. WW domains are especially prevalent in Homologous to the E6-AP Carboxyl Terminus (HECT) type E3 ligases within the NEDD4 family²⁶. The exact number of SLiMs in the human proteome remains unknown. bio-informatic prediction estimate more than 100,000 instances¹. The Eukaryotic Linear Motif (ELM) database has around 2,400 instances²⁷ (March 2026). The Motif Map of the Proteome (MoMaP) has more than 7,100 classified human SLiM instances²⁸ (March 2026). A subset of SLiM targets E3 ligases to their substrates for proteasomal degradation, and are therefore called degradation motifs, or degrons. Much is thus to be learned about the prevalence of SLiMs, as elaborated in **Paper I-IV**.

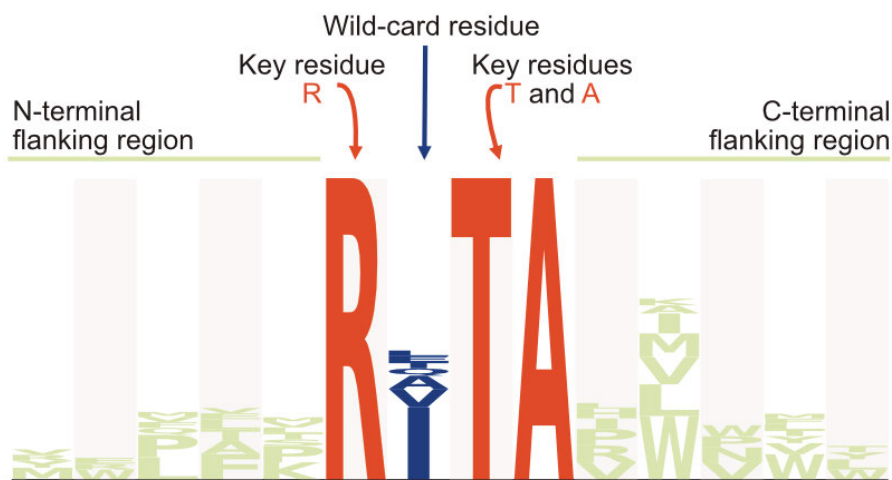


Figure 1: Example of a consensus motif representation of a SLiM. Shown is the SLiM binding to the UHRF1 tandem Tudor (TT) domain (defined in Paper II). Red residues are the key residues that drive the interaction. A wild-card position, indicated in blue, be occupied by any amino acid. The motif flanking regions are highlighted in light green. These positions typically have an amino acid preference and they fine-tune the interaction.

Different types of SLiM-based interactions

SLiM-mediated binding is typically defined by 3-4 key residues (**Figure 1**) that drive the interaction by contributing most of the binding free energy. These key residues are embedded within a context of flanking regions that modulate accessibility, affinity, and specificity. The folded proteins or protein domains that bind to SLiMs typically do not change the conformation upon binding, based on computational analysis of 103 high-resolution crystal structures of protein-peptide complexes²⁹. Instead, the disordered peptide adapts itself to the presented surface, in a so-called folding upon binding reaction^{30,31}. The peptide finds either a large pocket in which the entire peptide settles (a quarter of the cases), or it follows the knob-hole approach where a side chain inserts into a hole in the binding pocket²⁹. Consequential to the folding upon binding, a given IDR can take distinct folds upon binding to different binding pockets³², as exemplified by the C-terminal region of p53³³.

In the described coupled folding and binding mechanism, the IDR typically adopts a defined secondary structure upon binding. However, some bound IDRs remain conformationally heterogeneous. At the end of this spectrum are fully disordered complexes^{34,35}, in which both partners remain heterogeneous even in the bound state. These different modalities highlight the versatility of SLiM- and IDR-mediated recognition.

Affinity and specificity of SLiM based interactions

Affinity describes the intrinsic strength of a binding interaction³⁶. The interactions between two molecules are not discrete but a continuous spectrum, where there will always be a bound and an unbound fraction. The simplest way to describe this interaction is the 1:1 Hill-Langmuir-binding model (**Figure 2**)^{37,38}.

The ratio between the concentrations of the unbound and bound protein and ligand at equilibrium is denoted as the dissociation constant K_D (see equation 1). The dissociation constant, measured in mol/L, is the inverse of the



Figure 2: Schematic of a 1:1 protein-ligand interaction. The protein P is in green and the ligand L in blue. The interaction can be modelled by the 1:1 Langmuir-binding model as depicted to the right.

association constant K_A and can also be expressed as the ratio between the rate constants of dissociation and association (Equation 1).

$$K_D = \frac{1}{K_A} = \frac{k_{off}}{k_{on}} = \frac{[P][L]}{[PL]} \quad \text{Equation 1}$$

The change in Gibbs free energy upon binding (ΔG) determines if an interaction is thermodynamically favorable. Favorable reactions have a large negative energy change ΔG . The free energy change correlates with the dissociation constant K_D , universal gas constant R and temperature T (Equation 2)^{39,40}.

$$\Delta G = R T \ln K_D \quad \text{Equation 2}$$

The free energy change can also be described in the change of enthalpy H and the change of entropy S (Equation 3). Enthalpy H describes the total energy of the thermodynamic system (the change in heat and hence the strength of all individual bonds)^{41,41,42}, while entropy S describes the change of disorder in the system, including (1) the release of solvent (water) and ions (solvent entropy or hydrophobic enthalpy), (2) the loss in freedom of conformations by folding (conformational entropy, especially high in disordered regions), and (3) the loss of translational and rotational entropy^{39,43}.

$$\Delta G = \Delta H - T\Delta S \quad \text{Equation 3}$$

The affinity of SLiM-based interactions span wide range, from nanomolar to millimolar⁴⁴ and are often transient^{45,46}. Some interactions, such as the interactions between transcriptional co-activator CREB binding protein and its ligands⁴⁵ occur with very high affinity under given cell cycle conditions. Other interactions, such as those between importin-beta and the gradient of FG repeats in the nuclear pore complex occur with very low affinity (mid micromolar to millimolar), which allows the protein to slide through the nuclear pore^{47,48}. These examples highlight that the affinities of these interactions are tuned to the functional requirements.

Several features fine-tune affinity and specificity of SLiM-based interactions⁴⁹. As mentioned, most of the binding energy is coming from docking of motif key residues (**Figure 1**). Variation within the tolerated range of the key residues modulates the affinity. Wild-card residues and flanking regions can modulate the affinity of the interaction¹⁰. Amino acids variations of the on these positions may also serve a negative effect, by repelling other proteins with similar motif-binding preferences. PTMs of the wild-card of flanking

regions may further tune the affinities of the interactions^{50,51}. Phosphorylation sites introduce negative charges, acetylations neutralize charges, methylation inflates the side chain, ubiquitin prevent interactions altogether⁵².

Because of their compact size and their few defining amino acid, SLiMs evolve easily within a single IDR⁴⁹. Multiple instances enable multivalent binding. This may lead to increased apparent affinity of the interactions through allovalency or avidity effects which increase the apparent affinity and specificity⁵³. Allovalency can occur when multiple low-affinity SLiM instances within the same IDR compete for engaging the same binding pocket⁵⁴. This raises the local concentration of the binding motif. Avidity effects are typically caused by the binding of two or more SLiMs in a given IDR to multiple domains in a binding protein. In this case, binding of the first motif increases the local concentration for binding the second motif. Both allovalency and avidity increase the chance for rebinding before full dissociation of the complex^{10,39,55}. For avidity effects, typical examples are arrays of WW domains binding to arrays of PPxY motifs^{56,57}, as also described in **Paper I**.

Challenges in computationally identifying bona fide SLiM ligands

As SLiMs are defined amino acid patterns, it is possible to bioinformatically scan for additional motif-instances in a given proteome. Dedicated tools have been developed for this, such as SLiMsearch⁵⁸. Such scan may find bona fide ligands. The prediction is not trivial and experience has shown that there will frequent false positives^{10,59,60}. These false positive can be explained by the importance of the context (e.g. flanking regions) to binding. As mentioned above, while the core motif encodes the spatial arrangement of key side chains and thus provides sequence specificity, the flanking regions may both enhance and repel potential binders of the key motif sequence¹⁰. The repulsion between a binding domain and the flanking regions may make interactions so costly that binding is prevented. The understanding of the requirements of the flanking region for each interaction pair is still limited, and lowers the precision of SLiM binding predictions¹⁰. By pairing the motif-based searches with AlphaFold-based rankings of ligands, the precision may be improved⁶¹. Such predictions should be combined with experimental testing, for example peptide SPOT arrays as described later and used in **Papers I, II, and IV**.

1.2. The E3 ligase machinery

The ubiquitination machinery

Ubiquitination is a PTM that regulates protein stability and function through the covalent attachment of the small protein ubiquitin to target proteins (**Figure 3**). The discovery of the ubiquitin-dependent protein degradation was so scientifically impactful that Aaron Ciechanover, Avram Hershko, and Irwin Rose received the Nobel prize in Chemistry in 2004⁶². Their biochemical investigations in 1980 created a new field that has transformed our understanding of protein homeostasis.

Ubiquitination is executed by a machinery that comprises the cascade of E1 activating enzymes, E2 conjugating enzymes, E3 ligases⁶³, the 26S proteasome and ubiquitin-binding auxiliary proteins. E3 ligases play a central role in substrate selection, and thus provide specificity to the system.

Within the within this system, E3 ligases can form larger complexes for target recognition and regulation. Mechanistically, ubiquitination follows a conserved enzymatic cascade⁶⁴. The E1 enzyme is the first enzyme. It uses ATP to adenylate the C-terminal carboxylate of ubiquitin and then transfers the ubiquitin moiety onto its own cysteine to form a high energy thioester bond⁶⁵. The E2 ubiquitin-conjugating enzyme then accepts the activated ubiquitin on its cysteine. The final step is catalyzed by the E3 ligase, whose broadest

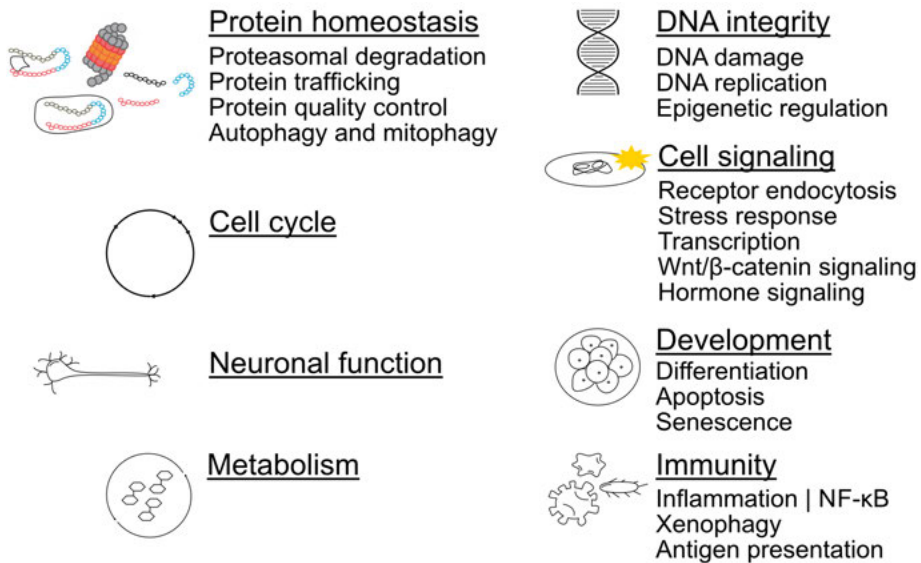


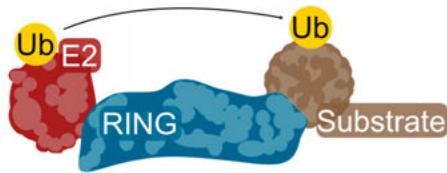
Figure 3: Overview of E3 ligase mediated cellular processes. The signalling in these processes require the correct ubiquitin chain formation depending on E3 ligase⁶⁸⁻⁷¹.

possible definition the ability to recruit an E2-Ub intermediate⁶⁶. The definition encompasses the traditional categories of HECT, RING-Between-RING (RBR) and Really Interesting New Gene (RING) E3 ligases, but also expands it to atypical E3 ligases, or catalytically inactive pseudo E3 ligases. HECT and RBR E3 ligases accept ubiquitin on a cysteine residue, while RING E3 ligases or atypical E3 ligases work as a scaffolding protein to bring the conjugated E2-enzyme in proximity to the substrate through adaptor proteins^{66,67} (**Figure 4**). Hence E3 ligases coordinate substrate specificity.

Beyond targeting substrates for degradation by the ubiquitin–proteasome system (UPS), E3 ligases also modulate protein activity, initiate biological signaling cascades, and regulate cellular processes such as cell cycle, development, stress response, DNA repair, or apoptosis (**Figure 3**)^{68–70}. They further influence intracellular trafficking, and coordinate complex signaling pathways⁷¹. Many E3 ligases also assemble into complexes such as the linear ubiquitin chain assembly (LUBAC) complex⁷², the C-Terminal to LisH (CTLH) complex⁷³, or the Anaphase-Promoting Complex/Cyclosome (APC/C) complex⁷⁴, to provide additional layers of selectivity and regulation. There are also cooperative E3 ligases where binding of two independent E3 ligases offers new substrates. The COP1–DET1 E3 ligase complex is one example of such a cooperative organization⁷⁵.

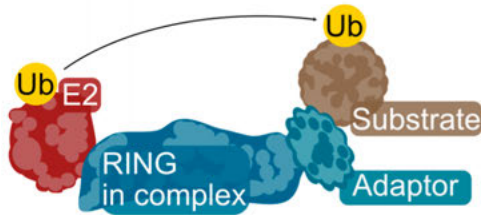
The complexity of E3 ligase machinery is further increased by the ability to ubiquitinate not only lysine but also cysteine, serine, threonine, and tyrosine. Once a substrate is ubiquitinated, the ubiquitin chain may remain as mon-ubiquitination moiety, or extend into complex branched polyubiquitin chains. Finally, there is the post-translational modification of ubiquitin itself, such as phosphorylation or acetylation. Adding to the functional scope of the E3 ligase machinery, emerging evidence has shown modifications of lipids as well as carbohydrates^{68,70}. This, however, is outside of the scope of this thesis.

RING E3 ligases and complex forming RING E3 ligases



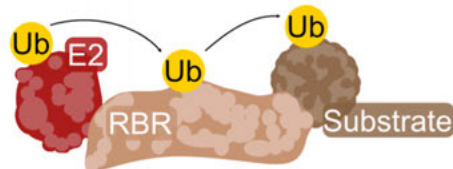
BARD1	RNF41
BRCA1	SH3RF1
MDM4	STUB1
MIB1	TRAF4
MIB2	UHRF1

Adaptor proteins



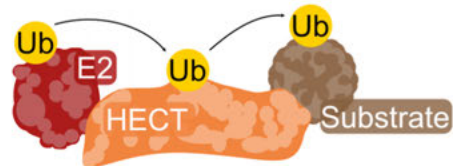
SPOPL
SPSB2
SPSB4

RBR E3 ligases



RNF31

HECT E3 ligases



HECW2	AREL1
NEDD4L	HECTD1
SMURF1	HERC1
SMURF2	HERC2
WWP1	UBR5
WWP	HECW1

Auxiliary proteins to the E3 ligases machinery



ATE1
RAD23B

Figure 4: Overview of the classical E3 ligase types and auxiliary proteins to the E3 ligase machinery with examples from the present thesis^{66-67,114}.

The diversity of the ubiquitin code

The regulation of the various ubiquitin-dependent processes relies on accurate E3 ligase substrate recognition and precise ubiquitin-chain assembly. The variety lies both in terms of which residue of ubiquitin is used to link to next ubiquitin molecules and the overall chain architecture. All seven lysine residues and the ubiquitin's initial methionine can be ubiquitinated to create extensive chains^{68,76}. Different linkages are then recognized by distinct ubiquitin-binding proteins. Monoubiquitin typically affects protein localization, activity, and interactions, rather than targeting proteins for proteasomal degradation. In homotypic polyubiquitin chains, all ubiquitin molecules are linked through the same residue, while branched chain polyubiquitination generates more complex structures. Without going into the details, it can be noted that polyubiquitination via K48 linked ubiquitination typically signals for proteasomal degradation and M1-linked ubiquitination is associated with signaling and immune response (e.g. the LUBAC complex)^{68,70,77}.

The polyubiquitin chain can further be modified by a set of 100 deubiquitinases (DUBs) remodeling ubiquitin chains by removing ubiquitin moieties from their substrates⁷⁶. How the DUBs find their substrates, and the extent to which they engage in SLiM-based interactions has also remained a fairly underexplored topic^{78,79}, which is also addressed in this thesis (**Paper IV**).

Adding to the complexity of ubiquitination are the modification on non-lysine residues (e.g. serines, threonines⁸⁰) as well as the combination with other ubiquitin-like proteins (interferon-stimulated gene 15 (ISG15)⁸¹, neural precursor cell-expressed developmentally downregulated protein 8 (NEDD8)⁸² and small ubiquitin-like modifier (SUMO)^{83,84}). PTMs further contribute to the complexity (phosphorylation⁸⁵⁻⁸⁷ or acetylation⁸⁸).

Together, this combinatorial assembly and editing of ubiquitin molecules serve as an intricate code for encoding information. With the realization that there is a code comes additional questions: how are specific ubiquitin signals generated and interpreted? In the context of this thesis, we focus on the question of how E3 ligases interact with other proteins for complex assembly and potential substrate recognition, but also venture into the SLiM-based interactions of DUBs.

E3 ligase substrate recognition

One of the remaining mysteries of the E3 ligase machinery is how the estimated 900 E3 ligases⁶⁶ are able to detect their substrates and ligate the correct ubiquitin-chain on the substrate⁷¹.

A prominent, yet not exclusive, strategy is SLiMs-mediated substrate recognition. Note, that degrons are SLiM for UPS degradation and non-degron SLiM may lead to other ubiquitin-signaling based cellular processes. SLiM-based recognition includes — among other examples — the seven in absentia homolog (SIAH) degron PxAxVxP which is recognized by SIAH RING-type ligases⁸⁹, while the MDM2-binding motif (FxxxWxx[VIL])⁹⁰, Figure 5) in p53 is recognized by the RING E3 ligase MDM2. The ETGE and DLG motifs are recognized by the Kelch domain of Keap1⁹¹. Although the established degrons are well appreciated, many E3 ligases are orphan in respect to their binding motifs.

In addition to SLiM-based interactions, E3 ligases can recognize substrates through alternative mechanisms. Some E3 ligases recognize their substrates by fold-fold interactions (F-box only 4 (FBXO4) binding to telomeric repeat factor 1⁹²), through a bipartite interface such as in the case of FBXL3-CRY2⁹³. Heterodimers can also engage with the high-order assembly such as the RING1B-BMI1 E3 ligase interacting with the whole nucleosome core particle including the histone octamer and the bound DNA⁹⁴. More special cases include Tripartite motif-containing protein 21 (TRIM21), which recognizes antibody-bound substrates via the Fc region⁹⁵.

At this stage, our understanding of E3 ligase substrate recognition remains incomplete. Known degron motifs account for only a fraction of E3 ligase–substrate interactions, and even for well-studied E3 ligase–substrate pairs, the motif determinants are often not fully resolved. Many E3 ligases lack clearly defined SLiM, and the extent of motif-based recognition versus alternative mechanisms for substrate recognition remains unclear^{71,96}. The contribution of flanking regions and PTMs are often still poorly understood¹⁰. Defining the substrate recognition determinants of E3 ligases is thus one of the main challenges of the E3 ligase field^{71,97–99}. Understanding of how E3 ligases find their interacting partners using SLiM-based interactions is a central theme of the thesis (**Paper II and III**).

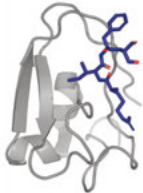
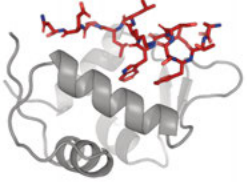
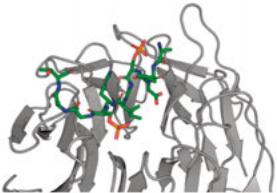
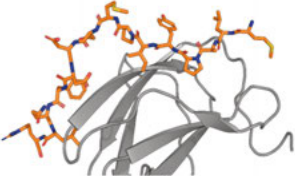
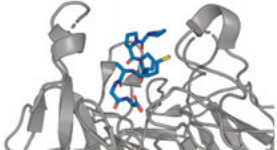
Degron type	Example	Structural representation
N-degion	UBR box of UBR2 recognizing a synthetic N-degion peptide PDB: 3NY3 - degion in blue	
Internal degion	SWIB domain of MDM2 recognizing p53 PDB: 1P22 - degion in red	
Phospho degion	WD40 domain of FBXW1 recognizing phosphorylated β -catenin PDB: 1P22 - degion in green	
Hydroxyproline degion	β -domain of pVHL binding to HIF-1 α with hydroxylated proline PDB: 1LM8 - degion in orange	
C-degion	Kelch-domain of KLHDC2 recognizing phosphorylated SELK PDB: 6DO3 - degion in blue	

Figure 5: Overview of well-studied degrons including their structures. Degrons can be coded N-terminally, internally and C-terminally. Additionally, degrons can be conditionally regulated through phosphorylation or proline hydroxylation. Other regulatory mechanisms have been reported too.

Regulation and affinity of E3 ligase substrate recognition

E3 ligase–substrate relationships are more intricate than a simple one E3–one degion relationship⁷¹. Binding motifs and degrons are diverse, varying in sequence and length, which influences the specificity and affinity. Multiple suboptimal degrons can collectively contribute to E3 ligase binding¹⁰⁰. The

following sections outline the biochemical determinants of the E3 ligase–substrate relationship.

Some degrons are regulated to achieve the desired function⁷¹ (**Figure 5**). Phosphorylation increases the affinity of so-called phosphodegrons such as those binding the substrate adaptors FBXW1¹⁰¹ and FBXW7¹⁰². The opposite has been reported, where phosphorylation inhibits the interaction. For example, phosphorylation of Thr18 prevents MDM2-binding and leads to p53 stabilization¹⁰³. A hydroxyproline-degron is a recognition mode that requires the conversion of proline to hydroxyproline for oxygen sensing^{104,105}. Other ways to activate degrons involve cofactors (SKP2 requires the cofactor CKS1)¹⁰⁶ or small molecules. Cereblon can recognize neo substrate degrons through bridging by the small molecule thalidomide¹⁰⁷.

E3 ligases are known to exhibit low affinities to their substrates¹⁰⁸. For example, the SCF complex recognizes its phosphorylated substrates in the range between 5 to 100 μM ¹⁰⁹. The low and narrow affinity range between E3 ligases and their substrates have functional implications. Firstly, low — yet specific — affinities are translated by E3 ligases to large covalent signal (compared to smaller PTMs)⁷¹. Secondly, the weak affinities may control the E3 ligase–substrate residence (E3 ligase processivity)¹¹⁰, that controls the length of the ubiquitin chain which is well studied in the APC/C complex^{110–113}. Thirdly, the narrow affinity range make E3 ligases sensitive to small changes in concentration of their competing ligands, which enables context-dependent control in respect to cell cycle or tissue expression^{71,114,115}. Lastly, E3 ligases have been observed to control proper complex formation^{116,117}, where they maintain the correct stoichiometry for nascent complexes.

Determining the affinities and affinity ranges of E3 ligase–substrate (or more broadly E3 ligase–ligand) binding aids understanding of these effects^{71,118–120}. Similarly, the dynamic relationship between the E3 ligase and the E2 conjugating enzyme control their residence time and the ubiquitin chain formation^{119–121}, which extends the functional implications of low and narrow affinity ranges beyond the E3 ligase–substrate relationships. In **Papers I, II, and III** I investigated the affinities of E3 ligases for their ligands. In **Paper IV** I focused on the impact of flanking regions on binding affinity.

Ubiquitination in health and disease

Given their central role in ubiquitin signaling, it is not surprising that defects in E3 ligase function are associated with disease. An estimated 19 % of all cancer driver genes are involved in the ubiquitin–proteasome system¹²². E3 ligase mutations are for example common in neurological diseases¹²³. Viruses are also known to hijack the E3 ligase machinery by either motif mimicry, rewiring of substrate recognition¹²⁴, or introduction of viral E3 ligases. Finally, dysregulation of TNF receptor-associated factor (TRAF)-family E3 ligases¹²⁵ or the LUBAC complex changes the immune and inflammatory response^{70,126}.

The fidelity of ubiquitin chain formation is of key concern for normal cell function and proteostasis. Consequently, the ubiquitin system has therefore attracted substantial pharmaceutical interest. Over the last decade, targeted protein degradation strategies such as proteolysis-targeting chimeras (PROTACs) and molecular glues have emerged where E3 ligases are used to selectively degrade target proteins^{127–130}. An improved understanding of the SLiM-based interactions of E3 ligases and their adaptor proteins will thus not only contribute to the understanding of cell function, but will also be of importance for potential drug discovery efforts. In this context, the SLiM containing peptides can serve either as scaffold for peptidomimetic ligand design, or as probes for more traditional inhibitor screening studies.

1.3. High-throughput methods for interaction and substrate discovery

High-throughput approaches employ different strategies to scan the human proteome for interactions. One way is to scan for binary interactions - that is pair-wise scanning for interactors. One longstanding method Yeast Two-Hybrid (Y2H) recognizes the proximity of the transcriptional activation domain and the DNA-binding domain of regulatory protein GAL4 which are fused to the proteins of interest¹³¹. Then there are affinity coupled approaches such as affinity purification coupled mass spectrometry (AP-MS) which co-isolates interaction partners from cell lysates and determines the binding partners after washing steps¹³².

Proximity labelling approaches like proximity-dependent Biotin Identification (BioID)¹³³ or ascorbate peroxidase (APEX)¹³⁴ covalently label proteins in the proximity of the target proteins. The labeled proteins are then identified through immunoprecipitation and MS. AlphaScreen (Amplified Luminescent Proximity Homogenous Assay Screen) is commercial high-through-put approach based in the Luminescent Oxygen Channeling Immunoassay (LOCI) by on the channeling of singlet oxygen species from a donor beads (which contains a photosensitizer) to an acceptor bead (containing a chemiluminescent)¹³⁵. The beads can be labeled with the respective binding partners. The singlet has a short travel distance of 200 nm which only exits acceptor beads in close proximity. AlphaScreens have been used to investigate E3 ligase substrates¹³⁶ as well as inhibitors¹³⁷. Recent developments include native mass spectrometry to examine the substrates recognition in the complexes of full E3 ligases, alone or in complex with PROTACs¹³⁸.

While these high-throughput approaches identify interactions between full length proteins, they lack the detection of transient and weak interactions which are the core of SLiM mediated PPI. To overcome these short-comings we use a proteomic peptide-phage display to search for SLiM-based interactions^{44,139}.

Phage display

Phage display was developed in 1985 by George P. Smith^{140,141} for which he received the Nobel prize in Chemistry in 2018. Coupling the genotype with the phenotype — the displayed peptide or protein — allows biopanning the binding sites¹⁴². Initially used for the development of antibodies and the

refinement of epitopes¹⁴², phage-display has a broad range of applications. One way is to use peptide libraries that present the peptide of interest as fusion proteins of the different coat proteins of the phage. While combinatorial peptide libraries — libraries that use a random combination of the 20 amino-acids — can identify biophysically optimal ligands among high sequence diversity, they do not report on proteomic binding partners¹⁴³. Additionally, the results of selections against combinatorial peptide-phage display libraries are biased towards overly hydrophobic sequences¹⁴⁴.

To investigate potential biologically relevant binding partners, proteomic phage peptide display (ProP-PD) was developed^{44,139}. The approach offers the detection of proteomic binding partners, as the search space is constrained to sequences from a proteome of interest. In ProP-PD, peptides of the intrinsically disordered regions are tiled and displayed either on the p3 or p8 coat protein of the M13 bacteriophage. While there are only five copies of the minor coat protein p3, there are 2,700 copies of the major coat protein p8. Using p8-display enables the investigation of low-affinity, transient interactions through multivalency¹⁴⁵. Different libraries have been developed for the ProP-PD platform. The second-generation human disorderome library (HD2 library) is the largest, displaying all the human disordered regions tiled in ca. 1 million peptides using an overlap of 12 amino acids⁴⁴. Other libraries are Riboviria Viral Disorderome (RiboVD) library, displaying the disordered regions of RNA-virus proteins¹²⁴, phosphomimetic libraries that mimic phosphorylated peptides¹⁴⁶, or genetic variation HD2 library (GenVar_HD2) that scans the effect of disease mutations¹⁴⁷.

ProP-PD is easily scalable as a high-throughput method, allowing to pan hundreds of proteins in one experimental set-up. A typical ProP-PD workflow starts by cloning the oligo pool into the corresponding phagemid backbone and then creating the actual phage library by electroporation of the library into *Escherichia coli* pre-infected with M13KO7 helper phages (**Figure 6**). Once the library has been prepared, it is used in biopanning. Purified baits are immobilized and incubated with the phage library. Presenting the phage library to a negative control, such as glutathione transferase (GST), prior to adding it to the bait allows the elimination of non-specifically binding peptides. After the incubation with the baits, non-binding phages are washed off and binding phages are eluted using exponentially growing *Escherichia coli*. The addition of the M13KO7 helper phage enables the construction of full phage particles. After an overnight incubation, the culture is centrifuged and pasteurized at 60 °C to prepare the phage supernatant for another round of phage selections.

Binding phages are typically enriched after four rounds of selections using the HD2 library; other library designs may require fewer or additional rounds.

Through barcoding and pooling, next-generation sequencing of hundreds of bait-enriched phage pools is performed in one go. Bioinformatic analysis of the sequences parses the respective binding protein and generates the consensus motif. Results are ranked by confidence, from score 1 to 4, according to the number of quality criteria each peptide has. The criteria are (I) replicated peptide in between two or more replicates (II) overlap in displayed peptide, (III) contains the generated motif, (IV) high normalized peptide count. Medium confidence peptides match two or three criteria, while high-confidence peptides match all four peptides. The data offers insights at different *resolutions*. Peptide-bait interactions, which are the individual enriched peptides. Region bait interactions, which are single peptides or a combination of two or more overlapping peptides into one. The protein-protein level is the broadest view, as it collapses all peptides from the same hit protein to one value and reports on the potential PPIs.

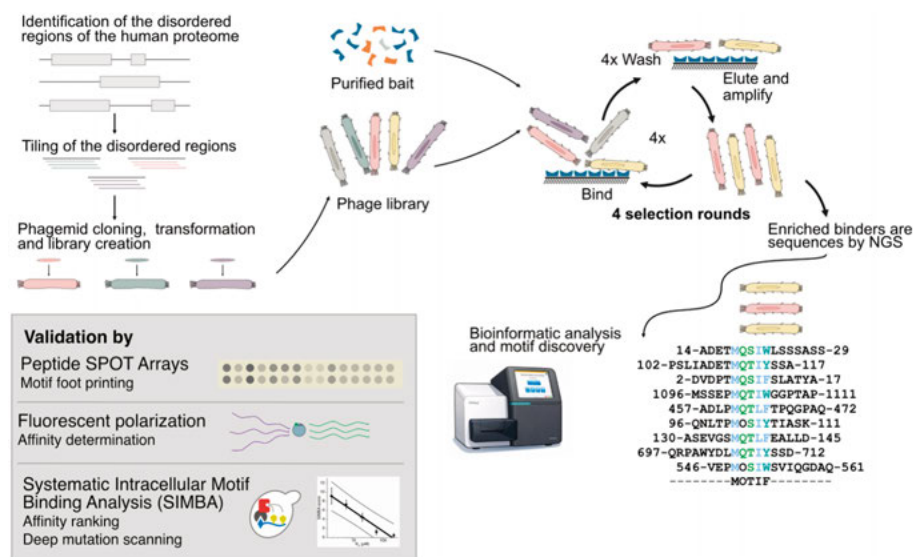


Figure 6: Overview of a ProP-PD approach and workflow and possible validations (adapted from Paper II). ProP-PD starts with the preparation of a suitable phage library. Here we tile all the disordered regions of the human proteome to 16 amino acid long peptides. The peptide coding sequences are cloned into the phagemid to produce the fusion protein of peptide and coat protein. The phage library is then presented to immobilized ligands. After four rounds of selections, which encompasses binding, washing, elution and phage amplification, a phage population is sent for next generation sequencing. The sequences are analysed and the motif bioinformatically determined⁴⁴. Necessary validations of the interactions are performed with SPOT arrays¹⁴⁹⁻¹⁵⁰, fluorescence polarization¹⁵¹ or SIMBA^{148,155}.

Validation of ProP–PD binders

ProP–PD is a powerful technique to identify peptide binders and SLiMs. However, although the confidence scoring allows us to infer the probability of those interactions, validation is needed (**Figure 6**). While SPOT arrays semi-quantitatively discriminate between binding and non-binding peptides, fluorescence polarization (FP) enables us to quantify the affinity between the binding domain and peptide. Systematic Intracellular Motif-Binding Analysis (SIMBA)¹⁴⁸ is an orthogonal method that allows deep mutational scanning of peptide-protein interactions and ranks the affinities of a multitude of peptides. Finally, co-immunoprecipitation (Co-IP) is a cell-based assay to validate the interaction in the full-length cellular context.

An accessible and scalable approach to validate the peptide-binding is the peptide **SPOT array** technology^{149,150}, and has been used to assess the peptide-binding within the E3 ligase machinery⁷⁸. Peptide arrays are created by solid phase peptide synthesis on a solid support, typically cellulose. Purified proteins are presented to these peptides. Binding is then detected by antibody recognition, typically with chemiluminescence as read out. Besides proteins, antibodies, serum and cell lysates can be tested for binding. In this thesis, we use peptides identified by ProP–PD, as well as ligands that we predicted with the SLiMSearch⁵⁸ algorithm. We also footprint the motifs through alanine-scanning, where we mutate each residue of the peptide to alanine, except for alanines which are mutated to glycine.

FP offers an immobilization-free method to detect interactions and quantify the affinities between proteins and their ligands¹⁵¹. Using fluorescently labeled peptides, such as with fluorescein isothiocyanate, the binding between peptide and peptide-binding domain can be followed. Polarized light is directed to the fluorescent probe, which — when unbound — rotates rapidly and emits a depolarized signal. Upon ligand-binding, the bound probe rotates slower, which leads to an increase in polarization, and this signal change upon binding can be used to measure the affinity of the interaction. To compare the affinity between the labeled peptide and other peptides, a labeled-peptide-protein-complex is formed and then displaced by titration of the unlabeled competitor peptide.

SIMBA is an intracellular yeast-based competitive growth assay to test the binding between peptide and peptide-binding regions¹⁴⁸. The core of the method is the pheromone-induced growth arrest in yeast, signaled through the mitogen-activated protein kinase (MAPK) dependent cascade. Cyclin-

dependent kinase (CDK) can block the growth arrest by phosphorylating a protein in the MAPK-dependent signaling pathway. The interaction between CDK and its target protein is mediated by a SLiM, which leads to a multisite phosphorylation of the target protein^{152,153}. The phosphorylations inhibit the interactions with the signaling protein, allowing the cells to grow in presence of the yeast mating pheromone that induced the growth arrest¹⁵⁴. Subbanna and colleagues engineered the system by fusing a foreign peptide-binding domain to native yeast cyclin Cln2. Consequentially, they replace the Cln2 docking motif in the CDK substrate, Ste20^{Ste5PM}, with the peptide of interest. The binding affinity between peptide and peptide-binding domain then correlates with the growth rate of the yeast clone¹⁵⁵. SIMBA allows deep mutational scanning of binding peptides for motif refinement and simultaneous affinity ranking of thousands of peptides.

Co-IP probes the biological context of the PPI of interest. It allows the detection of direct and indirect (bridged) interactions in the cell¹⁵⁶. Bait and/or prey proteins are fused to tags such as yellow fluorescent protein, green fluorescent protein, FLAG-tag, HA-tag, or Myc-tag¹⁵⁷. Both constructs are either transiently or stably expressed in mammalian cells. After lysing the cells, the fusion-proteins are purified through affinity chromatography, in small scale. The purification can be then analyzed by western blots, but also by AP-MS¹⁵⁸. The validation of SLiM based interactions through Co-IP becomes challenging when performing motif-mutant experiments because the interaction can be multivalently driven by multiple motif instances in the same binding partner. Mutating all the respective instances can be either technically challenging or expensive.

Global protein stability (GPS) profiling, briefly, is a fluorescence-activated cell sorting (FACS) approach to test peptides for degradation¹⁵⁹. The core of GPS profiling is a bicistronic lentiviral reporter, in which the peptide of interest is fused to GFP and cloned together with a second reporter like mCherry. Both fusion-GFP and mCherry are under the same translational control, expressed from the same mRNA, either by a self-cleaving peptide design or an internal ribosome entry site. If the peptide confers a degradation signal, the GFP-signal in the cell decreases in comparison to the mCherry signal¹⁵⁹. GPS profiling is a powerful approach to map degrons, and coupled with CRISPR screening it allows the identification of E3 ligase–degron pairs^{160,161}. GPS profiling has limitations: (i) it cannot reproduce conformational degrons, and (ii) the peptide-GFP fusion may interfere with the natural context of the degron — disrupting PTMs, altering E3 ligase binding or making lysines inaccessible⁷¹.

1.4. Current challenges in E3 ligase research

The field of E3 ligase research has several challenges:

1. *Determining which proteins exhibit which E3 ligase activity*

Large research efforts go into discovering and categorizing E3 ligases. There is little overlap between existing E3 ligase databases, recent unified efforts have begun to consolidate them into a unified E3-ome⁶⁶. Additionally, the traditional classification in HECT, RING and RBR E3 ligases is an oversimplification of the diverse protein landscape that exhibit E3 ligase function⁶⁶. Moreover, there are many E3 ligases that remain uncharacterized^{66,114}.

2. *Mapping the E3 ligase interactome*

Current methods do not capture the full biochemical span of E3 ligase interactions. E3 ligases detect their substrates through different processes. One of them being SLiM-based interactions which engage in low affinity and transient interactions with their ligands. Additionally, E3 ligase–substrates relationships are conditional. They depend on subcellular localization, tissue-specific expression, or condition-dependent expression¹¹⁴. Functionally, conformational degrons and PTM dependent degrons are difficult to capture and test for their degron potency⁷¹.

3. *Determining the ligand specificity determinants*

The E3 ligase ligand specificity determinants are flexible and SLiMs are often degenerate. Known binding SLiMs have been expanded and or redefined through biochemical studies^{10,71}. Same binding sites may accommodate different SLiM^{107,162,163}. Additionally, different domains within E3 ligases may recognize different ligands further increasing complexity¹⁶⁴.

4. *Limitations of high-throughput methods*

To accurately determine the E3 ligase ligand or substrate relationships, time-consuming, low-throughput studies have to be employed. Current high-throughput methods either are unable to accommodate the E3 ligase biochemistry, or alter the native recognition scenario⁷¹.

5. *Determining the ubiquitin chain dependent cellular outcomes*

E3 ligases use ubiquitination to signal to other proteins. The ubiquitin code is multilayered with different chain topologies and modifications, giving different signals. The signaling is further context dependent on the E2-E3 combination^{75,165}. Additional modifications and non-canonical ubiquitination sites increase the complexity¹⁶⁶. Overall, the systemic understanding of downstream events is challenging^{167,168}.

2. Aim of the thesis

The E3 ligase machinery is involved in regulating proteostasis and a multitude of cellular signaling pathways. A major unresolved question in the field concerns the specificity determinants that target the E3 ligase to substrates and interactors. As described, many of these interactions are expected to involve SLiM-mediated interactions. The overarching aim of the research presented in this thesis is thus to determine the SLiM-binding preferences of members of the E3 ligase machinery, with a particular emphasis on their modular peptide-binding domains.

Towards this aim I

- investigate the SLiM-based interactions of domains or full-length proteins of the E3 ligase machinery using ProP-PD (**Papers I-III**);
- explore the SLiM-based interactions of deubiquitinases (**Paper IV**);
- determine the specificity determinants within the ligand IDRs, including refining the SLiMs and their flanking regions (**Papers I-IV**);
- biochemically characterize the binding E3 ligase-ligand interfaces using biophysical methods, modelling, and mutational analysis (**Papers I-II**).

In doing so, I substantially expand the SLiM-based interactomes of the E3 ligases and associated proteins and uncover undescribed binding motifs. In a broader context of motif-biology, I contribute to charting the SLiM-based interactome of the cell.

3. Outline of the thesis

This thesis follows a thematic structure which I outline below.

To begin, I will elaborate on the expansion of the known ligand repertoire of E3 ligases and auxiliary proteins. I will illustrate this by using SH3 and WW domains (**Paper I**), and the Kelch domains (**Paper III**). I will also describe on the SLiM-binding of the ubiquitin-like domain of RAD23B (**Paper I**).

Next, I will present the discovery of new motifs. Starting with the B30.2/SPRY domains I explore the shared and domain-specific binding preferences of the novel binding motif of HERC1 alongside the known SLiMs of SPSB2 and SPSB4. I will then present the new asparagine-rich SLiM recognized by MIB-HERC2 domains (**Paper II**).

After exploring these new motifs, I present the redefinition of known motifs. BRCA1 and UHRF1 are known to bind to motifs that depend on PTMs. I discovered a phosphomimetic motif that binds to the BRCT domain of BRCA1. I establish that the TT domain of UHRF1 recognizes a SLiM that – beyond the known specificity to methylated lysine – binds unmodified ligands. Then I proceed to TRAF4 where I found a new SLiM that engages a previously reported binding site in the MATH domain. STUB1 is a well-established E3 ligase that recognizes substrates via its TPR domain in a chaperone-dependent manner. While chaperone-independent binding to the TPR has been reported, I present that the TPR domain of STUB1 can recognize an internal SLiM (**Paper II**).

I then illustrate the SLiM-mediated interplay between E3 ligases and DUBs. We were able to describe a SLiM binding to the CTD of RNF41 that deorphanizes the interaction between RNF41 and USP8. I also characterize the RNF41 binding site and the overlap in ligand repertoire between RNF41 and USP8 (**Paper II**). I will also further describe the impact of the flanking regions on the example of the zinc finger-ubiquitin binding protein (zf-UBP) domain of USP22, particularly peptide length (**Paper IV**).

I close by showing how the data can be used to characterize the relationships within E3 ligase–ligand network. I describe (i) efforts to determine degrons using GPS profiling, (ii) an approach to SLiM-based ligand prediction, (iii) how SLiMs can be used to explore multivalency in the E3 ligase–ligand binding (**Paper II**).

4. Findings

4.1. Proteome-wide discovery of ligands

This first section focuses on the expansion of the proteome-wide discovery of ligands. It elaborates on the peptide-binding of WW domains, SH3 domains, and Kelch domains. The section finishes with the expansion of the SLiM-mediated interactions on E3 ligase–auxiliary proteins such as RAD23B.

Proteome-wide discovery of SLiM-mediated interactions in the E3ligase machinery

In **Papers I–III**, I screened a total of 140 protein domains and selected full-length proteins of the E3 ligase machinery for SLiM-based interactions using the HD2 library, which tiles the IDRs of the human proteome. While ProP–PD by now is an established method for mapping motif-based PPIs with amino acid resolution⁴⁴, it had not previously been used to systematically screen for interactions of the E3 ligase machinery. The combined data resulted in 11,460 peptide-protein interactions translating into ca. 8,290 putative PPIs. This can be compared to the data set in the E3-atlas¹²⁷, which contains 12,380 interactions. In the following section I describe highlights from the findings starting from well-established SLiM-binding domains (e.g. WW and SH3 domains), followed by less explored domains of human E3 ligases.

Expanding the SLiM-based interactomes of SH3 and WW domains

In **Paper I**, we explore the SLiM binding of more than 800 protein domains, thereby vastly expanding the human SLiM-based interactome. Here, we focus on the SH3 domains from RING E3 ligases, and the WW domains from HECT E3 ligases respectively. Both SH3 and WW domains are known polyproline peptide-binding domains, and their consensus binding motifs have been defined in several studies over the last 30 years^{169–173}. The E3 ligases that harbour

these domains often have arrays of the domains, which offers possibilities for avidity effects^{170,174}.

SH3 domains typically bind to PxxP motifs¹⁷⁵, which have been further divided into two main classes depending on their orientation. Class I recognizes the [R/K]xxPxxP motif. Class II recognizes the PxxPx[R/K] motif^{176,177}. In addition, there are a number of different types of SH3 binding motifs. Here, I highlight the SH3 domains of the E3 ubiquitin-protein ligases SH3RF1 (POSH) and SH3RF3 (POSH3), which, in addition their E3 ligase function, serve as scaffolding platforms for the JNK signaling pathway¹⁷⁸. Our screens of 1st, 3rd, and 4th SH3 domains of SH3RF1 and 3rd SH3 domain of SH3RF3 yielded 944 peptide protein interactions and 649 corresponding PPIs. The consensus motif of these interactions from both SH3-domain-containing proteins was Px[ALPV]PxR (**Figure 7A**), as confirmed by alanine-scanning SPOT arrays (**Figure 7C**). Additionally, we were able to establish 370 new motif instances and determined the affinities for seven ligands, and found them to be in the 2–400 μ M range (**Figure 7D**), which is typical for SH3-mediated interactions. The identification of hundreds of new motif instances suggests that SH3RF1 and SH3RF3 SH3 domains function as interaction hubs, where ligand binding is likely governed by combinatorial motif recognition.

WW domains commonly recognize the PPxY motif²⁵. WW domains are part of the domain architecture of many E3 ligases such as NEDD4-like HECT-type E3 ligases¹⁷⁹. In total, I screened 17 WW domains of E3 ligases. The 1st WW domain of NEDD4L1 recognizes the SLiM [PL]PxY (**Figure 7B**). I could detect ca. 3,200 binding peptides that correspond to ca. 1,900 potential PPIs. While 220 PPIs of the data set have already been reported, we describe ca. 2,500 new motif instances. Motif co-occurrence analysis from the large-scale screen in **Paper I** further showed that PPxY motifs are among the most frequently co-occurring SLiMs. Notably, these motifs are often spaced at distances compatible with simultaneous engagement by tandem WW domains, likely supporting a model of multivalent binding. This is exemplified by the transmembrane protein PRR7 (proline-rich protein 7), which contains three closely spaced PPxY motifs, two of which are recognized by WW domains of HECT-type E3 ligases. Taken together, these findings indicate that PPxY motifs are not isolated recognition motifs but are often organized into multivalent assemblies that enable cooperative binding by WW domain arrays. This suggests that substrate recognition by NEDD4-family HECT-type E3 ligases is driven by motif density and spacing rather than by single high-affinity interactions, consistent with the fairly low affinities of isolated domains and motifs

In addition, many PPxY motifs appear to be subject to regulation by tyrosine phosphorylation, suggesting that their binding capacity can be dynamically modulated. Furthermore, PPxY motifs located in intracellular tails of ion channels are frequently affected by disease-associated mutations, which are likely to influence protein internalization, trafficking, and stability. Taken together, these findings indicate that PPxY motifs are not isolated recognition motifs but are often organized into multivalent assemblies that enable cooperative binding by WW domain arrays. This suggests that substrate recognition by NEDD4-family HECT-type E3 ligases is driven by motif density and spacing rather than by single high-affinity interactions, consistent with the fairly low affinities of isolated domains and motifs.

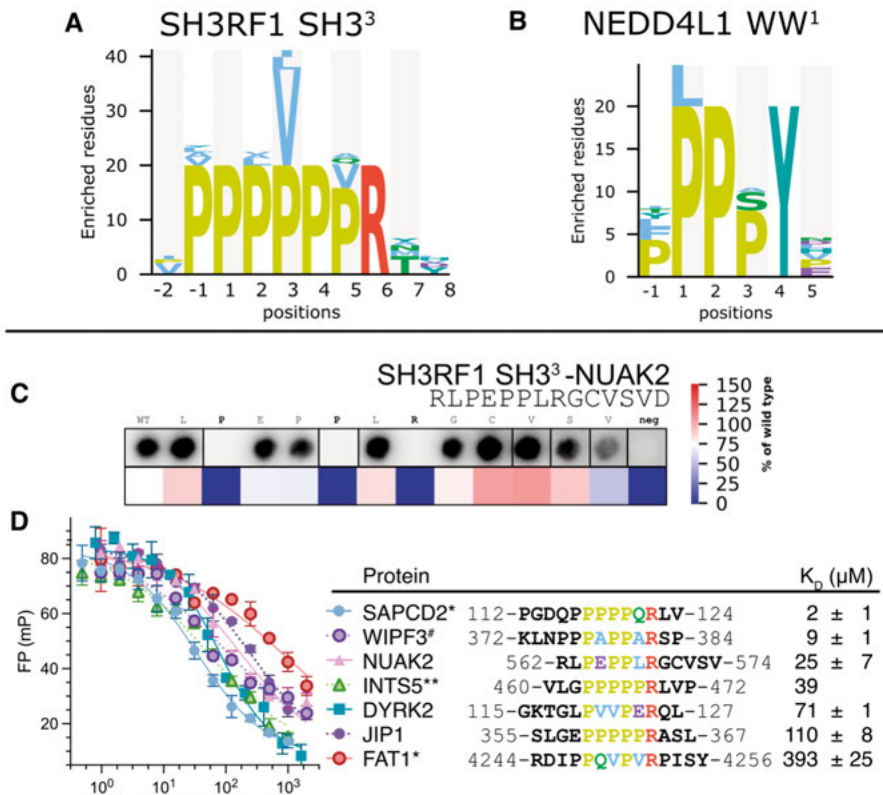


Figure 7: Peptide-binding results of the SH3 domain of SH3RF1 and WW domain of NEDD4L1. A) SLiM binding of the 3rd SH3 domain of SHRF1 (Px[ALPV]PxR). B) SLiM binding of the 1st WW domain of NEDD4L1 ([PL]PxY). C) Alanine scanning peptide SPOT array of the NUAK2 peptide binding to SHRF1 SH3-3. D) FP-based affinity determination of SH3RF1 ligands. # indicates reported interactions. * denotes only two replicates. ** denotes only 1 replicate. # denotes reported interactions. All other measurements are n = 3.

Expanding the SLiM-based interactomes of KELCH domains

BTB-Kelch proteins are adaptor proteins to the Cullin-RING E3 ligases complex 3 (CRL3). While the N-terminal BTB domain binds to the cullin-scaffold, the C-terminal Kelch domains recruit the substrates to the E3 ligase complex. Structurally, Kelch domains are composed of five to seven consecutive repeats that assemble into a β -propeller, adopting a donut-like structure. Loops that connect the inner and outer β -strands form the top face that accommodates the binding peptides^{180,181}. In **Paper III** we screened 18 BTB-Kelch domains for their peptide-binding using the HD2 library. This resulted in 572 peptide–domain interactions, corresponding to 498 candidate protein–protein interactions, of which only 15 had been previously reported. Notably, approximately 30% of the identified peptides contain residues subject to PTMs, suggesting that Kelch-mediated recognition is frequently regulated.

Based on the ligand peptides identified in our screens, we rediscovered the motifs of KLHL2, KLHL3, KLHL6, KLHL12, and KLHL20 and uncovered additional motif instances. The motifs rediscovered for KLHL3 and KLHL12 were similar to the motifs annotated in the ELM database²⁷. For example, KLHL3¹⁸² has been described to bind the motif ExEEExE[AV]DQH, whereas our data suggest that a simplified core motif, [APV]E[AGS], is sufficient for binding. Similarly, for KLHL12, the previously defined motif [PGA]PG[AG]PP¹⁸³ can be simplified to a core PG[GSA]P sequence that captures the essential determinants. These observations indicate that minimal motif cores are sufficient for engagement, while surrounding residues likely modulate binding. Our collaborators further investigated that the novel consensus motifs defined for KLHL21 and KLHL40, which take extended helical conformations upon binding. The results substantiate and expand the SLiM-based interaction space of the Kelch domain family and uncover unexpected diversity in Kelch domain binding, including the recognition of helical peptide motifs.

Exploration of peptide-binding for E3 ligase auxiliary proteins

We further expanded our investigation to auxiliary proteins to the E3 ligase system (**Paper I and II**). These proteins are included in the term E3 ligase machinery because they either modify substrates prior to binding E3 ligases or are involved after ubiquitination. The screens included: the arginyl-tRNA transferase ATE1 that prepares substrates for degradation in the ubiquitin pathway through the N-degron pathway¹⁸⁴, the chromo domain of the E3 SUMO-protein ligase CBX4, which interacts with the K9-trimethylated

histone H3¹⁸⁵ to recruit the E3 ligase complex PRC1 to regions of the chromatin^{186–188}, the melanoma-associated antigen (MAGE) domain of melanoma-associated antigen A4 (MAGEA4), the SOCS-BOX domain of Ras-related protein Rab-40A-like and the C-terminal domain REV1, and RAD23B. Here, we highlight the results on the ubiquitin like domain of RAD23B.

RAD23B is a multidomain protein that is involved in nucleotide excision repair (NER), but more importantly, functions as a ubiquitin shuttle factor. In NER, RAD23B is responsible for protecting the trimeric complex XPC–HR23B–Centrin-2 (CEN2) from proteasomal degradation¹⁸⁹. In addition to the aforementioned function, RAD23B binds to polyubiquitinated substrates and delivers them to the proteasome¹⁹⁰. The ubiquitin-like (UBL) domain of RAD23B recruits its ligands through hydrophobic residues, using a defined motif. It also contacts lysine 48 of ubiquitin, thereby blocking K48 chain elongation¹⁹¹. Because of its shuttling role, we investigated RAD23B peptide-binding capacity in **Paper I**. Our screen enriched 24 medium- to high-confidence peptides, translating to 19 putative PPIs. Among the identified PPIs, we found previously reported interactors: the deubiquitinase USP25^{192,193} and the proteasome regulatory subunit PSMD4 (26S proteasome non-ATPase regulatory subunit 4)^{194,195}. The enriched peptides also allowed us to confirm the consensus motif DxxxAIxLSL (**Figure 8A**). Using the HUWE1¹³⁷¹⁻¹³⁸²

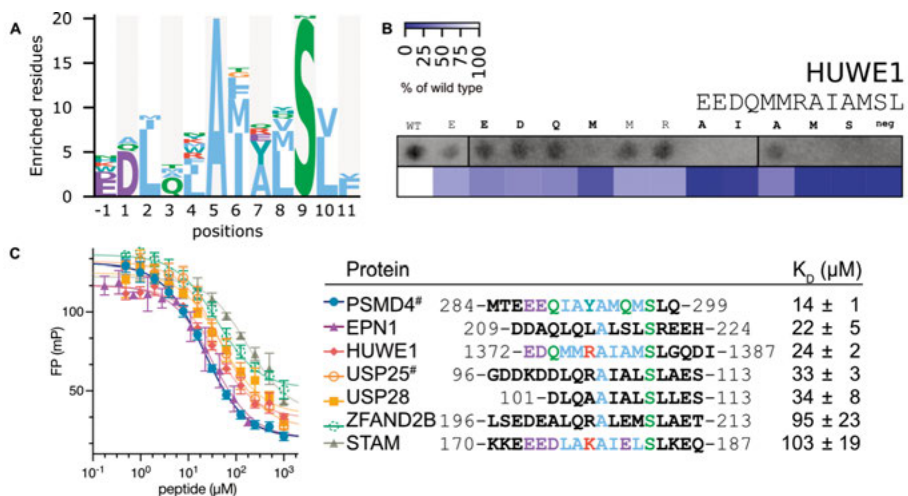


Figure 8: Overview of the data on the UBL domain of RAD23B. A) Consensus motif based on the ProP–PD derived peptides. B) Alanine-scanning SPOT array to footprint the motif using the HUWE1¹³⁷¹⁻¹³⁸² peptide. C) FP-based affinity determination of selected ligands from the ProP-PD selections. # indicates reported interactions. (n=3)

peptide we footprinted the motif (**Figure 8B**). Affinities were found to be in the 14–100 μM range (**Figure 8C**). PSMD4 is the interactor with the highest affinity, and also previously reported^{194,196}. From the analysis of the motif flanking regions, we further see a preference for acid residues on the C-terminal side.

Taken together, these findings refine the binding motif of the RAD23B ubiquitin-like domain binding motif, and show that it engages in interactions with multiple components of the proteasome and ubiquitin system. SLiM-based interactions are thus at play in different parts of the machinery.

4.2. Description of novel SLiMs of SPRY domains and MIB-HERC2 domains

B30.2/SPRY domains are known peptide binders¹⁹⁷, and the motifs of SPSB2 and SPSB4 — SPRY domain-containing SOCS box protein — have been described¹⁹⁸. In the following we illustrate how the B30.2/SPRY domain of HERC1 recognizes a SLiM that resembles the SLiM of SPSB2 and SPSB4. Additionally, we disclose the SLiM of the MIB-HERC2 domains of MIB1, MIB2, HECTD1 and HERC1. These domains show a strong preference to asparagine-rich motifs, which is reflected in their SLiM.

Shared and domain-specific binding preferences of B30.2/SPRY domains

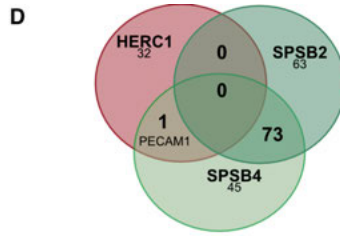
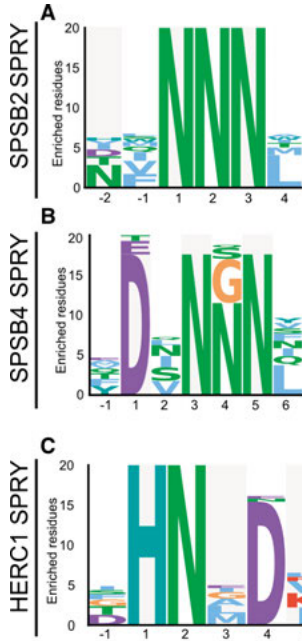
There are about 100 human B30.2/SPRY containing proteins that are associated with different pathways, including innate and adaptive immunity¹⁹⁹. As part of different E3 ligases they contribute to substrate binding. The domain is especially common in TRIM E3 ligases, where around half of the family members contain a B30.2/SPRY domain¹⁹⁹, but they are also found in other E3 ligases, such as the SPRY domain-containing SOCS box proteins (SPSBs)²⁰⁰, and the CTLH-adaptor proteins Ran-binding protein 9 (RANBP9)²⁰¹ and Ran-binding protein 10 (RANBP10)²⁰². B30.2/SPRY domains have a fairly low sequence similarity except for the regions that fold into the β -sheet motif²⁰⁰. The β -sheet motif offers a rigid, pre-formed pocket on the concave side of the sheet^{200,203}. This binding-site topology has been compared to the complementarity-determining regions (CDRs) of antibodies⁹⁵. The *Drosophila* ortholog of the human SPSB protein family, Gustavus, has been shown to bind to the DINNN-containing peptide from the protein VASA with nanomolar affinity²⁰³. Nevertheless, for most B30.2/SPRY domains it has not been established if they bind to peptides, and if so, what the SLiMs are^{200,204}.

In **Paper II** we systematically explored the peptide-binding ability of 20 B30.2/SPRY domains. The collection comprises domains from 12 TRIM domains, three CUL5 substrate adaptors (SPSB2, SPSB3 and SPSB4), as well as the CTLH adaptors RANBP9 and RANBP10, and three RING-finger-type E3 ligases. Most of them did not enrich high-confidence ligands, which may suggest that their main function is not SLiM binding. However, the domains SPSB2, SPSB4, and HERC1 enriched high-confidence peptides that could be used to determine a motif (**Figure 9A-C**). For the SPSB2 and SPSB4 SPRY

domains, the selections enriched asparagine-rich peptides, as expected from the previous knowledge on the VASA peptide. Additionally, the SPSB4 SPRY domain required an aspartic acid N-terminal to the central asparagine, just as described in the Gustavus-VASA interaction. Besides 17 previously reported PPI, we expanded the SPSB-binding repertoire with 320 potential interactions for SPSB2, SPSB3 and SPSB4 together. The ligand set of SPSB2 and SPSB4 SPRY overlap with 73 peptides, implying a redundancy in the substrate recognition of the CUL5 complex (**Figure 9D**). In contrast, HERC1 and SPSB4 share only one ligand, PECAM1 (**Figure 9D**). Using FP, we confirmed nanomolar affinities of endogenous ligands of SPSB4 (Figure 9G).

In our screen of B30.2/SPRY domains, the SPRY domain of HERC1 emerged as a previously unrecognized peptide-binding module. HERC1 selected a distinct set of ligands with limited overlap with the SPSB2/4 domains, consistent with its low sequence identity with other domains. Modeling and mutational analysis indicated that HERC1 engages its ligands through the same general surface region as SPSB2/4 proteins (**Figure 9E-F**), but employs different binding loops and sequence determinants. Affinity measurements showed that HERC1 binds its ligands with micromolar affinity, with a peptide derived from DLEU1 displaying the highest affinity ($K_D = 14 \mu\text{M}$), and additional ligands from PECAM1, HECW1, and ZNF507 binding in the range of 59–190 μM (**Figure 9H**). Notably, the HERC1 motif includes a histidine residue that appears to interact with a negatively charged patch of the domain. Consistent with this, binding was found to be pH-dependent, with an affinity optimum around pH 7.2, suggesting that ligand recognition may be modulated by local conditions. Finally, motif-guided prediction followed by SPOT validation confirmed additional ligands, demonstrating that the defined HERC1 SPRY motif has predictive power.

Together, these findings identify HERC1 as a motif-binding E3 ligase and show that SPRY domains have diversified ligand recognition while maintaining binding of asparagine-containing peptides. This expands the scope of SLiM-based recognition within the E3 ligase machinery and highlights how specificity can arise from variations within a shared fold, which is similar to findings on other large domain families such as WW, SH3 and PDZ domains.



E

```

HERC1 2035 -----GFDPEKAQ CCLVE -NQQLTHGSG- GK 2059
SPSB4 28 -RPARLDQLDMPAAGLAVQLRHWNPEDRSLNVFVKDDRLTFHRHPVAQST 79
SPSB2 22 LSCPEGLELLSAPPFDLGAQRHGWNPKDCSENI EVKEGGLYFERRPVAQST 74
HERC1 2060 GYGLASTGVTSQCYQKFIYVKNCGNECTCVGYSRMPVHDF - - - - - NHRIT 2106
SPSB4 80 DGIRKVGHARGLHAWQINWPAPQR GTHAVVGVATARAPLHSVGYTALVGS 131
SPSB2 75 DGARGKRYSRGLHAEISWPLEQR GTHAVVGVATAPLQTDHYALLGSN 126
HERC1 2107 SDMLYRAVSGNLYNG - - - - - EQTLT - - - - - SSFTQGFITCVLDNEARI 2149
SPSB4 132 AESNGWDLGRSLYHGGKQKQPGVAYPAFLGDEAFAL PDSLLVLDNDEGL 183
SPSB2 127 SESNGWDIGRGLYHQS KGGAPQYPAGTQGELEVP - - - - - ERLLVLDMEEGIL 177
HERC1 2150 SFGKNGEPEKLFEDVDAELVPCVMFSSNPGEKVKICDMQ - - - - - MR 2193
SPSB4 184 SFIVDQYLGVAERGLKGRKLVVSA - V - - - - - WGHCEVTMRYINGLDE - - - - - 228
SPSB2 178 GYAIQTYLGFARGLKGRKLVVSA - V - - - - - WQQCVIRIYRGERGSHH - - - - - 224

```

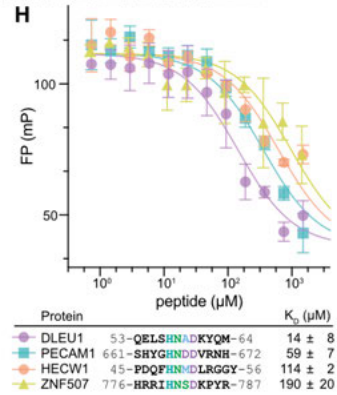
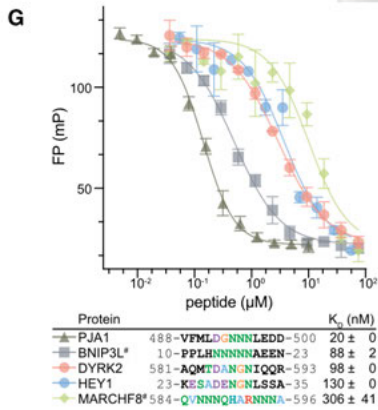
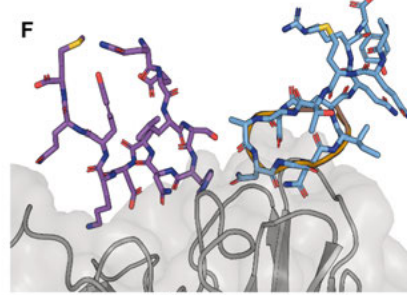


Figure 9: Overview of the SLiM binding of the SPRY domains in Paper II. A) Consensus motif of SPSB2: NNN. B) Consensus motif of SPSB2: DxN[NGS]N. C) Consensus motif of HERC1: HN_x[DE]. D) Overlap of SPRY domain binding peptides in the selections of HERC1, SPSB2, and SPSB4. E) Multiple sequence alignment of the SPRY domains of HERC1, SPSB4 and SPSB2, with the indication of the secondary structure elements of HERC1. Taupe arrows indicate beta-strands and grey-purple rectangles indicate loops or turns. Highlighted residues are conserved among the domains. Red dots indicate binding sites as identified by ColabFold or AlphaFold predictions. Underlined residues indicate mutants for binding site determination. Red curvise residues indicate binding regions. F) Alignment of the AlphaFold3 prediction of HEY1 (olive green) and the SPRY domain of SPSB2 with the VASA peptide (gold, PDB: 3F2O), the iNOS peptide (brown, PDB: 6KEY) and the ColabFold prediction of HERC1-SPRY domain binding to DLEU1⁵³⁻⁶⁴ peptide (purple). E) FP-based affinity determination of SPSB4 ligands. F) FP-based affinity determination of HERC1 ligands. Affinity measurements are in triplicates. # denotes reported interactions. Adapted from Paper II

The MIB-HERC2 domains also bind asparagine-rich sequences

In **Paper II** we further explored the peptide-binding capacity of the MIB-HERC2 domains. MIB-HERC2 domains appear in four E3 ligases: MIB1, MIB2, HECTD1, HERC2. Structurally, the MIB domains assume a fold similar to the SH3 domains. Both domains adopt a five-stranded antiparallel twisted β -sheet fold^{163,205} (**Figure 10A**). MIB1 and MIB2 are RING E3 ligases of similar domain architecture. Both contain tandem MIB-HERC2 domains, followed by the REP (tandem MIB-repeat) domain¹⁶³, 9 ankyrin repeats and the RING domains. HECTD1 and HERC2 belong to the class of HECT E3 ligases. Their immense size – HECTD1 has ca. 2,610 amino acids and HERC2 ca. 4,830 amino acids — offers space for many more domains (**Figure 10B**). And all domains share sequence identity, particularly at one of the binding sites (HECTD1 Gln1291, Figure 10C).

MIB proteins were first discovered in 2003 by Itoh and colleagues as an E3 ligase ubiquitinating Delta-like and Jagged, which are ligands to the Notch signaling pathway²⁰⁶. MIB1 ubiquitination leads to ligand internalization. Previous work on MIB1 illustrated a bipartite substrate recognition mechanism using the tandem MIB-HERC2 domain¹⁶³. It has also been speculated that HECTD1 and HERC2 use their other domains for a similar binding mechanism²⁰⁵. MacMillan and colleagues established that N-box-containing peptides are able to bind to MIB-HERC2 domains. Our selections enriched asparagine-containing peptides that resulted into SLiM, similar to, but distinct from the motifs found for the SPRY domains. When modelling the found peptides

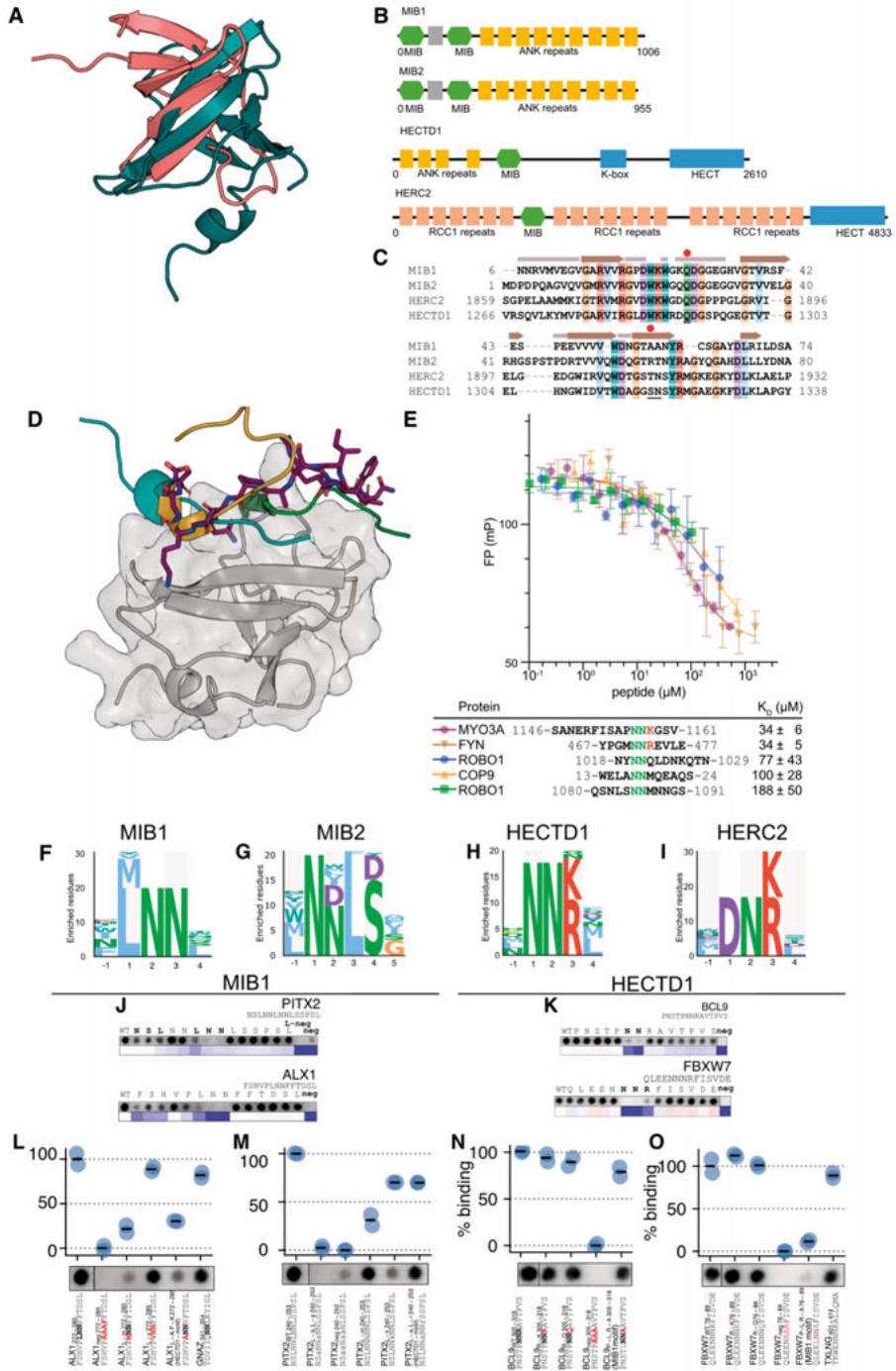


Figure 10: MIB-HERC2 domain contain E3 ligases. A) Structural alignment of a crystal structure of the 1st SH3 domain of SH3RF1 (red, pdb: 7NZC) and the AlphaFold3 model of MIB-HERC2 domain of HECTD1 (teal). B) Domain architecture of the MIB-HERC2 domain containing E3 ligases in this study. C) Multiple sequence alignment of the MIB-HERC2 domains. Taupe arrows indicate beta-strands and grey-purple rectangles indicate loops or turns. Highlighted residues are conserved among the domains. Red dots indicate binding sites as identified by or AlphaFold predictions. D) Alignment of AlphaFold3 model of ProP-PD peptides to MIB-HERC2 domain of HECTD1. The MIB1 peptide ALX1 is in green. The MIB2 peptide AKAP2 is in cyan. The HECT1 peptide MYO3A is in purple. The HERC2 peptide ZNF148 is in gold. E) FP-based affinity determination of HECTD1 ligands (n = 3). Followed by MIB-HERC2 binding motifs of MIB1 (F), MIB2 (G), HECTD1 (H) and HERC2 (I). Motif footprinting of (J) MIB1 binding motif and (k) HECDD1 binding motif. Peptide SPOT array analysis of MIB1 binding peptides (J-M) and HECTD1 binding peptide (N-O). Adapted from Paper II.

to their respective domains, we found that the models converge in the same site (**Figure 10D**). FP-based affinity measurements allowed us to determine the affinity range of the HECTD1 peptides to be between 34 and 188 μ M (**Figure 10E**). Subsequent mutational analysis of the binding site confirmed the importance of Gln1291 in peptide-binding. While the interactions appear anchored by the asparagine, specificity is provided by the neighboring residues that also form part of the motif (Figure 10F-I). While alanine-scanning SPOT arrays confirmed the importance of the key residues (**Figure 10J-K**), differentially mutating the distinctive residues of the MIB1 motif (the aliphatic residue) and HECTD1 motif (basic residues) shows only a limited disruptive effect on the interaction (**Figure 10L-O**). This observation underlines the importance of the asparagine in the motif and the physicochemical affinity contribution of the flanking region.

In summary, I defined a conserved MIB-HERC2 peptide-binding surface that recognizes an anchor asparagine within a distinct sequence context. Whether and how the other domains of the MIB-HERC2-containing E3 ligases contribute to ligand binding remain to be elucidated.

4.3. Redefining binding motifs

Our investigation broadened and redefined binding specificity determinants and established new motifs. The following work ranges from PTM-dependent interactions to novel SLiMs recognized by established motif binders.

BRCA1 is able to bind non-phosphorylated ligands and its hetero-dimerization partner BARD1 binds peptides too

BRCA1 is a heterodimeric RING E3 ligase that dimerizes with BRCA1-associated RING domain protein 1 (BARD1). While both E3 ligases are not robustly catalytically active, the BRCA1-BARD1 E3 ligase complex is functional²⁰⁷. The complex is reported to monoubiquitinate histone H2A at DNA damage sites²⁰⁸. It is involved in homologous recombination²⁰⁹, DNA checkpoint signaling^{210,211}, chromatin recruitment through the IDR of the BRCA1-BARD1 complex²¹². The complex is known to create K6²¹³ and K48²¹⁴ ubiquitin chains, which are here non-degradative.

In 2003, Manke *et al.* and Yu *et al.* describe how the BRCT domain recognizes the phosphoserine motif pSxxF^{215,216}. The phosphoserine is recognized by a positively charged pocket, while phenylalanine inserts into a hydrophobic cleft. BARD1 on the other hand recognizes unmethylated histone H4 with its ankyrin repeats²¹⁷. The detection of unmethylated H4 is important to target the complex to newly deposited histones, to use the sister chromatid as template for homologous recombination. Additionally, the same domain recognizes ubiquitinated histone H2A at lysine 15. Ubiquitination signals the DNA damage response signaling cascade. Through their heterodimeric-assembly, both E3 ligases harness multivalency to target histones for the DNA damage response and DNA repair functionality⁹⁸.

We included both the BRCT domain of BRCA1 and the ankyrin repeat containing domain of BARD1 in our investigations. Our binding peptides to BRCA1 refined the phosphomimetic motif ExxF (**Figure 11A**, where a glutamic acid mimics the negative charge of phosphorylated serine. Through SPOT arrays we can confirm the binding of the BRCT domain to the phosphomimetic motif (**Figure 11B**). Unsurprisingly, the binding affinity of the phosphorylated ACACA peptide showed a nanomolar affinity, while a non-modified peptide of AIRE⁵¹⁸⁻⁵²⁹ has an estimated binding affinity of around 100 μ M (**Figure 11C**). The ankyrin repeat containing domain of BARD1

enriched instead peptides displaying a putative consensus motif LxER (**Figure 11D**). The binding and the motif determinants of this putative motif needs to be further assessed. While we were unable to perform follow up experiments for the BARD1 ankyrin repeat binding due to time constraints, it is interesting to note that we found two overlapping putative PPIs between BARD1 and BRCA1 from our ProP-PD data: Zinc finger protein GLI2 (GLI2) and DAZ-associated protein 2 (DAZAP2) (**Figure 11E**). GLI2 is a zinc-finger transcription factor for the hedgehog pathway²¹⁸. DAZAP2 is a proline rich adaptor in Wnt signaling and the transforming growth factor (TGF)- β signaling through modulating transcription factor 4's (TCF4) DNA-binding affinity²¹⁹. When we expand the potential binding repertoire by searching for SLiM containing proteins with the discovered motifs of BRCA1 and BARD1 we find 30 overlapping proteins (**Figure 11F**).

In conclusion, the discovered motifs offer a new perspective on the BRCA1-BARD1 complex recognition²¹². The new motifs could broaden the functionality of the complex upon DNA damage or through the histone binding. The complex may send the new ligands for degradation if their concentrations in the nucleus exceed a threshold. Ubiquitination of transcription factors potentially increase their chromatin residence time as it has been shown in different cases^{220,221}.

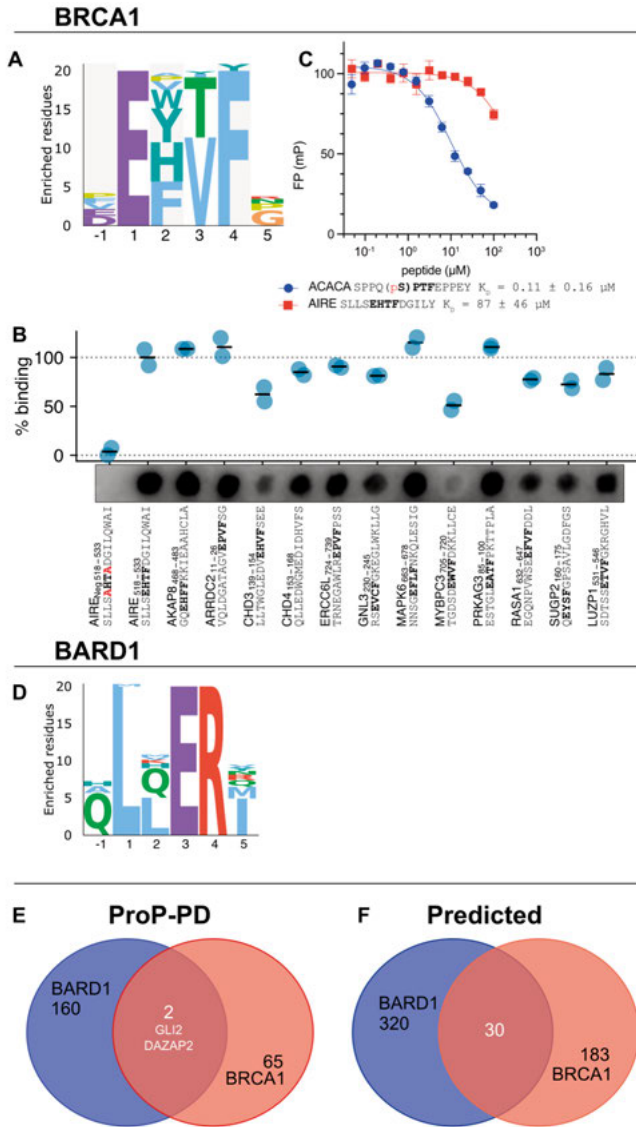


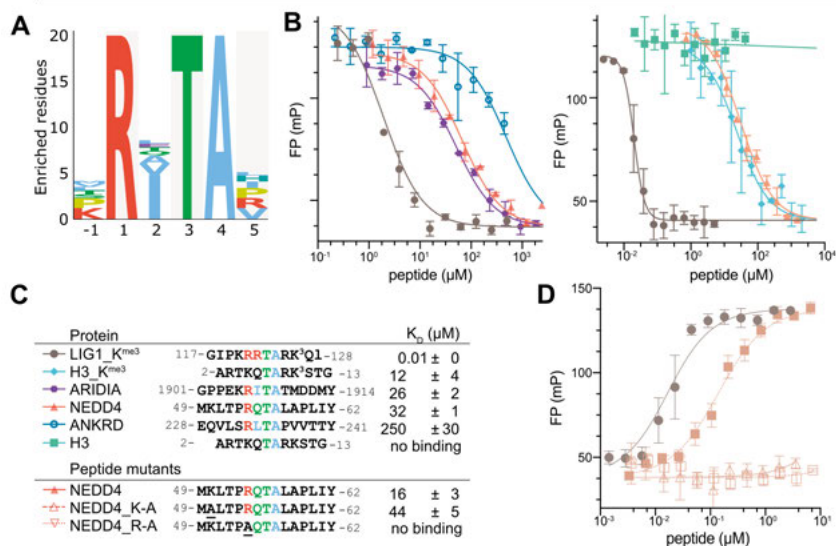
Figure 11: Data overview of the heterodimeric E3 ligases BRCA1 and BARD1. A) PSSM representation of the SLiM binding to the BRCT domain of BRCA1 (ExxF). B) Peptide SPOT array of found ProP-PD peptides to the BRCT domain of BRCA1. C) FP based affinity measurements of a phosphorylated ligand ACACA and the ProP-PD ligand AIRE. D) A) PSSM representation of the putative binding to the ANK domain of BARD1 (LxER). E) Venn diagram of ProP-PD ligands binding to the domains of interest of BARD1 (ANK) and BRCA1 (BRCT). F) Venn diagram of predicted ligands using SLiMsearch4 and the motifs of BARD1 and BRCA1. Affinity measurements are in triplicates. Adapted from Paper II.

UHRF1, a methyl-lysine reader and E3 ligase, recognizes a non-methylated SLiM

Just like the BARD1-BRCA1-complex, UHRF1 is a multidomain E3 ligase that is responsible for the maintenance of DNA methylation. Its TT domain is reported to recognize methylated substrates such as the Histone H3 (H3, methylated on lysine 9). To harness multivalency, the plant homeodomain (PH domain) of UHRF1 further contacts the unmodified N-terminal residues Ala1, Arg2 and Lys4²²². This interaction is required for the epigenetic inheritance of DNA methylation²²³. UHRF1 is also reported to bind to the methylated DNA ligase LIG1 (LIG1) with nanomolar affinity²²⁴.

Our data of the TT domain showed binding determinants beyond methylation. 46 medium to high confidence peptides were enriched from the ProP–PD selections, many of which containing a RxTA consensus motif (**Figure 12A**). Comparison with the LIG1 peptide reveals that it in fact contains the motif upstream to the methylation site, in contrast to the H3²⁻¹³ peptide that only partially contains the motif (**Figure 12C**). FP-based affinity determination showed that the affinity of methylated LIG1 and H3 to be in the nM and low μ M range respectively. Having the methylation in addition to the motif makes a 1,200-fold affinity difference for the LIG1¹¹⁷⁻¹²⁸ peptide according to the literature²²⁴, while we find that the unmethylated H3²⁻¹³ peptide is not able to bind. In contrast, some of the unmethylated RxTA containing ligands such as the ARID1A peptide bind UHRF1 with comparable affinities to the methylated H3 peptide (K_D values in the range of 26 to 250 μ M). The arginine at the first position of the motif is essential for binding, as mutating it disrupts the interactions (**Figure 12B, D**). These results suggest that UHRF1 recognition is not exclusively dependent on methylation. The RxTA motif contributed to binding, either independently or in combination with methylation, which may have potential implications for how the proteins engages both canonical and non-canonical targets in the context of DNA methylation maintenance.

UHRF1



TRAF4

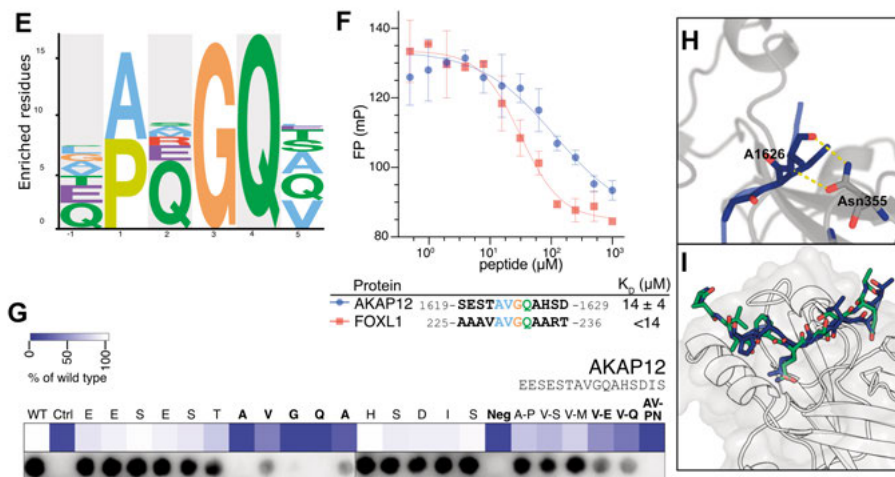


Figure 12: Overview of the Data of UHRF1 and TRAF4. A) PSSM representation of the SLiM binding the TT domain of UHRF1 (RxTA). B) FP-based affinity determination of binding ligands and investigation of the effect of histone H3 methylation, data in C. C) Affinity data of FP experiments from B and D. D) FP-based analysis of the affinity contribution of key residues and flanking residues. E) PSSM representation of the SLiM of the MATH domain of TRAF4 ([AP]xGQ). F) FP-based affinity determination of the ProP-PD ligands AKAP12 and FOXL1 (n=2). G) Motif footprinting with a SPOT array of the AKAP12 peptide (n=2). I) AlphaFold3 predictions of the AKAP12 peptide (blue) binding to TRAF4 (grey). The backbone of A1626 binds with hydrogen bonds to the sidechain of Asn355 E) AlphaFold3 Structural alignment of the Aflafold3 prediction of AKAP8L¹⁶²⁰⁻¹⁶³⁵ (blue) binding to the MATH domain of TRAF4 and the crystal structure of EGFR¹¹⁹⁸⁻¹²⁰⁷ (green, pdb: 9OLB). Adapted from Paper II.

The SLiM-binder TRAF4 recognizes an undiscovered SLiM

TRAF4, like other TRAF-family members, has been reported to be involved in immune signaling by inhibiting NOD2-induced NF- κ B activation and restricting IL-17-mediated signaling^{225,226}. Additionally, TRAF4 is also involved in development and morphogenesis, cell polarization - together with the HECT E3 ligase SMURF1²²⁷ - and in the integrity of tight junctions²²⁸. Other TRAF family members (TRAF2, TRAF3, and TRAF5) have been reported to recognize a major (P/S/A/T)x(Q/E)E motif and a minor (PxQxxD) motif²²⁹. In contrast, in 2017 Kim and colleagues described that the TRAF4 MATH domain binds to a RLxA based on co-crystallization of a GPIIb β 1¹⁷⁷⁻¹⁸¹ peptide²³⁰.

We included the MATH domain of TRAF4 in our analysis. Our ProP–PD selections enriched 276 peptides, potentially binding to 206 binding partners. From the peptides we were able to generate the motif [AP]xGQ (**Figure 12E**). Alanine SPOT array scanning confirmed the influence of the motif key residues (**Figure 12G**). FP-based affinity measurement showed an affinity in the 10 μ M range (**Figure 12F**). Our AlphaFold predictions indicate that the domain residues Asn355 mediate the interaction and that the modelled AKAP12¹⁶¹⁹⁻¹⁶²⁹ peptide binds into the same binding pocket as the published GPIIb β 1¹⁷⁷⁻¹⁸¹ peptide (**Figure 12H-I**). Site-directed mutagenesis of the binding domain verified that our [AP]xGQ containing peptides binds in the same binding pocket. In conclusion, we find that the general consensus binding motif of TRAF4 MATH domain is a [AP]xGQ motif, which binds to the same pocket as the previously described ligand.

Another E3 ligase with well-established ligand binding is the E3 ubiquitin-protein ligase CHIP (STUB1). Binding to chaperones it sends misfolded proteins for degradation. In the next section I describe how using ProP–PD we refine the STUB1 tetratricopeptide repeat (TPR)-domain binding motif.

Beyond its HSP-client nature, STUB1 binds also to internal SLiM

STUB1 is a prominent E3 ligase because of its interactions with HSP70²³¹ and HSP90²³². Dimeric STUB1 recognizes through its TPR domains the C-terminal tail of HSP70 or HSP90 respectively^{231,233}. The hydrophobic pocket of the TPR domain accommodates the C-terminal (I/M)EEVD motif present in HSP70 and HSP90^{232,234}. While binding to the chaperones, STUB1 does not recognize substrates itself, but uses the chaperones as substrate adaptors^{233,235}. In this configuration it degrades misfolded proteins that the chaperones can no longer refold. STUB1 has also been reported to bind to latent C-terminals revealed by caspase activity²³⁶. Additionally, the TPR-domains has been reported to bind a macrocyclic peptide with a CCIWC motif²³⁷. It has also been suggested to bind tryptophan containing alpha-helical peptides²³⁸.

Through our screen, we uncover that the TPR domain of STUB1 is able to bind to internal proteomic peptides through its reported binding pocket. Analysis of the bound peptide defined the consensus motif [SD]IW[PV] (**Figure 13A**). We confirmed the binding site through site-directed mutagenesis of residues suggested to be important for binding based on AlphaFold3 predictions, foot-printed the motif through alanine-scanning SPOT arrays, and determined that the affinities of the interactions are ranging from 6 to 222 μ M by FP (**Figure 13B,E**).

To assess these interactions in a cellular context, co-immunoprecipitation experiments were performed using GFP-tagged ligands. FLAG-STUB1 co-precipitated three out of four target proteins (RNA-binding E3 ubiquitin-protein ligase MEX3C (MEX3C), transcription factor E2 α (TCF3), and hematopoietically-expressed homeobox protein HHEX (HHEX), but not with homeobox protein OTX2 (OTX2). Mutation of motif residues (e.g., TCF3 Y340A and HHEX W127A) reduced, binding confirming the functional relevance of the identified motifs (**Figure 13F-G**).

Together, my findings demonstrate that the STUB1 TPR domain is not restricted to canonical C-terminal EEVD motifs, but can also recognize internal proteomic sequences through a hydrophobic core motif. This expands the substrate-recognition landscape of STUB1 and suggests that it may directly engage a broader set of proteins beyond chaperone-mediated interactions.

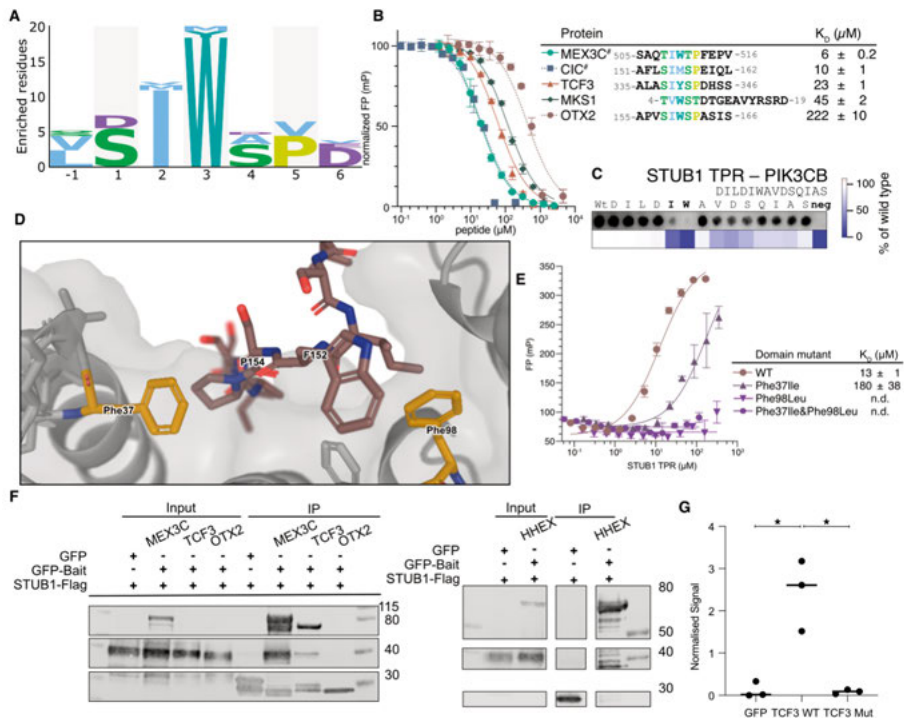


Figure 13: Overview of the data on the TPR domain of STUB1 A) PSSM representation of the STUB1 TPR binding motif ([SD]IW[PV]). B) FP-based affinity determination of Pro-PD ligands. C) Alanine-scanning SPOT array for motif printing using the PIK3CB peptide (n=2). D) AlphaFold3 model of the OTX2 peptide (brown) and the binding domain residues of STUB1 (yellow) as used for domain mutants in E. E) Affinity determination of STUB1 domain mutants. F) Western blot of GFP-based co-immunoprecipitation, capturing full-length GFP-tagged ligands and precipitating full-length FLAG-tagged STUB1 (n=3). # indicates reported interactions. G) TCF3-mutant binding to STUB1 in the experiment of F with their GFP-normalized signal. Significance was determined by a two tailed t-test. Adapted from Paper II

4.4. Interplay between SLiMs binding to E3 ligases and DUBs

The interplay of E3 ligases and DUBs has been established^{79,239}. However, the exact mechanism and determinants of these interactions remain elusive. In this section we illustrate how SLiMs mediate the interaction between writer (E3 ligase) and eraser (DUB).

Defining a RNF41 binding motif and uncovering competition between the E3 ligase and the DUB USP8

The RING-type E3 ligase RNF41 was described in 2002 to degrade the receptor tyrosine-protein kinase erbB-3 (ERBB3) in the proteasome^{240,241}. The US8-interacting domain of RNF41 binds to the intracellular domain of ERBB3 to recruit it for the catalytic N-terminal RING domain²⁴². RNF41 has been many reported functions: it influences MAPK²⁴³, PI3K (phosphatidylinositol 3-kinase)/Akt, and phospholipase C gamma (PLC γ)^{244,245} signaling pathways through degrading ERBB3. It interacts with Tyrosine-protein kinase JAK2-associated cytokine-receptor complexes²⁴⁶, and regulate the innate immunity and Toll-Like Receptor signaling through ubiquitinating MyD88²⁴⁷. RNF41 is also involved in antigen cross-representation through ubiquitinating the extracellular domains of the dendritic cell receptor CLEC9A (C-type lectin domain family 9 member A), endosomal sorting through the endosomal sorting complexes required for transport-0 (ESCRT-0) complex^{248,249}. It induces apoptosis through degradation of E3 ligase BIRC6 (Baculoviral IAP Repeat-Containing protein 6) that inhibits apoptosis²⁵⁰, triggers mitophagy by Parkin degradation^{251,252}.

Most important to RNF41 function are its interactions with the deubiquitinase USP8 (Ubiquitin carboxyl-terminal hydrolase). RNF41 is inherently unstable because its RING domain autoubiquitinates²⁴⁰. Catalytically inactive USP8 leads to inactivation of RNF41²⁵³. The MIT domain is a well-known peptide binder⁷⁸, which binds to the USP8-interacting motif of RNF41²⁵⁴. Additional contacts between USP8 and RNF41 are made through RNF41's C-terminal (CTD) USP8-interacting domain²⁴².

The rhodanese domain of USP8 was recently found to bind to peptide ligands⁷⁸. Intriguingly, the RNF41 CTD was also recently shown to bind to an IDR, although the motif determinants were not disclosed²⁵⁵. These reports led us to examine the peptide-binding capacity of the CTD of RNF41 in ProP-

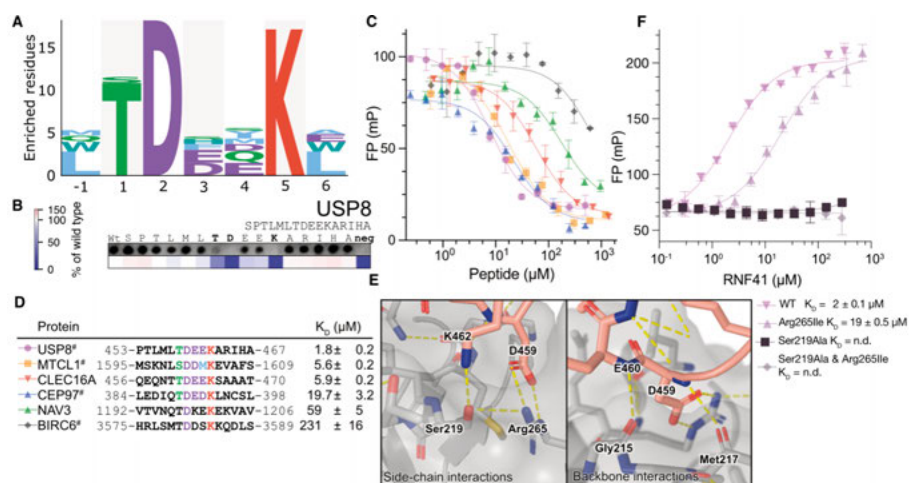


Figure 14: Overview of the data on RNF41. A) PSSM representation of the RNF41 CTD binding motif ([ST]DxxK). C) FP-based affinity determination of ProP-PD ligands. K_D values are in D. F) FP measurements to qualify the effect of domain mutants to the binding of USP8⁴⁵³⁻⁴⁶⁷. Affinity measurements are in triplicates. D) Overview of ProP-PD ligands and their affinities from C. E) AlphaFold3 models of the RNF41-USB8 interaction to identify the interaction driving residues. While domain residues Ser219 and Arg265 engage with the sidechains of the peptide, the residues Gly215 and Met217 engage with backbone interactions. Affinity measurements are in triplicates. # indicates reported interactions. Adapted from Paper II

PD. Using the HD2 library we enriched 25 medium to high confidence peptides that correspond to 22 putative PPIs. Eight of these PPI have been reported before, notably the BIRC6²⁵⁰ and USP8²⁵³, but also the interaction with HOMER2²⁵⁶⁻²⁵⁸ and KDM3B^{256,257}. The peptides defined the binding motif [ST]DxxK (**Figure 14A**). We proceeded to footprinting the motif through alanine-scanning SPOT arrays (Figure 14B), and recorded affinities through FP (Figure 14C). The alanine-scanning SPOT arrays confirmed the key residues of the motif. The affinities of the peptides to RNF41 range from ca. 2 μM (USP8⁴⁵³⁻⁴⁶⁷) to 231 μM (BIRC6³⁵⁷⁵⁻³⁵⁸⁹). The AlphaFold 3 models suggest the recognition of side-chain interactions between Ser219–K462 and Arg265–D459 (**Figure 14E**, the three-letter-codes denote the domain residues, while the one-letter-codes represent the peptide). Through site-directed mutagenesis we could determine that Ser219 is critical for the binding affinity, since mutating this residue completely disrupts the binding. We additionally predicted 40 RNF41-binding peptides using the defined consensus motif and tested them with SPOT arrays. 16 of these peptides bound to RNF41, including a second BIRC6²⁹⁵¹⁻²⁹⁶² peptide that was not discovered through ProP–PD.

Intrigued by the partial similarities between the recently defined USP8 rhodanese binding motif and the newly defined RNF41 CTD binding motif, we tested the cross-reactivity of the interactions, and found a partial overlap. This suggests that the E3 ligase and the DUB may compete for subsets of binding partners, thus adding complexity to the interplay between the two enzymes and highlighting the importance of knowing the SLiMs of both E3 ligases and DUBs.

DUBs also use SLiMs for ligand recognition

In **Paper IV** we further investigated the SLiM-binding of the auxiliary domains of DUBs and found that several of them bind to SLiMs. Here, we focus on one case, the zinc finger-ubiquitin binding protein (zf-UBP) domain of USP22. USP22 promotes cancer progression^{259,260}, and interacts with histone substrates as a component of the SAGA (Spt-Ada-Gcn5 acetyltransferase) complex, as well as other non-histone substrates^{261–263}. Screening the zf-UBP domain of USP22 in ProP–PD revealed five peptide ligands. The ATXN7L1^{36–52} peptide was the most relevant as it provides the binding site of the deubiquitination module of the SAGA complex^{264,265}. The binding motif was footprinted with SPOT arrays to be FLGxPWxxW. We determined the affinities using FP. The ProP–PD-sourced peptide ATXN7L1^{36–52} displayed an affinity of 11 μ M. From AlphaFold3 modelling we noted that a longer peptide resulted in a higher confidence model. Consistently, we found that expanding the peptide by five amino acids on the C-terminal side (ATXN7L1^{36–57}) increased binding affinity to 0.10 μ M. The same drastic increase could be observed for the ATXN7L2 peptide (**Figure 15**). Here, the extended flanking region has a large impact on binding affinity, besides the motif residues.

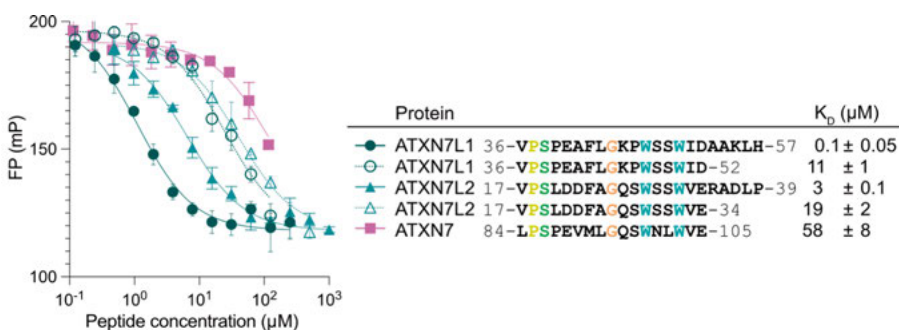


Figure 15: FP-based affinity determination of the ATXN7^{84–105}, ATXN7L1^{36–52} and ATXN7L2^{17–34} peptides. To investigate the impact of the flanking regions to the apparent affinity we extended the peptides of ATXN7L1 and ATXN7L2 C-terminally.

4.5. Additional analyses

We can use the discovered SLiM and interactions to further explore the E3 ligase machinery. Three approaches are illustrated below.

A global protein stability profiling approach did not explain the degron nature of our ligands

Global protein stability (GPS) profiling is an established method to determine the degradation potential^{159–161,266,267}. Naturally, we tried to determine the degron propensity of our dataset. For that, I, with the support of my collaborators, created a lentiviral library, consisting of 3,829 peptides including their internal replicates - using different codons - as well as known degrons (positive controls), and mutated and reversed sequences of the same degrons (negative controls). The library was N-terminally fused to eGFP, connected by a 14-amino-acid long linker. The peptide-eGFP fusion protein was in turn fused to mCherry, linked by the self-cleaving P2A peptide sequence. HEK cells were infected with the lentiviral library and then sorted in eight bins using FACS, based on the fluorescent ratio of eGFP and mCherry in each cell. A low ratio indicates the instability of the peptide-eGFP fusion protein, whereas a high ratio indicates the stability of the fusion protein.

After analyzing the peptide sequences, we could determine the distribution within the bins and calculate the protein stability index (PSI). The data showed that there were no significant differences between the positive controls (known degrons) and the negative controls (mutated degrons) as indicated by the PSI. Our approach to determine the degron potential of our peptides failed to generate conclusive data. This underlines the challenges in discovering degrons.

Determining SLiMs enables ligand prediction

A major strength of SLiM-based interaction analysis is that the identification of a binding motif enables prediction of additional ligands. Because SLiMs are autonomous sequence patterns that independently mediate interaction, they can be used to predict additional ligands. We defined the SLiMs based on ProP-PD data, alanine-scanning peptide arrays and SIMBA-based deep mutational scanning of the ligands of E3 ligases RNF31, RNF41, STUB1, and SPSB4. Having defined the consensus motifs, we searched the IDRs using SLiMSearch4⁵⁸. We then narrowed down our selection using predicted ligands from UbiBrowser2²⁶⁸. Focusing on putative degrons, we overlaid the SLiMSearch4 results with a proteome-wide map of short-lived proteins²⁶⁹.

Having determined potential binding peptides, we had them synthesized on cellulose membranes with the SPOT technology and then tested them for binding.

Using this strategy, the ligand repertoires of HERC1, RNF41, STUB1, and TRAF4 were expanded beyond the initial phage-display selections (**Paper II**). Of 80 tested interactions 34 were validated. This combined computational and experimental workflow demonstrates that the generated SLiM information is predictive, and provides a path toward systematic expansion of E3 ligase interaction networks. These results underline the balance between key residues and flanking regions. The key residues drive the interaction, and the flanking region modulates the peptide acceptance.

SLiM scanning of identified ligands enables us to explore multivalency that may increase the binding potential

E3 ligases may use multivalency to amplify the apparent affinity. As described in the introduction, avidity arises when a single interacting protein engages two or more recognition modules of the same E3 ligase. E3 ligases can exploit these effects via dimerization (STUB1²⁷⁰, MIB1²⁰⁵), repetition of binding domains (WW-domain-containing proteins²⁷¹), or employment of multiple substrate recognition domains (MIB1^{163,205}). By analyzing our ligands of **Paper II** we identify that many ligands of the MATH and SPRY domains contain multiple binding regions within the same ligand (**Figure 16**), supporting a potential multivalent mode of binding.

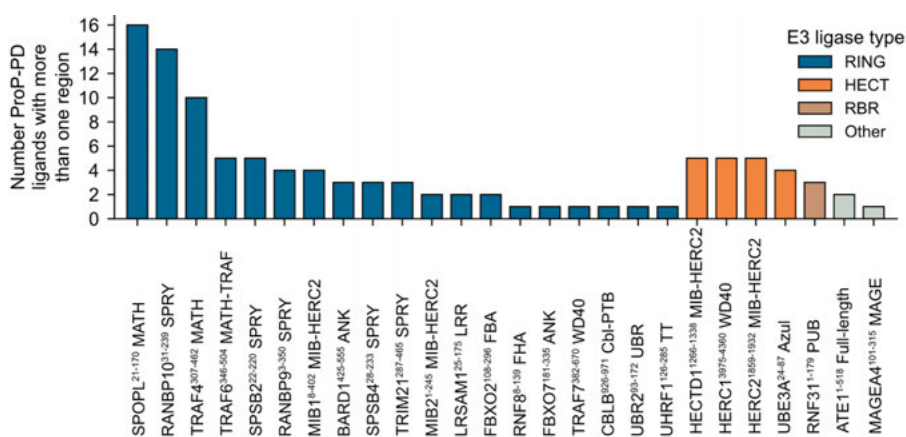


Figure 16: Overview of peptide-binding regions that map to the same ligand from E3 ligase machinery proteins of Paper II. This illustrates possible multivalent binding or avidity effects.

We can expand this analysis by scanning the identified ligands for additional instances of the hit protein, which offers further insights into potential multivalency related effects. Using the SLiMs of MIB1, MIB2, HECTD1, HERC2, RNF31, RNF41, STUB1, TRAF4 and UHRF1 we detect multiple motif instances in the sequences of ProP–PD determined ligands and in the disordered regions (**Figure 17**). For each of these cases the multivalency might increase the local concentration of the ligand. For STUB1 the search was performed with the SIMBA-based deep mutational scanning motif [MVLI][MFYW]x[PVICT]. Particularly the STUB1 ligand TCF3 has several predicted motif instances in its disordered regions. However, we need to keep in mind that these are only predicted motifs, and experimental biochemical assessment of sequence-based predictions remains essential.

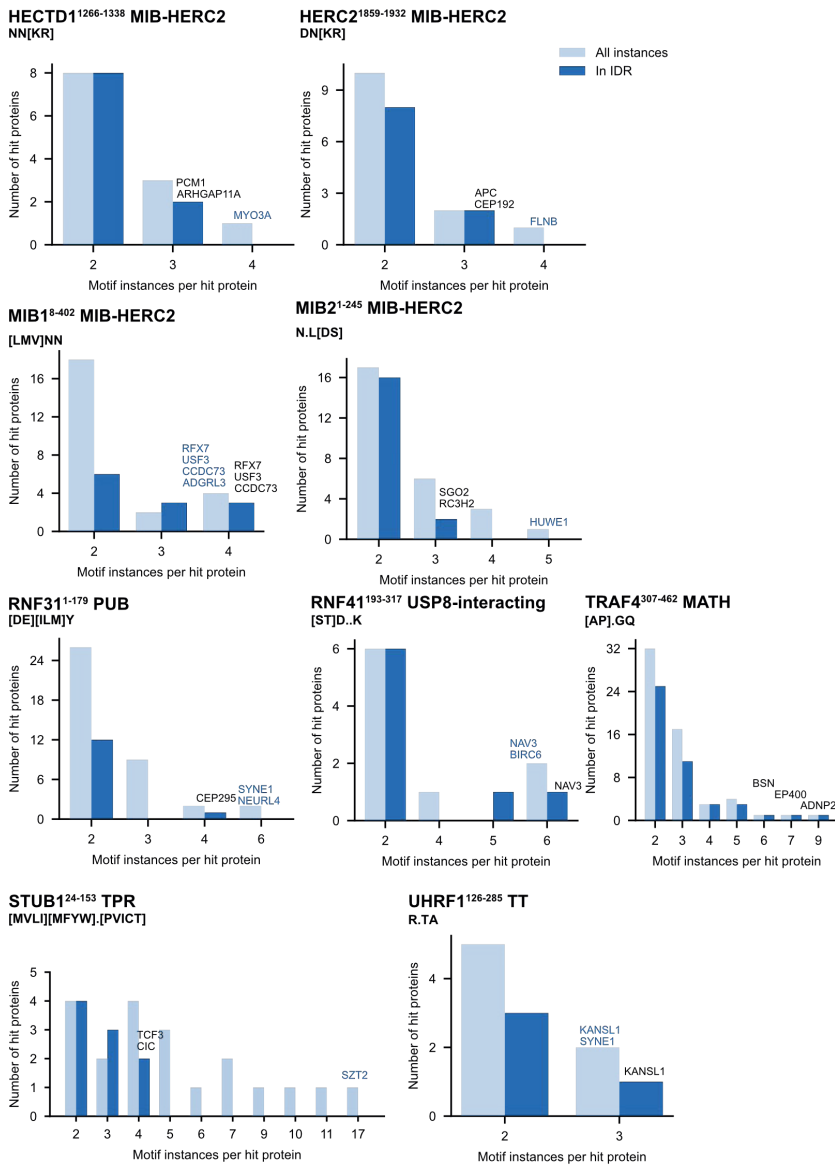


Figure 17: Analysis of possible binding motifs beyond the ligands found in ProP-PD predicted using the discovered motifs and screening of the full protein sequences of found ligands for additional motif instances and filtering for disordered regions (IUpred score larger than 0.4). The black labels indicate the gene names of proteins with the highest number of instances in the disordered regions per E3 ligase. Except for TRAF4 where the gene names of the three highest classes are indicated. The blue label indicates gene names with the highest number of instances in the whole ligand sequence – ordered and disordered. For STUB1 the deep mutational scanning motif [MVL][MFYW]_x[PVICT] was used for the analysis.

5. Summary of findings

Unlocking the hidden codes of SLiM-based interactions of the E3 ligase machinery is one of the keys to understanding this system. Different high-throughput approaches have been employed by various groups to identify ligands or potential substrates^{160,266,267}. Nevertheless, identifying the interaction partners and substrates of E3 ligases remains a major challenge. There are many reasons for this: (i) many interactions are transient and of low affinity^{98,272,273}, (ii) substrate interactions are often conditional and context dependent, employing intricate control mechanisms^{71,114}, (iii) degrons may be cryptic or conformationally masked^{71,98}, (iv) the E3 ligase machinery is redundant with many substrates targeted by multiple E3 ligases^{114,272}, (v) the motif sequence is variable in degrons^{71,98}, and (vi) there may be cooperativity between E3 ligases that target different substrates^{98,114}. Moreover, the dynamic level of substrates which are targeted for degradation, are further modified by DUBs, which complicates detection^{68,71,272}. Some of these challenges are specific for the E3 ligase system. Others are general challenges related to the nature of the SLiM-based interactions.

Here, I took a reductionistic approach to search for interactions, using a SLiM-centered strategy. I employed ProP-PD to search for E3 ligase binding peptides (**Papers I, II, and III**).

In **Paper I**, which constitutes a massive screen for SLiM-based interactions, I contributed to the screening of ligands for WW and SH3 domains from E3 ligases. Although the core motifs for these domains were already known, our work substantially expanded their SLiM-based interactomes. Additionally, I established a SLiM of the UBL domain of the RAD23B. The motif broadens the functionality of RAD23B beyond its ubiquitin-chain shuttle for the proteasome.

In **Paper II**, which represents the bulk of the thesis research, I performed a large screen for SLiM-based interactions of a variety of domains, and some full-length proteins, of the E3 ligase machinery. In this study, we uncovered several novel binding motifs of E3 ligases, and provide deeper insights into

the interactions in terms of affinities and specificity determinants using a combination of assays, including FP-monitored affinity measurements, peptide SPOT arrays, and mutational analysis of peptides and domains. Paired with AlphaFold modelling we provide structural insights into the interactions. We define for example the SLiM-recognition determinants of the CTD of RNF41. We illustrate that the novel SLiM detected by the SPRY domain of HERC1 has a preference for asparagine-containing peptides and uses a similar binding site as the SPRY domains of SPSB2 and SPSB4. We also establish that MIB-HERC2 domains of the MIB1, MIB2, HECTD1 and HERC2, recognize asparagine-rich peptide motifs. We further find that domains such as the BRCT domain of BRCA1 and the TT domain of UHRF1 recognize also peptides in a PTM-independent manner. We surprisingly find a novel SLiM that binds to the MATH domain of TRAF4. The most striking finding of **Paper II** is that SLiM-based peptide recognition is pervasive across the E3 ligase machinery, where it adds an additional layer of specificity to ligand recognition.

In **Paper III** I expand the ligand repertoire of Kelch domain containing E3 ligase adaptor proteins. Additionally, I show that fewer determinants — key motif residues — are required to mediate binding.

In **Paper IV** I investigate the contribution of the flanking regions by testing, peptides of different length, to the affinity of DUB ligand-binding.

The study showed that the ubiquitin system contains a broader and more diverse network of peptide-recognition modules than previously appreciated. Notably, the SLiM-based interactions found for the E3 ligases do not appear to correspond directly to degron function as tested in GPS profiling, potentially reflecting that many ubiquitination events serve regulatory roles other than proteasomal degradation. Taken together, these studies provide a framework for understanding how SLiM-based interactions contribute to specificity, connectivity and regulation across the ubiquitin system.

6. Conclusions and future perspective

In the present research, I show that motif-based interactions are widespread across the E3 ligase machinery and its associated factors. ProP–PD enabled large-scale identification of SLiM-mediated interactions, while biochemical validation and modeling provide information on the underlying specificity determinants. Across multiple domains and domain classes, the data revealed both new motifs and revised versions of previously established motifs, often showing that known consensus sequences are less restrictive than assumed. These findings support a view of E3 ligase specificity in which core motif residues, flanking-region chemistry, and domain context together determine ligand recognition.

The affinity measurements quantify the residue preferences of the binding domains and rank the potential ligands, placing them into biochemical context. Our AlphaFold-supported structural models predict the likely binding sites. Through the annotation of PTMs and SNPs within the ProP–PD, we can estimate the context-dependence of the interaction, and the effect of disease on the E3 ligase-ligand interaction.

My research largely expanded the motif-based interactions across all proteins and in particular the E3 ligase machinery. The present data expands the known E3 ligase interactome by ca. 70 %, including 14 novel or redefined binding motifs. We performed FP-based affinity measurements for 9 different E3 ligase domains. In detail I describe the SLiM binding of peptides of all known MIB-HERC2-domain-containing proteins, the SPRY domain of HERC1, the TPR domain of STUB1 and the CTD of RNF41.

My research illustrates not only the powerful application of ProP–PD for discovering peptide binders and defining SLiM, but also the broader value of SLiM-based biochemistry in mapping PPIs. My research gives insights into the balance between key residues, wildcard residues, and the flanking region of SLiMs. My data will enable the field to study the E3 ligase machinery in more detail. The described E3 ligase-ligand relationship offers an entry for

future functional investigations. Determining binding peptides will aid structural analysis through co-crystallizations.

E3 ligases are involved in many different pathologies, ranging from cancer, neurodegenerative diseases to bacterial and viral infections. The defined binding motifs may in the future be leveraged for the development of E3-ligase-targeting inhibitors and drugs such as PROTACs or molecular glues. The combined knowledge about the binding specificity determinants and the affinities will aid the process²⁷⁴.

In conclusion, this thesis demonstrates that the “hidden code” of the ubiquitin system is encoded not only in ubiquitin chains, but also in the SLiM-based interactions that govern complex assembly and substrate recognition.

7. Populärvetenskaplig sammanfattning

Proteiner utför de funktioner som håller våra celler och kroppar vid liv. Vårt DNA kodar för uppskattningsvis 20 000 till 30 000 proteiner. Vid varje given tidpunkt uppgår antalet proteiner i en given kroppscell till cirka 10 000 till 12 000. Varje protein har en bestämd funktion, såsom att driva metabolismen, den aeroba andningen eller att upprätthålla DNA:ts integritet. Liksom vi själva åldras även proteiner. Under ett proteins livstid kan produktions- och veckningsfel uppstå. Slutligen kan även cellens yttre eller inre förhållanden förändras på ett sådant sätt att vissa proteiner inte längre behövs. Då kommer de så kallade E3 ligaserna in i bilden.

E3 ligaser är en typ av proteiner med uppskattningsvis 600 till 1 000 medlemmar. Denna proteinfamilj är ansvarig för att märka in proteiner för nedbrytning av ubiquitin-proteasomsystemet. För detta ändamål har det i naturen utvecklats ett markeringssystem: E1 enzymet aktiverar ubiquitin för att vidarebefordra det till det E2 konjugerande enzymet. E3 ligaser tar sedan antingen upp det aktiverade ubiquitinnet eller binder till E2 enzymet. Det väsentliga steget är nu substratigenkänningen. E3 ligaser måste tillförlitligt identifiera substratet och märka det med ubiquitin. Hur E3 ligaser känner igen sina substrat och andra bindningspartners är vad jag har undersökt under min forskarutbildning.

Ubiquitin är i sig själv ett litet protein. Endast 76 aminosyror långt bestämmer det inte bara över nedbrytningen av andra proteiner, utan har även andra signalfunktioner. Det kan bland annat avgöra lokaliseringen av proteiner, till exempel i cellkärnan eller vid cellmembranet, men även signalera DNA-skador. Dessa signaler kommuniceras genom olika kedjeformationer av ubiquitin eller ubiquitinliknande proteiner. Genom de olika format som ubiquitinkedjor tar, och de effekter de har på olika signalvägar utvidgas E3 ligasernas funktionella roller mångfaldigt. Avgörande för detta är såväl placeringen av ubiquitinnet på substratproteinet som hur den resulterande ubiquitinkedjan byggs upp. Medan den katalytiska vägen för E3-ligaser är väl beskriven saknas kunskap om vilka E3 ligaser som binder till vad och om ubiquitinerings

olika cellulära följdverkningar. Den centrala frågan som vi fokuserar på här är hur olika E3 ligaserna känner igen sina substrat.

E3 ligaser känner ofta igen sina substrat bland annat genom så kallade Short Linear Motifs (SLiM, eller SLiMar i plural), vilket kan översättas som "korta linjära bindningsmotiv". För detta utnyttjar E3 ligaserna sin modulära uppbyggnad. Proteiner är i grunden modulärt organiserade, där varje modul – ett veckat domän – har en tertiär veckning som ger domänet en bestämd funktion. Vissa domäner utför katalytiska funktioner, medan andra utför sysslan att specifikt känna igen proteiner, delvis genom de redan nämnda bindningsmotiven. Ett SLiM är ett 3–10 aminosyror långt avsnitt av ett protein som ofta finns i de delar av proteiner som är ostrukturerade. Detta innebär att dessa delar av proteinerna inte antar någon veckad struktur. En SLiM känns ofta igen av sina bindningspartners utifrån 3–4 aminosyror som fungerar som nycklar för den molekylära igenkänningen.

Utöver E3 ligaser finns det ett stort antal andra proteiner som bidrar till att reglera vilka proteiner som ubiquitineras och vilka konsekvenser detta i sin tur får. I det följande stycket använder jag E3-ligas-maskineriet som ett samlingsbegrepp. Ytterligare komponenter är exempelvis de enzymer som bryter ner ubiquitin-kedjan (deubiquitinaser) och proteiner såsom RAD23B, vilket vidarebefordrar ubiquitinerade proteiner till proteasomen.

Under samlingsbegreppet determinanter för substratspecificitet har jag undersökt hur E3 ligaser känner igen sina substrat. För detta har jag byggt upp en samling med över 130 proteiner (Manuskript I–III). Tillsammans med mina kollegor och studenter har jag testat över 140 proteindomäner med avseende på deras SLiM-bindande förmåga. För detta har jag använt fagpresentation av peptider från vårt proteom (Proteomic Peptide Phage Display; ProP–PD), en metod som har utvecklats av den grupp där jag utförde mitt avhandlingsarbete. Fagpresentation är en metod som först utvecklades för att utveckla antikroppar. Genom vidareutveckling kan fagpresentation även användas för att undersöka humana protein-proteininteraktioner. I detta fall undersöks vilket protein som binder vilken SLiM. Med hjälp av denna metod har jag kartlagt ett stort antal nya interaktioner.

Vidare har vi karakteriserat de upptäckta interaktionerna och SLiMarna genom biokemiska metoder. Vi har undersökt de molekylära determinanter som avgör vilket protein som binder vilken SLiM. Tack vare utvecklingen av artificiell intelligens kunde vi datorbaserat modellera interaktionerna med programmet AlphaFold och validera dem experimentellt. Genom

fluorescensanisotropi, alltså genom utnyttjande av de fluorescerande ljusstrålarnas egenskaper, har vi bestämt interaktionsstyrkorna mellan peptiderna och deras domäner. Vidare har vi undersökt domänernas SLiM-bindning genom peptid SPOT-array försök. Peptid SPOT-arrayer är membran på vilka peptider syntetiserats, som sedan används för att testa olika interaktioner.

Slutligen har vi undersökt interaktionerna i cellulärt sammanhang med fullängdsproteiner. Vi har också utifrån de funna SLiMarna bioinformatiskt sökt efter och validerat ytterligare interaktioner. I mitt fjärde manuskript utvidgar vi forskningen till att undersöka de biokemiska grunderna för SLiM-bindning med deubiquitinaser. Utöver detta har jag undersökt inverkan av peptidlängden och av aminosyror utanför nyckelaminosyror.

Sammantaget kunde vi identifiera 11 460 peptider som interagerar med E3-ligaskomplexet. Dessa interaktioner representerar ungefär 8 290 potentiella protein-protein-interaktioner. Detta utvidgar de för närvarande 12 380 kända E3 ligas interaktionerna med 67 %. Vi karakteriserade nya SLiM för bland annat MIB-domän-E3-ligaserna MIB1, MIB2, HECTD1 och HERC2. Jag kunde även definiera en ny SLiM som binder E3-ligaset HERC1. För E3 ligaserna RNF41 och STUB1 kunde vi bestämma SLiM som förmedlar bindningen till redan kända interaktionspartner.

Mitt arbete utgör en del av de vetenskapliga ansträngningarna att karakterisera E3-ligassystemet. Resultaten ger ett väsentligt bidrag till utvidgningen av det kända E3-ligas-interaktionsnätverket och tydliggör SLiMarnas centrala betydelse för substratigenkänning. Därmed utgör data en värdefull resurs som framtida studier kan falla tillbaka på för att motivera och driva vidare undersökningar. Därutöver har resultaten potentiell relevans för utvecklingen av terapeutiska strategier, i synnerhet med avseende på modulering av ubiquitinberoende signalvägar.

8. Populärwissenschaftliche Zusammenfassung

Proteine führen Funktionen aus, die unsere Zellen und Körper am Leben erhalten. Unsere DNA kodiert schätzungsweise 20.000 bis 30.000 Proteine. Zu jedem Zeitpunkt beträgt die Anzahl der Proteine in jeder beliebigen Körperzelle etwa 10.000 bis 12.000. Jedes Protein hat eine bestimmte Funktion, wie zum Beispiel das Betreiben des Metabolismus, der aeroben Atmung oder der Wahrung der DNA-Integrität. Wie wir altern auch Proteine. In der Lebenszeit von Proteinen können Produktions- und Faltungsfehler auftreten. Schließlich können sich auch die äußeren oder inneren Bedingungen der Zelle so ändern, dass gewisse Proteine nicht mehr benötigt werden.

Dann kommen die sogenannten E3-Ligasen ins Spiel. E3-Ligasen sind eine Proteinfamilie mit schätzungsweise 600 bis 1.000 Mitgliedern. Typischerweise ist diese Proteinfamilie dafür zuständig, Proteine innerhalb des Ubiquitin-Proteasom-Systems abzubauen. Zu diesem Zweck hat die Natur ein auf einem Stoffwechselweg basierendes Markierungssystem entwickelt: Das E1-Enzym aktiviert Ubiquitin, um es an das E2-konjugierende Enzym weiterzugeben. Die E3-Ligase nimmt dann entweder das aktivierte Ubiquitin auf oder bindet an das E2-Enzym. Der wesentliche Schritt ist nun die Substraterkennung. Die E3-Ligase muss zuverlässig das Substrat identifizieren und mit Ubiquitin markieren.

Ubiquitin selbst ist auch ein Protein. Nur 76 Aminosäuren lang, bestimmt es nicht nur über den Abbau anderer Proteine, sondern hat auch weitere Signalwirkungen. Es kann unter anderem über die Lokalisation von Proteinen entscheiden, zum Beispiel im Zellkern oder an der Zellmembran, aber auch DNA-Schäden melden. Diese Signale werden durch verschiedene Kettenformationen von Ubiquitin oder ubiquitinähnlichen Proteinen kommuniziert. Durch die unterschiedlichen Formate der Ubiquitinketten und die entsprechende Signalwirkung erweitert sich der funktionale Horizont der E3-Ligasen um ein Vielfaches. Entscheidend dabei ist die Platzierung des Ubiquitins auf dem Substratprotein wie auch die Synthetisierung der entsprechenden Ubiquitinkette.

E3-Ligasen erkennen ihre Substrate unter anderem durch sogenannte *Short Linear Motifs* (SLiMs), was als „kurze lineare Motive“ übersetzt werden kann. Hierfür nutzen E3-Ligasen ihren modularen Aufbau. Proteine sind grundsätzlich modular organisiert, wobei jedes Modul – eine Domäne – eine tertiäre Faltung hat, welche der Domäne eine bestimmte Funktion verleiht. Einige Domänen übernehmen katalytische Funktionen, während andere der spezifischen Erkennung von Proteinen dienen, teilweise durch die bereits erwähnten SLiMs. Ein SLiM ist ein 3–10 Aminosäuren langer Abschnitt eines Proteins mit geringer Sekundärstruktur. Dies impliziert, dass dieser Abschnitt keine höhere Ordnung ausbildet und folglich als intrinsisch ungeordnet klassifiziert wird. Ein so kurzer Abschnitt von Aminosäuren wird als Peptid bezeichnet. Das SLiM erkennt aufgrund von 3–4 Schlüsselaminosäuren die Bindungspartner.

Während der katalytische Weg der E3-Ligasen gut beschrieben ist, mangelt es am Wissen über die Interaktionspartner und über die jeweiligen zellulären Auswirkungen der Ubiquitinierung. Die zentrale Frage hierbei ist, wie die verschiedenen E3-Ligasen ihre Substrate erkennen.

Neben den E3-Ligasen tragen zahlreiche weitere Proteine zu diesem Stoffwechselweg bei. Im Folgenden fasse ich sowohl die E3-Ligasen als auch diese begleitenden Komponenten unter dem Begriff der E3-Ligase-Maschinerie zusammen. Weitere Komponenten sind die ubiquitinkettenabbauenden Deubiquitinasen oder Proteine wie RAD23B, welches ubiquitinierte Proteine an das Proteasom weiterleitet.

Unter dem Oberbegriff der Determinanten der Substratspezifität untersuchen wir, wie genau E3-Ligasen ihre Substrate erkennen. Dafür habe ich eine Sammlung mit über 130 Domänen oder Volllängen-Proteine aufgebaut (Manuscript I-III). Zusammen mit meinen Kollegen und Studenten habe ich über 140 Domänen auf ihre SLiM-Bindefähigkeit geprüft. Dafür habe ich *Proteomic Peptide Phage Display* (ProP-PD) angewandt. Phagen-Display ist eine Display-Methode, welche zuerst zum Design von Antikörpern entwickelt wurde. Durch Weiterentwicklung kann man Phagen-Display auch zur Untersuchung von humanen Proteininteraktionen verwenden. In diesem Falle wird untersucht, welche Domäne welches Peptid bindet, was somit Rückschlüsse auf das jeweilige SLiM ermöglicht. Des Weiteren haben wir die entdeckten Interaktionen und SLiMs durch biochemische Verfahren charakterisiert. Wir haben die genauen substratspezifischen Determinanten auf Seite der Peptide wie auch der Domänen untersucht. Durch Fluoreszenzanisotropie, also die Ausnutzung der Welleneigenschaft von fluoreszierenden Lichtstrahlen, haben

wir die Interaktionsstärken zwischen den Peptiden und deren Domänen bestimmt. Weiterhin haben wir durch SPOT-Arrays die Peptidbindung der Domänen untersucht. SPOT-Arrays sind Membranen, welche mit synthetisierten Peptiden beschichtet sind. Dank der Entwicklung von künstlicher Intelligenz konnten wir computergestützt mit dem Programm AlphaFold die Interaktionen modellieren und experimentell validieren. Letztendlich haben wir die Interaktionen im zellulären Kontext mit Vollängen-Proteinen untersucht und durch die gefundenen SLiMs weitere Interaktionen prognostiziert und validiert. In Manuscript IV erweitern wir die biochemischen Grundlagen der SLiM-Bindung am Beispiel von Deubiquitinasen. Ich habe den Einfluss der Peptidlänge und der Aminosäuren außerhalb der Schlüsselaminosäuren untersucht.

Insgesamt konnten wir 11.460 Peptide finden, die mit der E3-Ligasen-Maschinerie interagieren. Diese Interaktionen stellen ungefähr 8.290 potentielle Protein-Protein-Interaktionen dar. Bei den momentan 12.380 bekannten Interaktionen stellt dies eine Erweiterung der globalen bekannten E3-Ligasen-Interaktionen von 67 % dar. Insbesondere konnten wir neue SLiMs charakterisieren, unter anderem die der MIB-Domänen-E3-Ligasen MIB1, MIB2, HECTD1 und HERC2. Ich konnte auch ein neues SLiM für die E3-Ligase HERC1 bestimmen. Für die E3-Ligasen RNF41 und STUB1 konnten wir SLiMs bestimmen, welche das Binden von bereits bekannten Interaktionspartnern vermitteln.

Meine Arbeit ist Teil der wissenschaftlichen Anstrengungen, das E3-Ligasen-System zu charakterisieren. Die Ergebnisse leisten einen substanziellen Beitrag zur Erweiterung des bekannten E3-Ligase-Interaktionsnetzwerks und verdeutlichen die zentrale Bedeutung von SLiMs für die Substraterkennung. Somit stellen die Daten eine wertvolle Ressource dar, auf die zukünftige Studien zurückgreifen können, um weiterführende Untersuchungen zu begründen und voranzutreiben. Darüber hinaus besitzen die Ergebnisse potenzielle Relevanz für die Entwicklung pharmakologischer Strategien, insbesondere in Hinblick auf die gezielte Modulation ubiquitin-abhängiger Signalwege.

9. Reflections and Acknowledgments

Whether Alexander von Humboldt or someone else has started initiated *the invention of nature*²⁷⁵ remains an engaging topic of discussion, but ultimately it is a historical debate. What drives us natural scientists is the admiration for nature. And here I wholeheartedly agree with von Humboldt's observation following his visits to the Americas: the greatness of nature reveals itself when we pause at what is little. I am deeply grateful to have had the opportunity to study something so very small and yet so remarkable — short linear motifs — and to appreciate the complexity and significance they embody admire the greatness of it: short linear motifs. While there was often little time to pause over the past years, as is typical in doctoral training, I value the journey nonetheless. Looking back, I can recognize how I have changed over time, in both subtle and profound ways.

Being part of this journey is my supervisor **Ylva**. I am happy you chose me six years ago to join your lab on this journey – not up the Orinoco, but down to the phage lab. I appreciate your untiring scientific curiosity, which inspired me many times. I am grateful that the lab you built up with so much heart has given me the opportunity to go on my very own expedition into nature. I am also thankful for your supervision and guidance. Your advice and companionship both in scientific matters and beyond the lab have made these years very impactful for me. Thank you.

My sincere thanks also go to the other supervisors and professors who have guided me along this path. In particular, I would like to thank my co-supervisor **Helena** – thank you for being a mentor to me since 2016, when I first attended your biosensors course. I am also grateful to the professors from the Ubimotif network: **Anton, Katrin, Helen, Mathias, Ora** and **Yogesh**. Your input mattered – thank you. I would like to express my special thanks to **Norman**, whose dedicated work on SLiMs and the design of libraries made my project possible. Thank you also for trying to teach us how to code.

I also would like to thank my professors from my undergraduate studies at the *Universität für Bodenkultur* in Vienna, who taught me to understand the

beauty of nature through a scientific lens, as well as my professors and teachers from my master's studies at Uppsala University. Representatively, I would like to thank you, **Suparna**.

I am very grateful to my students **Sofie, Biqin, Kristof, and Lowie**. I suspect I may have learned more from you than you did from me.

I would also like to thank my lab for making this journey possible. Firstly, **Emily**, thank You for Your companionship and for all the great memories we made over the years – it was great to work and celebrate with you. Also **Leandro**, thank you for your warmth, almost unlimited coding support, and accepting me into your office; thank you for processing my data. I also want to thank **Ali** and **Caroline** for developing ProP-PD and making it accessible for my work – this was invaluable to my project. My sincere thanks go to **Priyanka, Susanne, and Vishaka** for being such supportive and inspiring team members. **Lidia**, it has been a pleasure. I thank **Johanna**, whose onboarding in the early days encouraged me to find my own way.

Thank you, **Mostafa** for introducing me to Egyptian pop music and for being such a wonderful office mate. I also thank the later additions **Candice** and **Yalin**, who truly turned our office into a home. **Nadine**, it was fantastic to have you next door – thank you for always lending your ear. Thank you **Julia** for not only modelling my interactions, being awake so late and available for me, but also to be such a fun person. Also thanks to other Ubimotif students who accompanied me on this journey: **Caio, Katja, Javier, Marcel, Preethi,** and **Sophia**. Special thanks **Alicia, Andrea, Awa, Hazem, and Marie**, who dared to come to Uppsala. Of course, I also want to thank the rest of our biochemistry corridor. Thank you, **Moritz** for our trips to Australia and Ekoln. **Taru, Petra, Szabolcs, Kendra, and Joanna** for your open ears and laughs throughout the day. Naturally, **Marija, Leonardo, Anna-Lena, Lucas, Sad-dat, Ivan** and many more – thank you for keeping up the spirit in the corridor.

I would also like to thank the people who have had my back throughout these years and made the time outside the lab so enjoyable. My thanks goes to the members of the rowing club **UARS** in Uppsala for creating a great place to recharge. **Konstantin**, danke für deine gefühlt ewige Freundschaft und den unlimitierten Spaß. **Benedikt** und **Barbara**, danke, dass ihr mein Zuhause in Niederösterreich und Burgenland seid, zu dem ich immer wieder kommen darf. Danke an **Chema** und **Robert**, die mich in diesen Jahren ebenfalls begleitet haben.

Thank you **Rachael, Jean, and Raymond**. Whatever I write here would not do justice to your impact. I am also thankful to **Juulia, Marianna, Milena, Mostafa, Tea, and Umut**, who made sure that no opportunity for celebration was missed.

To **Alba and Petar**, thank you for having my back from the very beginning. It is hard to imagine that I would have made it without you. Our missions are not over yet – there is still more to explore. I also want to thank **Bella**, who made sure I went outside often enough while writing this thesis.

Ich möchte mich natürlich bei **Mama** für die vielen Telefonminuten bedanken, die mich immer wieder weitermachen ließen. Danke an meinen Vater[†]. Es ist schade, dass Du nicht hier sein kannst. Ebenfalls danke an Julia, André, Amalia, Adele und jetzt auch Lotti. Natürlich auch ein Dank an **A+T** für Eure unerschöpflichen Weisheiten. To **Rafael**, thank you for stepping into my life and opening your heart to me. Within a few months you changed my life – beyond caipirinhas and pudims.

10. References

1. Tompa, P., Davey, N. E., Gibson, T. J. & Babu, M. M. A million peptide motifs for the molecular biologist. *Mol Cell* **55**, 161–169 (2014).
2. Vacic, V. *et al.* Characterization of Molecular Recognition Features, MoRFs, and Their Binding Partners. *J. Proteome Res.* **6**, 2351–2366 (2007).
3. Pawson, T. & Nash, P. Assembly of Cell Regulatory Systems Through Protein Interaction Domains. *Science* **300**, 445–452 (2003).
4. Pancsa, R. & Tompa, P. Structural disorder in eukaryotes. *PLoS One* **7**, e34687 (2012).
5. Uversky, V. N., Gillespie, J. R. & Fink, A. L. Why are ‘natively unfolded’ proteins unstructured under physiologic conditions? *Proteins* **41**, 415–427 (2000).
6. Granata, D. *et al.* The inverted free energy landscape of an intrinsically disordered peptide by simulations and experiments. *Sci Rep* **5**, 15449 (2015).
7. Das, R. K., Ruff, K. M. & Pappu, R. V. Relating sequence encoded information to form and function of intrinsically disordered proteins. *Current Opinion in Structural Biology* **32**, 102–112 (2015).
8. Mao, A. H., Lyle, N. & Pappu, R. V. Describing sequence-ensemble relationships for intrinsically disordered proteins. *Biochem J* **449**, 307–318 (2013).
9. Mao, A. H., Crick, S. L., Vitalis, A., Chicoine, C. L. & Pappu, R. V. Net charge per residue modulates conformational ensembles of intrinsically disordered proteins. *Proc Natl Acad Sci U S A* **107**, 8183–8188 (2010).
10. Holehouse, A. S. & Kragelund, B. B. The molecular basis for cellular function of intrinsically disordered protein regions. *Nat Rev Mol Cell Biol* **25**, 187–211 (2024).
11. Moses, D., Ginell, G. M., Holehouse, A. S. & Sukenik, S. Intrinsically disordered regions are poised to act as sensors of cellular chemistry. *Trends Biochem Sci* **48**, 1019–1034 (2023).
12. Due, A. D. Short linear motifs and disordered context of Arabidopsis transcription factors in coregulator binding. (Kopenhagen University, Kopenhagen, 2024).

13. Mohan, A. *et al.* Analysis of Molecular Recognition Features (MoRFs). *Journal of Molecular Biology* **362**, 1043–1059 (2006).
14. Davey, N. E. *et al.* Attributes of short linear motifs. *Mol Biosyst* **8**, 268–281 (2012).
15. Teyra, J., Sidhu, S. S. & Kim, P. M. Elucidation of the binding preferences of peptide recognition modules: SH3 and PDZ domains. *FEBS Letters* **586**, 2631–2637 (2012).
16. Edwards, R. J., Davey, N. E., Brien, K. O. & Shields, D. C. Interactome-wide prediction of short, disordered protein interaction motifs in humans^{†‡}. *Mol Biosyst* **8**, 282–295 (2012).
17. Guo, Z. *et al.* Proteomics strategy to identify substrates of LNX, a PDZ domain-containing E3 ubiquitin ligase. *J Proteome Res* **11**, 4847–4862 (2012).
18. Dodson, E. J., Fishbain-Yoskovitz, V., Rotem-Bamberger, S. & Schueler-Furman, O. Versatile communication strategies among tandem WW domain repeats. *Exp Biol Med (Maywood)* **240**, 351–360 (2015).
19. Ivarsson, Y. Plasticity of PDZ domains in ligand recognition and signaling. *FEBS Letters* **586**, 2638–2647 (2012).
20. Tonikian, R. *et al.* A specificity map for the PDZ domain family. *PLoS Biol* **6**, e239 (2008).
21. Nie, J. *et al.* LNX functions as a RING type E3 ubiquitin ligase that targets the cell fate determinant Numb for ubiquitin-dependent degradation. *EMBO J* **21**, 93–102 (2002).
22. Young, P. W. LNX1/LNX2 proteins: functions in neuronal signalling and beyond. *Neuronal Signal* **2**, NS20170191 (2018).
23. Mehrabipour, M., Jasemi, N. S. K., Dvorsky, R. & Ahmadian, M. R. A Systematic Compilation of Human SH3 Domains: A Versatile Superfamily in Cellular Signaling. *Cells* **12**, 2054 (2023).
24. Kurochkina, N. & Guha, U. SH3 domains: modules of protein-protein interactions. *Biophys Rev* **5**, 29–39 (2013).
25. Hu, H. *et al.* A map of WW domain family interactions. *Proteomics* **4**, 643–655 (2004).
26. Huang, S.-S., Hsu, L.-J. & Chang, N.-S. Functional role of WW domain-containing proteins in tumor biology and diseases: Insight into the role in ubiquitin-proteasome system. *FASEB BioAdvances* **2**, 234–253 (2020).
27. Kumar, M. *et al.* ELM-the Eukaryotic Linear Motif resource-2024 update. *Nucleic Acids Res* **52**, D442–D455 (2024).
28. Krystkowiak, I. & Davey, N. E. The MoMaP resource: a curated motif map of the proteome. <https://slim-tools.org/momap/>.

29. London, N., Movshovitz-Attias, D. & Schueler-Furman, O. The Structural Basis of Peptide-Protein Binding Strategies. *Structure* **18**, 188–199 (2010).
30. Gianni, S., Dogan, J. & Jemth, P. Coupled binding and folding of intrinsically disordered proteins: what can we learn from kinetics? *Curr Opin Struct Biol* **36**, 18–24 (2016).
31. Kjaer, L. F. *et al.* Hierarchical folding-upon-binding of an intrinsically disordered protein. *Nat Commun* **16**, 11346 (2025).
32. Wright, P. E. & Dyson, H. J. Linking folding and binding. *Curr Opin Struct Biol* **19**, 31–38 (2009).
33. Rustandi, R. R., Baldissari, D. M. & Weber, D. J. Structure of the negative regulatory domain of p53 bound to S100B(beta-beta). *Nat Struct Biol* **7**, 570–574 (2000).
34. Tuttle, L. M. *et al.* Gcn4-Mediator Specificity Is Mediated by a Large and Dynamic Fuzzy Protein-Protein Complex. *Cell Rep* **22**, 3251–3264 (2018).
35. Berlow, R. B., Dyson, H. J. & Wright, P. E. Multivalency enables unidirectional switch-like competition between intrinsically disordered proteins. *Proc Natl Acad Sci U S A* **119**, e2117338119 (2022).
36. Lim, W. A. & Mayer, B. J. *Cell Signaling: Principles and Mechanisms*. (CRC Press, Boca Raton, 2024).
37. Finlay, D. B., Duffull, S. B. & Glass, M. 100 years of modelling ligand–receptor binding and response: A focus on GPCRs. *Br J Pharmacol* **177**, 1472–1484 (2020).
38. Langmuir, I. THE ADSORPTION OF GASES ON PLANE SURFACES OF GLASS, MICA AND PLATINUM. *J. Am. Chem. Soc.* **40**, 1361–1403 (1918).
39. Skriver, K., Theisen, F. F. & Kragelund, B. B. Conformational entropy in molecular recognition of intrinsically disordered proteins. *Current Opinion in Structural Biology* **83**, 102697 (2023).
40. Pollard, T. D. A guide to simple and informative binding assays. *Mol Biol Cell* **21**, 4061–4067 (2010).
41. Spolar, R. S. & Record, M. T. Coupling of local folding to site-specific binding of proteins to DNA. *Science* **263**, 777–784 (1994).
42. Chodera, J. D. & Mobley, D. L. Entropy-enthalpy compensation: role and ramifications in biomolecular ligand recognition and design. *Annu Rev Biophys* **42**, 121–142 (2013).
43. Theisen, F. F. *et al.* Quantification of Conformational Entropy Unravels Effect of Disordered Flanking Region in Coupled Folding and Binding. *J Am Chem Soc* **143**, 14540–14550 (2021).

44. Benz, C. *et al.* Proteome-scale mapping of binding sites in the unstructured regions of the human proteome. *Mol Syst Biol* **18**, e10584 (2022).
45. Dogan, J., Jonasson, J., Andersson, E. & Jemth, P. Binding Rate Constants Reveal Distinct Features of Disordered Protein Domains. *Biochemistry* **54**, 4741–4750 (2015).
46. Shammass, S. L., Rogers, J. M., Hill, S. A. & Clarke, J. Slow, reversible, coupled folding and binding of the spectrin tetramerization domain. *Biophys J* **103**, 2203–2214 (2012).
47. Pyhtila, B. & Rexach, M. A gradient of affinity for the karyopherin Kap95p along the yeast nuclear pore complex. *J Biol Chem* **278**, 42699–42709 (2003).
48. Timney, B. L. *et al.* Simple kinetic relationships and nonspecific competition govern nuclear import rates in vivo. *J Cell Biol* **175**, 579–593 (2006).
49. Ivarsson, Y. & Jemth, P. Affinity and specificity of motif-based protein-protein interactions. *Curr Opin Struct Biol* **54**, 26–33 (2019).
50. Kliche, J. & Ivarsson, Y. Orchestrating serine/threonine phosphorylation and elucidating downstream effects by short linear motifs. *Biochem J* **479**, 1–22 (2022).
51. Sundell, G. N. *et al.* Proteome-wide analysis of phospho-regulated PDZ domain interactions. *Mol Syst Biol* **14**, e8129 (2018).
52. Suskiewicz, M. J. The logic of protein post-translational modifications (PTMs): Chemistry, mechanisms and evolution of protein regulation through covalent attachments. *BioEssays* **46**, 2300178 (2024).
53. Mammen, M., Choi, S.-K. & Whitesides, G. M. Polyvalent Interactions in Biological Systems: Implications for Design and Use of Multivalent Ligands and Inhibitors. *Angew Chem Int Ed Engl* **37**, 2754–2794 (1998).
54. Klein, P., Pawson, T. & Tyers, M. Mathematical modeling suggests cooperative interactions between a disordered polyvalent ligand and a single receptor site. *Curr Biol* **13**, 1669–1678 (2003).
55. Bugge, K. *et al.* Interactions by Disorder – A Matter of Context. *Front. Mol. Biosci.* **7**, 110 (2020).
56. Baker, K. *et al.* Yorkie-Warts Complexes are an Ensemble of Interconverting Conformers Formed by Multivalent Interactions. *Journal of Molecular Biology* **433**, 166776 (2021).
57. Liu, Y. *et al.* Structural basis for the regulatory role of the PPxY motifs in the thioredoxin-interacting protein TXNIP. *Biochem J* **473**, 179–187 (2016).
58. Krystkowiak, I. & Davey, N. E. SLiMSearch: a framework for proteome-wide discovery and annotation of functional modules in intrinsically disordered regions. *Nucleic Acids Res* **45**, W464–W469 (2017).

59. Gibson, T. J., Dinkel, H., Van Roey, K. & Diella, F. Experimental detection of short regulatory motifs in eukaryotic proteins: tips for good practice as well as for bad. *Cell Commun Signal* **13**, 42 (2015).
60. Davey, N. E., Shields, D. C. & Edwards, R. J. Masking residues using context-specific evolutionary conservation significantly improves short linear motif discovery. *Bioinformatics* **25**, 443–450 (2009).
61. Lee, C. Y. *et al.* Systematic discovery of protein interaction interfaces using AlphaFold and experimental validation. *Mol Syst Biol* **20**, 75–97 (2024).
62. Celebrating the death of proteins. *Nat Struct Mol Biol* **11**, 1025–1025 (2004).
63. Hershko, A. & Ciechanover, A. The ubiquitin system. *Annu Rev Biochem* **67**, 425–479 (1998).
64. Hershko, A., Heller, H., Elias, S. & Ciechanover, A. Components of ubiquitin-protein ligase system. Resolution, affinity purification, and role in protein breakdown. *J Biol Chem* **258**, 8206–8214 (1983).
65. Schulman, B. A. & Wade Harper, J. Ubiquitin-like protein activation by E1 enzymes: the apex for downstream signalling pathways. *Nat Rev Mol Cell Biol* **10**, 319–331 (2009).
66. Chua, N. K. *et al.* The E3-ome gene-centric compendium reveals the human E3 ligase landscape. *Cell* S0092867426001169 (2026) doi:10.1016/j.cell.2026.01.029.
67. Morreale, F. E. & Walden, H. Types of Ubiquitin Ligases. *Cell* **165**, 248–248.e1 (2016).
68. Dikic, I. & Schulman, B. A. An expanded lexicon for the ubiquitin code. *Nat Rev Mol Cell Biol* **24**, 273–287 (2023).
69. Swatek, K. N. & Komander, D. Ubiquitin modifications. *Cell Res* **26**, 399–422 (2016).
70. Komander, D. & Rape, M. The Ubiquitin Code. *Annual Review of Biochemistry* **81**, 203–229 (2012).
71. Zhang, Z., Mena, E. L., Timms, R. T., Koren, I. & Elledge, S. J. Degrons: defining the rules of protein degradation. *Nat Rev Mol Cell Biol* **26**, 868–883 (2025).
72. Kirisako, T. *et al.* A ubiquitin ligase complex assembles linear polyubiquitin chains. *EMBO J* **25**, 4877–4887 (2006).
73. Maitland, M. E. R., Lajoie, G. A., Shaw, G. S. & Schild-Poulter, C. Structural and Functional Insights into GID/CTLH E3 Ligase Complexes. *Int J Mol Sci* **23**, 5863 (2022).
74. Okoye, C. N., Rowling, P. J. E., Itzhaki, L. S. & Lindon, C. Counting Degrons: Lessons From Multivalent Substrates for Targeted Protein Degradation. *Front. Physiol.* **13**, (2022).

75. Wang, S., Teng, F., Stjepanovic, G., Rao, F. & Su, M.-Y. Cryo-EM structure of the human COP1-DET1 ubiquitin ligase complex. *Nat Commun* **17**, 543 (2026).
76. Mevissen, T. E. T. & Komander, D. Mechanisms of Deubiquitinase Specificity and Regulation. *Annu Rev Biochem* **86**, 159–192 (2017).
77. Kelsall, I. R., Zhang, J., Knebel, A., Arthur, J. S. C. & Cohen, P. The E3 ligase HOIL-1 catalyses ester bond formation between ubiquitin and components of the Myddosome in mammalian cells. *Proceedings of the National Academy of Sciences* **116**, 13293–13298 (2019).
78. Konstantinou, A. *et al.* Elucidation of short linear motif-based interactions of the MIT and rhodanese domains of the ubiquitin-specific protease 8. *Biol Direct* **20**, 59 (2025).
79. Konstantinou, A. Expanding the motif-based interactome: Insights into the recognition landscape of deubiquitinases. (Acta Universitatis Upsalensis, Uppsala, 2025).
80. McDowell, G. S. & Philpott, A. New Insights Into the Role of Ubiquitylation of Proteins. *Int Rev Cell Mol Biol* **325**, 35–88 (2016).
81. Haas, A. L., Ahrens, P., Bright, P. M. & Ankel, H. Interferon induces a 15-kilodalton protein exhibiting marked homology to ubiquitin. *J Biol Chem* **262**, 11315–11323 (1987).
82. Kamitani, T., Kito, K., Nguyen, H. P. & Yeh, E. T. Characterization of NEDD8, a developmentally down-regulated ubiquitin-like protein. *J Biol Chem* **272**, 28557–28562 (1997).
83. Mahajan, R., Delphin, C., Guan, T., Gerace, L. & Melchior, F. A small ubiquitin-related polypeptide involved in targeting RanGAP1 to nuclear pore complex protein RanBP2. *Cell* **88**, 97–107 (1997).
84. Matunis, M. J., Coutavas, E. & Blobel, G. A novel ubiquitin-like modification modulates the partitioning of the Ran-GTPase-activating protein RanGAP1 between the cytosol and the nuclear pore complex. *J Cell Biol* **135**, 1457–1470 (1996).
85. Swaney, D. L. *et al.* Global analysis of phosphorylation and ubiquitylation cross-talk in protein degradation. *Nat Methods* **10**, 676–682 (2013).
86. Peng, J. *et al.* A proteomics approach to understanding protein ubiquitination. *Nat Biotechnol* **21**, 921–926 (2003).
87. Wauer, T. *et al.* Ubiquitin Ser65 phosphorylation affects ubiquitin structure, chain assembly and hydrolysis. *EMBO J* **34**, 307–325 (2015).
88. Kienle, S. M. *et al.* Electrostatic and steric effects underlie acetylation-induced changes in ubiquitin structure and function. *Nat Commun* **13**, 5435 (2022).
89. House, C. M. *et al.* A binding motif for Siah ubiquitin ligase. *Proc Natl Acad Sci U S A* **100**, 3101–3106 (2003).

90. Kussie, P. H. *et al.* Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain. *Science* **274**, 948–953 (1996).
91. Tong, K. I. *et al.* Keap1 recruits Neh2 through binding to ETGE and DLG motifs: characterization of the two-site molecular recognition model. *Mol Cell Biol* **26**, 2887–2900 (2006).
92. Zeng, Z. *et al.* Structural basis of selective ubiquitination of TRF1 by SCFFbx4. *Dev Cell* **18**, 214–225 (2010).
93. Xing, W. *et al.* SCF(FBXL3) ubiquitin ligase targets cryptochromes at their cofactor pocket. *Nature* **496**, 64–68 (2013).
94. McGinty, R. K., Henrici, R. C. & Tan, S. Crystal structure of the PRC1 ubiquitylation module bound to the nucleosome. *Nature* **514**, 591–596 (2014).
95. James, L. C., Keeble, A. H., Khan, Z., Rhodes, D. A. & Trowsdale, J. Structural basis for PRYSPRY-mediated tripartite motif (TRIM) protein function. *Proc Natl Acad Sci U S A* **104**, 6200–6205 (2007).
96. Mészáros, B., Kumar, M., Gibson, T. J., Uyar, B. & Dosztányi, Z. Degrons in cancer. *Sci Signal* **10**, eaak9982 (2017).
97. Sherpa, D., Chrustowicz, J. & Schulman, B. A. How the ends signal the end: Regulation by E3 ubiquitin ligases recognizing protein termini. *Mol Cell* **82**, 1424–1438 (2022).
98. Zheng, N. & Shabek, N. Ubiquitin Ligases: Structure, Function, and Regulation. *Annu. Rev. Biochem.* **86**, 129–157 (2017).
99. Crews, C. M. Feeding the machine: mechanisms of proteasome-catalyzed degradation of ubiquitinated proteins. *Curr Opin Chem Biol* **7**, 534–539 (2003).
100. Nash, P. *et al.* Multisite phosphorylation of a CDK inhibitor sets a threshold for the onset of DNA replication. *Nature* **414**, 514–521 (2001).
101. Wu, G. *et al.* Structure of a beta-TrCP1-Skp1-beta-catenin complex: destruction motif binding and lysine specificity of the SCF(beta-TrCP1) ubiquitin ligase. *Mol Cell* **11**, 1445–1456 (2003).
102. Welcker, M. & Clurman, B. E. FBW7 ubiquitin ligase: a tumour suppressor at the crossroads of cell division, growth and differentiation. *Nat Rev Cancer* **8**, 83–93 (2008).
103. Sakaguchi, K. *et al.* Damage-mediated phosphorylation of human p53 threonine 18 through a cascade mediated by a casein 1-like kinase. Effect on Mdm2 binding. *J Biol Chem* **275**, 9278–9283 (2000).
104. Yu, F., White, S. B., Zhao, Q. & Lee, F. S. HIF-1alpha binding to VHL is regulated by stimulus-sensitive proline hydroxylation. *Proc Natl Acad Sci U S A* **98**, 9630–9635 (2001).

105. Ivan, M. *et al.* HIF α targeted for VHL-mediated destruction by proline hydroxylation: implications for O₂ sensing. *Science* **292**, 464–468 (2001).
106. Ganoth, D. *et al.* The cell-cycle regulatory protein Cks1 is required for SCF(Skp2)-mediated ubiquitinylation of p27. *Nat Cell Biol* **3**, 321–324 (2001).
107. Fischer, E. S. *et al.* Structure of the DDB1-CRBN E3 ubiquitin ligase in complex with thalidomide. *Nature* **512**, 49–53 (2014).
108. Coyaud, E. *et al.* BioID-based Identification of Skp Cullin F-box (SCF) β -TrCP1/2 E3 Ligase Substrates. *Mol Cell Proteomics* **14**, 1781–1795 (2015).
109. 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* **26**, 131–143 (2007).
110. Rape, M., Reddy, S. K. & Kirschner, M. W. The processivity of multi-ubiquitination by the APC determines the order of substrate degradation. *Cell* **124**, 89–103 (2006).
111. Lu, Y., Wang, W. & Kirschner, M. W. Specificity of the anaphase-promoting complex: a single-molecule study. *Science* **348**, 1248737 (2015).
112. Lu, Y., Lee, B., King, R. W., Finley, D. & Kirschner, M. W. Substrate degradation by the proteasome: a single-molecule kinetic analysis. *Science* **348**, 1250834 (2015).
113. Hartooni, N., Sung, J., Jain, A. & Morgan, D. O. Single-molecule analysis of specificity and multivalency in binding of short linear substrate motifs to the APC/C. *Nat Commun* **13**, 341 (2022).
114. Dutta, A. *et al.* Multi-scale classification decodes the complexity of the human E3 ligome. *Nat Commun* **16**, 11382 (2025).
115. Cruz Walma, D. A., Chen, Z., Bullock, A. N. & Yamada, K. M. Ubiquitin ligases: guardians of mammalian development. *Nat Rev Mol Cell Biol* **23**, 350–367 (2022).
116. Juszkievicz, S. & Hegde, R. S. Quality Control of Orphaned Proteins. *Mol Cell* **71**, 443–457 (2018).
117. Padovani, C., Jevtić, P. & Rapé, M. Quality control of protein complex composition. *Mol Cell* **82**, 1439–1450 (2022).
118. Meyer, H.-J. & Rape, M. Enhanced protein degradation by branched ubiquitin chains. *Cell* **157**, 910–921 (2014).
119. Pierce, N. W., Kleiger, G., Shan, S. & Deshaies, R. J. Detection of sequential polyubiquitylation on a millisecond timescale. *Nature* **462**, 615–619 (2009).

120. Kleiger, G., Saha, A., Lewis, S., Kuhlman, B. & Deshaies, R. J. Rapid E2-E3 assembly and disassembly enable processive ubiquitylation of cul-1-RING ubiquitin ligase substrates. *Cell* **139**, 957–968 (2009).
121. Ye, Y. & Rape, M. Building ubiquitin chains: E2 enzymes at work. *Nat Rev Mol Cell Biol* **10**, 755–764 (2009).
122. Tokheim, C. *et al.* Systematic characterization of mutations altering protein degradation in human cancers. *Mol Cell* **81**, 1292–1308.e11 (2021).
123. George, A. J., Hoffiz, Y. C., Charles, A. J., Zhu, Y. & Mabb, A. M. A Comprehensive Atlas of E3 Ubiquitin Ligase Mutations in Neurological Disorders. *Front Genet* **9**, 29 (2018).
124. Mihalič, F. *et al.* Large-scale phage-based screening reveals extensive pan-viral mimicry of host short linear motifs. *Nat Commun* **14**, 2409 (2023).
125. Scheffner, M., Werness, B. A., Huibregtse, J. M., Levine, A. J. & Howley, P. M. The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53. *Cell* **63**, 1129–1136 (1990).
126. Fiil, B. K. & Gyrd-Hansen, M. The Met1-linked ubiquitin machinery in inflammation and infection. *Cell Death Differ* **28**, 557–569 (2021).
127. Liu, Y. *et al.* Expanding PROTACtable genome universe of E3 ligases. *Nat Commun* **14**, 6509 (2023).
128. Hinterdorfer, M., Spiteri, V. A., Ciulli, A. & Winter, G. E. Targeted protein degradation for cancer therapy. *Nat Rev Cancer* **25**, 493–516 (2025).
129. Hsia, O. *et al.* Targeted protein degradation via intramolecular bivalent glues. *Nature* **627**, 204–211 (2024).
130. Békés, M., Langley, D. R. & Crews, C. M. PROTAC targeted protein degraders: the past is prologue. *Nat Rev Drug Discov* **21**, 181–200 (2022).
131. Fields, S. & Song, O. A novel genetic system to detect protein-protein interactions. *Nature* **340**, 245–246 (1989).
132. Gavin, A.-C. *et al.* Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* **415**, 141–147 (2002).
133. Roux, K. J., Kim, D. I., Raida, M. & Burke, B. A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *J Cell Biol* **196**, 801–810 (2012).
134. Rhee, H.-W. *et al.* Proteomic mapping of mitochondria in living cells via spatially restricted enzymatic tagging. *Science* **339**, 1328–1331 (2013).
135. Ullman, E. F. *et al.* Luminescent oxygen channeling immunoassay: measurement of particle binding kinetics by chemiluminescence. *Proc Natl Acad Sci U S A* **91**, 5426–5430 (1994).

136. Takahashi, H. *et al.* Establishment of a Wheat Cell-Free Synthesized Protein Array Containing 250 Human and Mouse E3 Ubiquitin Ligases to Identify Novel Interaction between E3 Ligases and Substrate Proteins. *PLoS One* **11**, e0156718 (2016).
137. Weber, E., Rothenaigner, I., Brandner, S., Hadian, K. & Schorpp, K. A High-Throughput Screening Strategy for Development of RNF8-Ubc13 Protein-Protein Interaction Inhibitors. *SLAS Discov* **22**, 316–323 (2017).
138. Sternicki, L. M., Crowe, C., Wieske, L. H. E., Ciulli, A. & Poulsen, S.-A. Native Mass Spectrometry Analysis of Cullin RING Ubiquitin E3 Ligase Complexes in the Context of Targeted Protein Degradation. *bioRxiv* 2025.12.04.692288 (2026) doi:10.64898/2025.12.04.692288.
139. Davey, N. E. *et al.* Discovery of short linear motif-mediated interactions through phage display of intrinsically disordered regions of the human proteome. *The FEBS Journal* **284**, 485–498 (2017).
140. Smith, G. P. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* **228**, 1315–1317 (1985).
141. Parmley, S. F. & Smith, G. P. Antibody-selectable filamentous fd phage vectors: affinity purification of target genes. *Gene* **73**, 305–318 (1988).
142. Scott, J. K. & Smith, G. P. Searching for peptide ligands with an epitope library. *Science* **249**, 386–390 (1990).
143. Zhang, Y. *et al.* Inhibition of Wnt signaling by Dishevelled PDZ peptides. *Nat Chem Biol* **5**, 217–219 (2009).
144. Luck, K. & Travé, G. Phage display can select over-hydrophobic sequences that may impair prediction of natural domain-peptide interactions. *Bioinformatics* **27**, 899–902 (2011).
145. Sidhu, S. S. Engineering M13 for phage display. *Biomol Eng* **18**, 57–63 (2001).
146. Kliche, J. *et al.* Large-scale phosphomimetic screening identifies phospho-modulated motif-based protein interactions. *Mol Syst Biol* **19**, e11164 (2023).
147. Kliche, J. *et al.* Proteome-scale characterisation of motif-based interactome rewiring by disease mutations. *Mol Syst Biol* **20**, 1025–1048 (2024).
148. Örd, M. *et al.* High-throughput investigation of cyclin docking interactions reveals the complexity of motif binding determinants. *Nat Commun* **16**, 7622 (2025).
149. Frank, R. The SPOT-synthesis technique. Synthetic peptide arrays on membrane supports--principles and applications. *J Immunol Methods* **267**, 13–26 (2002).

150. Yim, Y. Y., Betke, K. & Hamm, H. Using peptide arrays created by the SPOT method for defining protein-protein interactions. *Methods Mol Biol* **1278**, 307–320 (2015).
151. Moerke, N. J. Fluorescence Polarization (FP) Assays for Monitoring Peptide-Protein or Nucleic Acid-Protein Binding. *Curr Protoc Chem Biol* **1**, 1–15 (2009).
152. Bhaduri, S. *et al.* A docking interface in the cyclin Cln2 promotes multi-site phosphorylation of substrates and timely cell-cycle entry. *Curr Biol* **25**, 316–325 (2015).
153. Bhaduri, S. & Pryciak, P. M. Cyclin-specific docking motifs promote phosphorylation of yeast signaling proteins by G1/S Cdk complexes. *Curr Biol* **21**, 1615–1623 (2011).
154. Strickfaden, S. C. *et al.* A mechanism for cell-cycle regulation of MAP kinase signaling in a yeast differentiation pathway. *Cell* **128**, 519–531 (2007).
155. Subbanna, M. S., Winters, M. J., Örd, M., Davey, N. E. & Pryciak, P. M. A quantitative intracellular peptide-binding assay reveals recognition determinants and context dependence of short linear motifs. *Journal of Biological Chemistry* **301**, 108225 (2025).
156. DeCaprio, J. & Kohl, T. O. Immunoprecipitation. *Cold Spring Harb Protoc* **2020**, (2020).
157. Brizzard, B. Epitope tagging. *Biotechniques* **44**, 693–695 (2008).
158. Hubner, N. C. *et al.* Quantitative proteomics combined with BAC TransgeneOmics reveals in vivo protein interactions. *J Cell Biol* **189**, 739–754 (2010).
159. Yen, H.-C. S., Xu, Q., Chou, D. M., Zhao, Z. & Elledge, S. J. Global protein stability profiling in mammalian cells. *Science* **322**, 918–923 (2008).
160. Zhang, Z. *et al.* Elucidation of E3 ubiquitin ligase specificity through proteome-wide internal degron mapping. *Mol Cell* **83**, 3377-3392.e6 (2023).
161. Timms, R. T. *et al.* Defining E3 ligase–substrate relationships through multiplex CRISPR screening. *Nat Cell Biol* **25**, 1535–1545 (2023).
162. Nguyen, T. V. *et al.* Glutamine Triggers Acetylation-Dependent Degradation of Glutamine Synthetase via the Thalidomide Receptor Cereblon. *Mol Cell* **61**, 809–820 (2016).
163. McMillan, B. J. *et al.* A tail of two sites: a bipartite mechanism for recognition of notch ligands by mind bomb E3 ligases. *Mol Cell* **57**, 912–924 (2015).

164. Mancias, J. D. *et al.* Ferritinophagy via NCOA4 is required for erythropoiesis and is regulated by iron dependent HERC2-mediated proteolysis. *Elife* **4**, e10308 (2015).
165. Brown, N. G. *et al.* Mechanism of polyubiquitination by human anaphase-promoting complex: RING repurposing for ubiquitin chain assembly. *Mol Cell* **56**, 246–260 (2014).
166. Kelsall, I. R. Non-lysine ubiquitylation: Doing things differently. *Front. Mol. Biosci.* **9**, (2022).
167. Wu, X. *et al.* Structure-guided engineering enables E3 ligase-free and versatile protein ubiquitination via UBE2E1. *Nat Commun* **15**, 1266 (2024).
168. Mali, S. M., Singh, S. K., Eid, E. & Brik, A. Ubiquitin Signaling: Chemistry Comes to the Rescue. *J Am Chem Soc* **139**, 4971–4986 (2017).
169. Staub, O. *et al.* WW domains of Nedd4 bind to the proline-rich PY motifs in the epithelial Na⁺ channel deleted in Liddle's syndrome. *EMBO J* **15**, 2371–2380 (1996).
170. Fotia, A. B. *et al.* The role of individual Nedd4-2 (KIAA0439) WW domains in binding and regulating epithelial sodium channels. *FASEB J* **17**, 70–72 (2003).
171. Ren, R., Mayer, B. J., Cicchetti, P. & Baltimore, D. Identification of a ten-amino acid proline-rich SH3 binding site. *Science* **259**, 1157–1161 (1993).
172. Yu, H. *et al.* Solution structure of the SH3 domain of Src and identification of its ligand-binding site. *Science* **258**, 1665–1668 (1992).
173. Chen, H. I. & Sudol, M. The WW domain of Yes-associated protein binds a proline-rich ligand that differs from the consensus established for Src homology 3-binding modules. *Proc Natl Acad Sci U S A* **92**, 7819–7823 (1995).
174. Sugiura, M. *et al.* Ceramide kinase, a novel lipid kinase. Molecular cloning and functional characterization. *J Biol Chem* **277**, 23294–23300 (2002).
175. Teyra, J. *et al.* Comprehensive Analysis of the Human SH3 Domain Family Reveals a Wide Variety of Non-canonical Specificities. *Structure* **25**, 1598-1610.e3 (2017).
176. Hou, T., Li, N., Li, Y. & Wang, W. Characterization of domain-peptide interaction interface: prediction of SH3 domain-mediated protein-protein interaction network in yeast by generic structure-based models. *J Proteome Res* **11**, 2982–2995 (2012).
177. Kay, B. K., Williamson, M. P. & Sudol, M. The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. *FASEB J* **14**, 231–241 (2000).

178. Tapon, N., Nagata, K., Lamarche, N. & Hall, A. A new rac target POSH is an SH3-containing scaffold protein involved in the JNK and NF-kappaB signalling pathways. *EMBO J* **17**, 1395–1404 (1998).
179. Rotin, D. & Kumar, S. Physiological functions of the HECT family of ubiquitin ligases. *Nat Rev Mol Cell Biol* **10**, 398–409 (2009).
180. Myers, N. E. M. Biochemical strategies for ligand discovery against cancer targets. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology* (Acta Universitatis Upsaliensis, Uppsala, 2024).
181. Li, X., Zhang, D., Hannink, M. & Beamer, L. J. Crystal Structure of the Kelch Domain of Human Keap1*. *Journal of Biological Chemistry* **279**, 54750–54758 (2004).
182. Takahashi, D. *et al.* KLHL2 interacts with and ubiquitinates WNK kinases. *Biochem Biophys Res Commun* **437**, 457–462 (2013).
183. Chen, Z. *et al.* Identification of a PGXPP degron motif in dishevelled and structural basis for its binding to the E3 ligase KLHL12. *Open Biol* **10**, 200041 (2020).
184. Kim, B. H. *et al.* Crystal structure of the Ate1 arginyl-tRNA-protein transferase and arginylation of N-degron substrates. *Proceedings of the National Academy of Sciences* **119**, e2209597119 (2022).
185. Bernstein, E. *et al.* Mouse Polycomb Proteins Bind Differentially to Methylated Histone H3 and RNA and Are Enriched in Facultative Heterochromatin. *Mol Cell Biol* **26**, 2560–2569 (2006).
186. Kaustov, L. *et al.* Recognition and Specificity Determinants of the Human Cbx Chromodomains. *J Biol Chem* **286**, 521–529 (2011).
187. Buchwald, G. *et al.* Structure and E3-ligase activity of the Ring-Ring complex of polycomb proteins Bmi1 and Ring1b. *EMBO J* **25**, 2465–2474 (2006).
188. Connelly, K. E. *et al.* Engagement of DNA and H3K27me3 by the CBX8 chromodomain drives chromatin association. *Nucleic Acids Res* **47**, 2289–2305 (2019).
189. Ng, J. M. Y. *et al.* A novel regulation mechanism of DNA repair by damage-induced and RAD23-dependent stabilization of xeroderma pigmentosum group C protein. *Genes Dev* **17**, 1630–1645 (2003).
190. Liang, R.-Y. *et al.* Rad23 Interaction with the Proteasome Is Regulated by Phosphorylation of Its Ubiquitin-Like (Ubl) Domain. *Journal of Molecular Biology* **426**, 4049–4060 (2014).
191. Ryu, K.-S. *et al.* Binding surface mapping of intra- and interdomain interactions among hHR23B, ubiquitin, and polyubiquitin binding site 2 of S5a. *J Biol Chem* **278**, 36621–36627 (2003).

192. Huttlin, E. L. *et al.* Architecture of the human interactome defines protein communities and disease networks. *Nature* **545**, 505–509 (2017).
193. Luck, K. *et al.* A reference map of the human binary protein interactome. *Nature* **580**, 402–408 (2020).
194. Fujiwara, K. *et al.* Structure of the ubiquitin-interacting motif of S5a bound to the ubiquitin-like domain of HR23B. *J Biol Chem* **279**, 4760–4767 (2004).
195. Walters, K. J., Kleijnen, M. F., Goh, A. M., Wagner, G. & Howley, P. M. Structural studies of the interaction between ubiquitin family proteins and proteasome subunit S5a. *Biochemistry* **41**, 1767–1777 (2002).
196. Rual, J.-F. *et al.* Towards a proteome-scale map of the human protein-protein interaction network. *Nature* **437**, 1173–1178 (2005).
197. Perfetto, L. *et al.* Exploring the diversity of SPRY/B30.2-mediated interactions. *Trends in Biochemical Sciences* **38**, 38–46 (2013).
198. Nishiya, T. *et al.* Regulation of inducible nitric-oxide synthase by the SPRY domain- and SOCS box-containing proteins. *J Biol Chem* **286**, 9009–9019 (2011).
199. D’Cruz, A. A., Babon, J. J., Norton, R. S., Nicola, N. A. & Nicholson, S. E. Structure and function of the SPRY/B30.2 domain proteins involved in innate immunity. *Protein Sci* **22**, 1–10 (2013).
200. Masters, S. L. *et al.* The SPRY domain of SSB-2 adopts a novel fold that presents conserved Par-4-binding residues. *Nat Struct Mol Biol* **13**, 77–84 (2006).
201. Kobayashi, N. *et al.* RanBPM, Muskelin, p48EMLP, p44CTLH, and the armadillo-repeat proteins ARMC8alpha and ARMC8beta are components of the CTLH complex. *Gene* **396**, 236–247 (2007).
202. Sherpa, D. *et al.* Modular UBE2H-CTLH E2-E3 complexes regulate erythroid maturation. *Elife* **11**, e77937 (2022).
203. Woo, J.-S., Suh, H.-Y., Park, S.-Y. & Oh, B.-H. Structural Basis for Protein Recognition by B30.2/SPRY Domains. *Molecular Cell* **24**, 967–976 (2006).
204. Wang, H.-T. & Hur, S. Substrate recognition by TRIM and TRIM-like proteins in innate immunity. *Semin Cell Dev Biol* **111**, 76–85 (2021).
205. Guo, B., McMillan, B. J. & Blacklow, S. C. Structure and function of the Mind bomb E3 ligase in the context of Notch signal transduction. *Current Opinion in Structural Biology* **41**, 38–45 (2016).
206. Itoh, M. *et al.* Mind Bomb Is a Ubiquitin Ligase that Is Essential for Efficient Activation of Notch Signaling by Delta. *Developmental Cell* **4**, 67–82 (2003).

207. Baer, R. & Ludwig, T. The BRCA1/BARD1 heterodimer, a tumor suppressor complex with ubiquitin E3 ligase activity. *Curr Opin Genet Dev* **12**, 86–91 (2002).
208. Kalb, R. *et al.* Histone H2A monoubiquitination promotes histone H3 methylation in Polycomb repression. *Nat Struct Mol Biol* **21**, 569–571 (2014).
209. Tarsounas, M. & Sung, P. The antitumorigenic roles of BRCA1-BARD1 in DNA repair and replication. *Nat Rev Mol Cell Biol* **21**, 284–299 (2020).
210. Sobhian, B. *et al.* RAP80 targets BRCA1 to specific ubiquitin structures at DNA damage sites. *Science* **316**, 1198–1202 (2007).
211. Wang, B. *et al.* Abraxas and RAP80 form a BRCA1 protein complex required for the DNA damage response. *Science* **316**, 1194–1198 (2007).
212. Witus, S. R. *et al.* BRCA1/BARD1 intrinsically disordered regions facilitate chromatin recruitment and ubiquitylation. *EMBO J* **42**, e113565 (2023).
213. Nishikawa, H. *et al.* Mass spectrometric and mutational analyses reveal Lys-6-linked polyubiquitin chains catalyzed by BRCA1-BARD1 ubiquitin ligase. *J Biol Chem* **279**, 3916–3924 (2004).
214. Wu-Baer, F., Lagrazon, K., Yuan, W. & Baer, R. The BRCA1/BARD1 heterodimer assembles polyubiquitin chains through an unconventional linkage involving lysine residue K6 of ubiquitin. *J Biol Chem* **278**, 34743–34746 (2003).
215. Yu, X., Chini, C. C. S., He, M., Mer, G. & Chen, J. The BRCT domain is a phospho-protein binding domain. *Science* **302**, 639–642 (2003).
216. Manke, I. A., Lowery, D. M., Nguyen, A. & Yaffe, M. B. BRCT repeats as phosphopeptide-binding modules involved in protein targeting. *Science* **302**, 636–639 (2003).
217. Nakamura, K. *et al.* H4K20me0 recognition by BRCA1-BARD1 directs homologous recombination to sister chromatids. *Nat Cell Biol* **21**, 311–318 (2019).
218. Bai, C. B., Auerbach, W., Lee, J. S., Stephen, D. & Joyner, A. L. Gli2, but not Gli1, is required for initial Shh signaling and ectopic activation of the Shh pathway. *Development* **129**, 4753–4761 (2002).
219. Lukas, J. *et al.* Dazap2 modulates transcription driven by the Wnt effector TCF-4. *Nucleic Acids Res* **37**, 3007–3020 (2009).
220. Salghetti, S. E., Caudy, A. A., Chenoweth, J. G. & Tansey, W. P. Regulation of transcriptional activation domain function by ubiquitin. *Science* **293**, 1651–1653 (2001).

221. von der Lehr, N. *et al.* The F-box protein Skp2 participates in c-Myc proteasomal degradation and acts as a cofactor for c-Myc-regulated transcription. *Mol Cell* **11**, 1189–1200 (2003).
222. Structural Insight into Coordinated Recognition of Trimethylated Histone H3 Lysine 9 (H3K9me3) by the Plant Homeodomain (PHD) and Tandem Tudor Domain (TTD) of UHRF1 (Ubiquitin-like, Containing PHD and RING Finger Domains, 1) Protein* - Journal of Biological Chemistry. [https://www.jbc.org/article/S0021-9258\(20\)46676-1/fulltext](https://www.jbc.org/article/S0021-9258(20)46676-1/fulltext).
223. Rothbart, S. B. *et al.* Multivalent histone engagement by the linked tandem Tudor and PHD domains of UHRF1 is required for the epigenetic inheritance of DNA methylation. *Genes Dev* **27**, 1288–1298 (2013).
224. Kori, S. *et al.* Structure of the UHRF1 Tandem Tudor Domain Bound to a Methylated Non-histone Protein, LIG1, Reveals Rules for Binding and Regulation. *Structure* **27**, 485-496.e7 (2019).
225. Zepp, J. A. *et al.* Cutting edge: TNF receptor-associated factor 4 restricts IL-17-mediated pathology and signaling processes. *J Immunol* **189**, 33–37 (2012).
226. Marinis, J. M., Homer, C. R., McDonald, C. & Abbott, D. W. A novel motif in the Crohn's disease susceptibility protein, NOD2, allows TRAF4 to down-regulate innate immune responses. *J Biol Chem* **286**, 1938–1950 (2011).
227. Kalkan, T., Iwasaki, Y., Park, C. Y. & Thomsen, G. H. Tumor necrosis factor-receptor-associated factor-4 is a positive regulator of transforming growth factor-beta signaling that affects neural crest formation. *Mol Biol Cell* **20**, 3436–3450 (2009).
228. Kédinger, V. *et al.* Tumor Necrosis Factor Receptor-Associated Factor 4 Is a Dynamic Tight Junction-Related Shuttle Protein Involved in Epithelium Homeostasis. *PLOS ONE* **3**, e3518 (2008).
229. Ye, H., Park, Y. C., Kreishman, M., Kieff, E. & Wu, H. The structural basis for the recognition of diverse receptor sequences by TRAF2. *Mol Cell* **4**, 321–330 (1999).
230. Kim, C. M., Son, Y.-J., Kim, S., Kim, S. Y. & Park, H. H. Molecular basis for unique specificity of human TRAF4 for platelets GPIIb/IIIa and GPVI. *Proc Natl Acad Sci U S A* **114**, 11422–11427 (2017).
231. Ballinger, C. A. *et al.* Identification of CHIP, a novel tetratricopeptide repeat-containing protein that interacts with heat shock proteins and negatively regulates chaperone functions. *Mol Cell Biol* **19**, 4535–4545 (1999).
232. Qian, S.-B., McDonough, H., Boellmann, F., Cyr, D. M. & Patterson, C. CHIP-mediated stress recovery by sequential ubiquitination of substrates and Hsp70. *Nature* **440**, 551–555 (2006).

233. Jiang, J. *et al.* CHIP is a U-box-dependent E3 ubiquitin ligase: identification of Hsc70 as a target for ubiquitylation. *J Biol Chem* **276**, 42938–42944 (2001).
234. Zhang, M. *et al.* Chaperoned Ubiquitylation—Crystal Structures of the CHIP U Box E3 Ubiquitin Ligase and a CHIP-Ubc13-Uev1a Complex. *Molecular Cell* **20**, 525–538 (2005).
235. Connell, P. *et al.* The co-chaperone CHIP regulates protein triage decisions mediated by heat-shock proteins. *Nat Cell Biol* **3**, 93–96 (2001).
236. Ravalin, M. *et al.* Specificity for latent C-termini links the E3 ubiquitin ligase CHIP to caspases. *Nat Chem Biol* **15**, 786–794 (2019).
237. Ng, S. *et al.* Discovery and Structure-Based Design of Macrocyclic Peptides Targeting STUB1. *J. Med. Chem.* **65**, 9789–9801 (2022).
238. Tokareva, O. S. *et al.* Recognition and reprogramming of E3 ubiquitin ligase surfaces by α -helical peptides. *Nat Commun* **14**, 6992 (2023).
239. Clague, M. J., Urbé, S. & Komander, D. Breaking the chains: deubiquitylating enzyme specificity begets function. *Nat Rev Mol Cell Biol* **20**, 338–352 (2019).
240. Qiu, X.-B. & Goldberg, A. L. Nrdp1/FLRF is a ubiquitin ligase promoting ubiquitination and degradation of the epidermal growth factor receptor family member, ErbB3. *Proceedings of the National Academy of Sciences* **99**, 14843–14848 (2002).
241. Diamonti, A. J. *et al.* An RBCC protein implicated in maintenance of steady-state neuregulin receptor levels. *Proceedings of the National Academy of Sciences* **99**, 2866–2871 (2002).
242. Bouyain, S. & Leahy, D. J. Structure-based mutagenesis of the substrate-recognition domain of Nrdp1/FLRF identifies the binding site for the receptor tyrosine kinase ErbB3. *Protein Sci* **16**, 654–661 (2007).
243. Ben-Levy, R., Paterson, H. F., Marshall, C. J. & Yarden, Y. A single autophosphorylation site confers oncogenicity to the Neu/ErbB-2 receptor and enables coupling to the MAP kinase pathway. *EMBO J* **13**, 3302–3311 (1994).
244. Fazioli, F. *et al.* The erbB-2 mitogenic signaling pathway: tyrosine phosphorylation of phospholipase C-gamma and GTPase-activating protein does not correlate with erbB-2 mitogenic potency. *Mol Cell Biol* **11**, 2040–2048 (1991).
245. Peles, E., Levy, R. B., Or, E., Ullrich, A. & Yarden, Y. Oncogenic forms of the neu/HER2 tyrosine kinase are permanently coupled to phospholipase C gamma. *EMBO J* **10**, 2077–2086 (1991).
246. Wauman, J., De Ceuninck, L., Vanderroost, N., Lievens, S. & Tavernier, J. RNF41 (Nrdp1) controls type 1 cytokine receptor degradation and ectodomain shedding. *J Cell Sci* **124**, 921–932 (2011).

247. Wang, C. *et al.* The E3 ubiquitin ligase Nrdp1 ‘preferentially’ promotes TLR-mediated production of type I interferon. *Nat Immunol* **10**, 744–752 (2009).
248. Tullett, K. M. *et al.* RNF41 regulates the damage recognition receptor Clec9A and antigen cross-presentation in mouse dendritic cells. *Elife* **9**, e63452 (2020).
249. De Ceuninck, L., Wauman, J., Masschaele, D., Peelman, F. & Tavernier, J. Reciprocal cross-regulation between RNF41 and USP8 controls cytokine receptor sorting and processing. *J Cell Sci* **126**, 3770–3781 (2013).
250. Qiu, X.-B., Markant, S. L., Yuan, J. & Goldberg, A. L. Nrdp1-mediated degradation of the gigantic IAP, BRUCE, is a novel pathway for triggering apoptosis. *EMBO J* **23**, 800–810 (2004).
251. Pearson, G. *et al.* Clec16a, Nrdp1, and USP8 Form a Ubiquitin-Dependent Tripartite Complex That Regulates β -Cell Mitophagy. *Diabetes* **67**, 265–277 (2018).
252. Zhong, L., Tan, Y., Zhou, A., Yu, Q. & Zhou, J. RING Finger Ubiquitin-Protein Isopeptide Ligase Nrdp1/FLRF Regulates Parkin Stability and Activity*. *Journal of Biological Chemistry* **280**, 9425–9430 (2005).
253. Wu, X., Yen, L., Irwin, L., Sweeney, C. & Carraway, K. L. Stabilization of the E3 ubiquitin ligase Nrdp1 by the deubiquitinating enzyme USP8. *Mol Cell Biol* **24**, 7748–7757 (2004).
254. Avvakumov, G. V. *et al.* Amino-terminal dimerization, NRDP1-rhodanese interaction, and inhibited catalytic domain conformation of the ubiquitin-specific protease 8 (USP8). *J Biol Chem* **281**, 38061–38070 (2006).
255. Gingerich, M. A. *et al.* Reciprocal regulatory balance within the CLEC16A-RNF41 mitophagy complex depends on an intrinsically disordered protein region. *J Biol Chem* **299**, 103057 (2023).
256. Masschaele, D. *et al.* High-Confidence Interactome for RNF41 Built on Multiple Orthogonal Assays. *J Proteome Res* **17**, 1348–1360 (2018).
257. Masschaele, D. *et al.* RNF41 interacts with the VPS52 subunit of the GARP and EARP complexes. *PLoS One* **12**, e0178132 (2017).
258. Rolland, T. *et al.* A proteome-scale map of the human interactome network. *Cell* **159**, 1212–1226 (2014).
259. Kosinsky, R. L. *et al.* USP22 Suppresses SPARC Expression in Acute Colitis and Inflammation-Associated Colorectal Cancer. *Cancers (Basel)* **13**, 1817 (2021).
260. McCann, J. J. *et al.* USP22 Functions as an Oncogenic Driver in Prostate Cancer by Regulating Cell Proliferation and DNA Repair. *Cancer Res* **80**, 430–443 (2020).

261. Melo-Cardenas, J., Zhang, Y., Zhang, D. D. & Fang, D. Ubiquitin-specific peptidase 22 functions and its involvement in disease. *Oncotarget* **7**, 44848–44856 (2016).
262. Zhang, X.-Y. *et al.* The putative cancer stem cell marker USP22 is a subunit of the human SAGA complex required for activated transcription and cell-cycle progression. *Mol Cell* **29**, 102–111 (2008).
263. Zhao, Y. *et al.* A TFTC/STAGA module mediates histone H2A and H2B deubiquitination, coactivates nuclear receptors, and counteracts heterochromatin silencing. *Mol Cell* **29**, 92–101 (2008).
264. Ingvarsdottir, K. *et al.* H2B ubiquitin protease Ubp8 and Sgf11 constitute a discrete functional module within the *Saccharomyces cerevisiae* SAGA complex. *Mol Cell Biol* **25**, 1162–1172 (2005).
265. Köhler, A. *et al.* The mRNA export factor Sus1 is involved in Spt/Ada/Gcn5 acetyltransferase-mediated H2B deubiquitinylation through its interaction with Ubp8 and Sgf11. *Mol Biol Cell* **17**, 4228–4236 (2006).
266. Voutsinos, V. *et al.* A near-complete map of human cytosolic degrons and their relevance for disease. *Sci Adv* **12**, eadz3483 (2026).
267. Larsen, F. B. *et al.* Comprehensive degron mapping in human transcription factors. 2025.05.16.654404 Preprint at <https://doi.org/10.1101/2025.05.16.654404> (2025).
268. Wang, X. *et al.* UbiBrowser 2.0: a comprehensive resource for proteome-wide known and predicted ubiquitin ligase/deubiquitinase-substrate interactions in eukaryotic species. *Nucleic Acids Res* **50**, D719–D728 (2022).
269. Li, J. *et al.* Proteome-wide mapping of short-lived proteins in human cells. *Molecular Cell* **81**, 4722–4735.e5 (2021).
270. Balaji, V. *et al.* A dimer-monomer switch controls CHIP-dependent substrate ubiquitylation and processing. *Mol Cell* **82**, 3239–3254.e11 (2022).
271. Vogel, A., McWhorter, M., Kwok, E. & Nyarko, A. Multivalent AMOTL1 Engages NEDD4-1 and KIBRA Through Distinct Cooperative Binding Mechanisms. *Journal of Molecular Biology* **438**, 169652 (2026).
272. Iconomou, M. & Saunders, D. N. Systematic approaches to identify E3 ligase substrates. *Biochem J* **473**, 4083–4101 (2016).
273. Suiter, C. C. *et al.* Combinatorial mapping of E3 ubiquitin ligases to their target substrates. *Mol Cell* **85**, 829–842.e6 (2025).
274. Myers, N. E. M. *et al.* Identification of ZFTA as a Novel KLHL20 Substrate and Mechanistic Insights Into Fuzzy Binding of Disordered Peptides via Biosensor Analysis and Computational Modelling. *Chembiochem* **27**, e70237 (2026).

275. Wulf, A. *The Invention of Nature: Alexander Von Humboldt's New World*. (Knopf Doubleday Publishing Group, Westminster, 2015).

Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 2662

Editor: The Dean of the Faculty of Science and Technology

A doctoral dissertation from the Faculty of Science and Technology, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology. (Prior to January, 2005, the series was published under the title “Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology”.)

Distribution: publications.uu.se
urn:nbn:se:uu:diva-583655



ACTA UNIVERSITATIS
UPSALIENSIS
2026