UPPSALA
UNIVERSITET

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 2204

Exploration of analytical methods to study motif-mediated host-virus protein-protein interactions

ESZTER KASSA

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2022

ISSN 1651-6214
ISBN 978-91-513-1633-8
URN urn:nbn:se:uu:diva-486917

Dissertation presented at Uppsala University to be publicly examined in BMC A1:111, Husargatan 3, Uppsala, Thursday, 8 December 2022 at 09:15 for the degree of Doctor of Philosophy. The examination will be conducted in English. Faculty examiner: Gilles Travé (University of Strasbourg).

Abstract

Kassa, E. 2022. Exploration of analytical methods to study motif-mediated host-virus protein-protein interactions. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology* 2204. 74 pp. Uppsala: Acta Universitatis Upsaliensis. ISBN 978-91-513-1633-8.

Proteins are responsible for countless processes in living creatures, but most often they do not perform these tasks alone. Rather, they engage in interactions with other proteins, creating whole protein-protein interaction (PPI) networks. Some of these interactions are formed between a folded protein domain and a short linear motif (SLiM), which is a small, 3-10 amino acid long stretch usually in the intrinsically disordered regions of proteins. These interactions tend to be low-to-medium affinity and transient, therefore their capture requires special tools. Furthermore, viruses often hijack the human cellular machinery through PPIs as they have limited genomes and are obligate cellular parasites. Therefore, the investigation of viral-host PPIs is of great importance and can lead to the development of novel antivirals.

In my thesis, I used mostly peptide-based and mass spectrometry (MS) techniques to validate and further explore motif-based PPIs. The main objectives were to: i) evaluate and compare synthetic peptide-based pulldown approaches, ii) validate and further explore the interaction between viral peptides and human polyadenylate-binding protein (PABP) using green fluorescent protein (GFP)-tagged peptide repeats, iii) confirm interactions, define and refine human interaction motifs that engage in interactions with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) proteins by employing peptide SPOT (synthetic peptide arrays on membrane support technique) arrays and alanine scanning, iv) investigate the change in the interactome of the nuclear pore complex protein 153 (NUP153) between uninfected and tick-borne encephalitis virus (TBEV)-infected states using GFP-tagged full-length protein for pulldown.

First, we explored the potential of affinity purification-mass spectrometry (AP-MS) and protein interaction screen on peptide matrix (PRISMA) to capture SLiM-based PPIs. The peptide pulldown approach proved to be more applicable over a wide range of affinities and interactions, however, protein concentration and the local concentration of presented motifs were limiting factors in certain cases. We then investigated SLiM-based interactions between RNA-viruses and human proteins. Here, using green fluorescent-peptide pulldowns I confirmed the interaction between viral peptides and the human poly-A binding protein. Next, we uncovered that some human SLiMs interact with SARS-CoV-2 proteins, and I was able to highlight the interaction motif using peptide arrays when only a handful of peptides were available. Lastly, I identified different enriched proteins in NUP153-pulldowns from mock-infected and TBEV-infected cell lysate, that were complementary to the changes observed with other techniques.

In conclusion, I explored a range of techniques that are valuable in the validation of PPIs, which is crucial in combination with high-throughput approaches. As more and more SLiM-based interactions are explored and predicted, the value of these tools continues to increase.

Keywords: protein-protein interaction, short linear motif, affinity purification-mass spectrometry, virus, SARS-CoV-2, TBEV

Eszter Kassa, Department of Chemistry - BMC, Analytical Chemistry, Box 599, Uppsala University, SE-75124 Uppsala, Sweden.

ISSN 1651-6214
ISBN 978-91-513-1633-8
URN urn:nbn:se:uu:diva-486917 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-486917)

Anyának, Apának és Johannak

The cover art was made by David Ribar
using the crystal structure PDB ID 7K28.

List of Papers

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

I. **Kassa, E**., Jamshidi, S., Mihalič, F., Simonetti, L., Kliche, J., Jemth, P., Bergström Lind, S., and Ivarsson, Y. (2022) Evaluation of affinity-purification coupled to mass spectrometry approaches for capture of short linear motif-based interactions. https://doi.org/10.1101/2022.10.19.512833. *Submitted Manuscript*

II. Mihalič, F., Simonetti, L., Giudice, G., Sander, M. R., Lindqvist, R., Akprioro Peters, M. B., Benz, C., **Kassa, E.**, Badgujar, D., Inturi, R., Ali, M., Krystkowiak, I., Sayadi, A., Andersson, E., Aronsson, H., Söderberg, O., Dobritzsch, D., Petsalaki, E., Överby, A. K., Jemth, P., Davey, N. E., and Ivarsson, Y. (2022) Large-scale phage-based screening reveals extensive pan-viral mimicry of host short linear motifs. *bioRxiv*. 10.1101/2022.06.19.496705 *Submitted manuscript*

III. Mihalič, F., Benz, C., **Kassa, E**., Lindqvist, R., Simonetti, L., Inturi, R., Aronsson, H., Andersson, E., Chi, C. N., Davey, N. E., Överby, A. K., Jemth, P., and Ivarsson, Y. (2022) Identification of motif-based interactions between SARS-CoV-2 protein domains and human peptide ligands pinpoint antiviral targets. *bioRxiv*. 10.1101/2022.10.07.511324 *Submitted manuscript*

IV. Peters, M. B. A., **Kassa, E**., Yau, W-L., Lindqvist, R., Nilsson, E., Siljedahl, M., Ivarsson, Y., Lundmark, R., and Överby, A. K. (2022) Tick-borne flaviviruses recruits the pro viral factor NUP153 to the replication site. *Manuscript*

The following publications are not part of this thesis.

I. Valdés, A., Bitzios, A., **Kassa, E.**, Shevchenko, G., Falk, A., Malmström, P-U., Dragomir, A., Segersten, U., and Bergström Lind, S. (2021) Proteomic comparison between different tissue preservation methods for identification of promising biomarkers of urothelial bladder cancer. *Sci Rep.* 11(1):7595

II. Benz, C., **Kassa, E.**, Tjärnhage, E., Bergström Lind, S., and Ivarsson, Y. (2020) Chapter 1: Identification of Cellular Protein–Protein Interactions, in Inhibitors of Protein–Protein Interactions: Small Molecules, Peptides and Macrocycles, *RSC*. pp. 1-39 DOI: 10.1039/9781839160677-00001

Contribution to manuscripts

I. Planned the experiments and carried out most of them. Data analysis. Main responsible for preparing the manuscript.

II. Took part in planning the experiment. Performed pulldowns and analysed data. Took part in writing and editing the manuscript.

III. Took part in planning the experiment. Performed pulldowns and analysed data. Took part in editing the manuscript.

IV. Took part in planning the experiment. Performed pulldowns and analysed data. Took part in writing and editing the manuscript.

Contents

Abbreviations

AA	amino acid
AE	affinity enrichment
ANOVA	analysis of variance
AP-MS	affinity purification-mass spectrometry
C	capsid protein
CEM	chain ejection model
CID	collision-induced dissociation
Co-IP	co-immunoprecipitation
CompPASS	comparative proteomic analysis software suite
CRAPome	contaminant repository for affinity purification
CRISPR	clustered regularly interspaced short palindromic repeats
CRM	charge residue model
DMT	dimethyl labelling
DNA	deoxyribonucleic acid
DTT	dithiothreitol
ECD	electron capture dissociation
EGFP	enhanced green fluorescent protein
ELM	eukaryotic linear motif
ER	endoplasmic reticulum
ESCRT	endosomal sorting complexes required for transport
ESI	electrospray ionization
ETD	electron transfer dissociation
FDR	false discovery rate
FP	fluorescence polarization
GFP	green fluorescent protein
HCD	high-energy collisional dissociation
HCl	hydrochloric acid
HCoV229E	human coronavirus 229E
HEK293	human embryonic kidney 293
HeV	hendra virus
HIV	human immunodeficiency virus
HPLC	high-performance liquid chromatography
HSQC	heteronuclear single quantum coherence
IAA	iodoacetamide
ICAT	isotope-coded affinity tag

IDP	intrinsically disordered protein
IDR	intrinsically disordered region
IEM	ion evaporation model
ITC	isothermal calorimetry
iTRAQ	isobaric tags for relative and absolute quantitation
K_D	dissociation constant
kDa	kilodalton
LC-MS/MS	liquid chromatography-tandem mass spectrometry
LFQ	label-free quantification
LGTV	langat virus
MiST	mass spectrometry interaction statistics
MDM2	E3 ubiquitin-protein ligase/mouse double minute 2 homolog
MOI	multiplicity of infection
MRBL-pep	microspheres with ratiometric barcode lanthenide encoding coupled to peptides
mRNA	messenger ribonucleic acid
MS	mass spectrometry
nanoESI	nano electrospray ionization
NMR	nuclear magnetic resonance
NPC	nuclear pore complex
NSP, NS	nonstructural protein
NUP	nuclear pore complex protein
ORF	open reading frame
PABC	polyadenylate-binding protein 1 C-terminal domain
PABPi	PABP inhibitor
PBS	phosphate buffered saline
PLA	proximity ligation assay
POI	protein of interest
PPI	protein-protein interaction
PRISMA	protein interaction screen on peptide matrix
ProP-PD	proteomic peptide-phage display
RNA	ribonucleic acid
RRM	RNA-binding domain
SAINT	significance analysis of interactome
SARS-CoV-2	severe acute respiratory syndrome coronavirus 2
SDS	sodium dodecyl sulfate
SILAC	stable isotope labelling by amino acids in cell culture
SLiM	short linear motif
SPOT	synthetic peptide arrays on membrane support technique
SPR	surface plasmon resonance
ssRNA	single-stranded RNA
TBEV	tick-borne encephalitis virus
TP53	cellular tumor antigen p53
TSG101	tumor susceptibility gene 101 protein

TMT	tandem mass tag
YFP	yellow fluorescent protein

Amino acids

A	Alanine
C	Cysteine
D	Aspartic acid
E	Glutamic acid
F	Phenylalanine
G	Glycine
H	Histidine
I	Isoleucine
K	Lysine
L	Leucine
M	Methionine
N	Asparagine
P	Proline
Q	Glutamine
R	Arginine
S	Serine
T	Threonine
V	Valine
W	Tryptophan
Y	Tyrosine

SLiM symbol	**Meaning**[1][2]
A	The given amino acid (here A)
x	Unknown or any amino acid
[ABC]	Either A, B, or C
[^A]	Amino acid A is excluded at this position. (Also for multiple amino acids.)
Φ	Hydrophobic amino acid
(ABC)	A, B and C in any order

Introduction

In the effort to understand how living cells function in health and disease, scientists have long been investigating biomolecules such as proteins, lipids, DNA, RNA, and metabolites. Many functions, such as signal transduction, transport, and repair in the human cell are controlled by the intricate balance of abundance and interactions of proteins. Elucidation of protein-protein interactions is crucial for the fundamental understanding of biological processes, and therefore many techniques have been developed for the discovery and validation of them.

The focus of this thesis is to explore methods for studying short linear motif (SLiM)-based protein-protein interactions (PPIs) using proteomics tools, mainly mass spectrometry. These interactions typically engage a folded domain and a short (3-10 amino acid long), flexible, or intrinsically disordered part of another protein [3], which results in usually low-to-medium affinity and transient interactions [4]. SLiM-based interactions have great biological significance, as they are often involved in e.g. signalling [5], where exactly this kind of interaction is needed. SLiMs are also important in the context of the takeover of the host cell by viruses [6][7][8], therefore investigating these interactions can lead to further understanding and new therapies for viral infections.

Despite their importance, only a fraction of the predicted hundred thousand to one million SLiM-based interactions is known today [9], as the SLiM-based interactions often escape capture by many methods traditionally used for high-throughput interaction screening [3].

During my PhD I have investigated various methods to capture proteins that interact with SLiM-containing peptides and identify the enriched proteins them using mass spectrometry-based bottom-up proteomics (**Papers I, II and IV**). I also worked on determining key residues (motifs) in human peptides binding viral protein domains (**Paper III**). This endeavour also included working on elucidating protein-protein interactions of different viruses, such as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (**Paper II and III**) and tick-born encephalitis (TBEV) (**Paper IV**).

In the following sections, I give a background to the field and introduce the methods used before describing my contribution to the field.

Proteins and interactions

Protein-protein interactions

Most often proteins are not fulfilling their functions alone, but they take part in a number of interactions with other proteins (and other molecules). These protein-protein interactions are responsible for a wide array of cellular processes, such as signalling, transport, expression of proteins and protein degradation.

PPIs can be characterized in many ways [10][11]. If we look at the participants in the interaction, it can be homo-oligomeric (multiple copies of the same protein form a complex), or hetero-oligomeric, when different proteins interact. Based on the duration or dynamics of the interaction, we can talk about stable or permanent interactions (such as the ribosome, where the protein complex is mostly permanent) or transient ones, such as most of the motif-mediated interactions (described below). PPIs can be obligate or non-obligate, depending on if the interacting proteins can exist on their own *in vivo*. The interactions can be further described by the affinity of the interaction. For this, the dissociation constant (K_D) of the process of complex forming reaction is used (where the protein (P) and the ligand (L) form a complex (PL), Reaction 1). The higher the affinity, the lower the K_D value. The unit of K_D is M.

$$P + L \leftrightarrows PL \qquad \text{(Reaction 1)}$$

$$K_D = \frac{[P][L]}{[PL]} \qquad \text{(Equation 1)}$$

Protein interaction networks are often altered in the case of disease. Studying them can reveal the mechanism of disease progression, and the key interactions can be druggable targets [12][13].

Intrinsic disorder

Correct protein folding is often crucial for protein function. The structure taken up by a protein depends on its amino acid sequence, with more hydrophobic residues often getting packed on the inside of globular structures and charged residues facing outwards [14]. Many of the folded globular protein domains have specific binding pockets for other ligands. However, regions without a defined three-dimensional structure can also have a wide range of functions [15]. These regions are called intrinsically disordered regions (IDRs) and they can constitute a part of the protein, wedged in between structured domains, or can build up whole intrinsically disordered proteins (IDPs). It has been estimated that 35-50% of the human proteome is disordered [16][17].

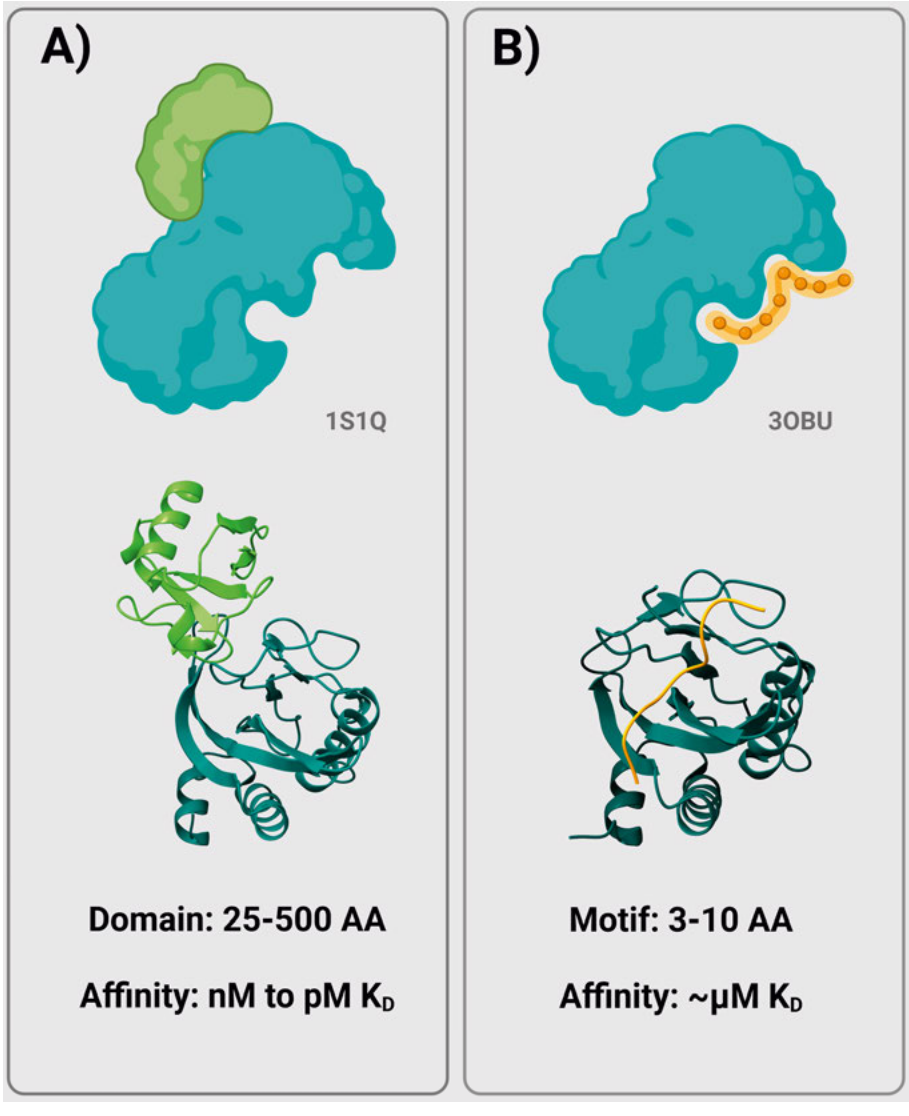


Figure 1: Protein interactions. A) Interaction between globular domains, exemplified by the interaction of human TSG101 UEV domain (dark teal) and the human ubiquitin (lime green). In this case bigger, globular domains interact, the interaction surface is bigger which leads to higher affinity interactions. B) Interaction between a globular domain and a peptide, containing a short linear motif (SLiM) exemplified by the interaction of human TSG101 UEV domain (dark teal) and the HIV-1 $Gag_{453-461}$ peptide (orange). In this case, a domain and a short motif interact, where only a few residues take part in forming the interaction. The interaction surface is small, which leads to lower affinity interactions. Created with BioRender.com

Short Linear Motifs (SLiMs)

A subgroup of PPIs occurs when a globular protein domain engages in the interaction with only a short linear stretch of another protein [3]. These amino acid stretches are called interaction motifs or short linear motifs (SLiMs). SLiMs are 3-10 amino acid long stretches in the protein, which are capable of interactions, sometimes only requiring a few key amino acid residues. SLiMs are most often found in the intrinsically disordered regions of proteins [3].

Because the interaction surface is quite small, these motif-based interactions are often of low affinity, with K_D values in the micromolar range [18] and transient, which means the interactions can be short-lived (Figure 1). This can be advantageous in for example signalling, where quick reactions are needed.

When surrounded by unstructured regions, SLiMs are often more conserved than their surrounding amino acids, and the conservation thus indicates a functional region [19]. While mutating or removing the key residues of a SLiM results in the disruption of the interaction, flanking regions surrounding the motif can be also of great importance, just as well as the disordered nature of the protein in general [3][20][21]. Proteins often have multiple copies of the same SLiM, or several different kinds of SLiMs in close vicinity, which can lead to several consequences [22]. When different SLiMs are located very close to each other in the protein sequence it can result in the interactions via these SLiMs being mutually exclusive [23]. Having multiple copies of the same SLiMs that bind to an array of domains in another protein can result in higher apparent affinity through multivalency and avidity effects [4].

In the effort to discover SLiMs, different prediction methods/software tools are often used. As the motifs are often conserved through species, conserved residues of intrinsically disordered parts of proteins can be used to predict functional motifs using for example SLiMPrints [24]. The recently developed AlphaFold2 [25] structure prediction tool is further expanding the prediction horizon by predicting folded structures and IDR more precisely [26].

Currently, the Eukaryotic Linear Motif (ELM) [27] resource provides information about SLiMs with a collection of known motifs with experimental validation. In this database, there are 2056 true positive human motif instances, which is far below the predicted number of a hundred thousand to a million [9]. This highlights the need for investigating motif-based interactions further.

An applied aspect of SLiM-based interactions is that they can be targeted by drugs that block the SLiM-binding pockets. The development of inhibitors might start from the structure of the domain and the binding pocket [28]. A group of such examples is MDM2 inhibitors (e.g. Nutlins, RG7112, and Idasanutlin) which occupy the same binding pocket as the TP53 interaction peptide. They have been shown to inhibit the TP53-MDM2 interactions [29] and Idasanutlin demonstrated antitumor activity [30].

Viruses use motif mimicry for hijacking the host cell

Viruses do not have all the tools they need for reproduction and proliferation, therefore they are obligate intracellular parasites, relying on hijacking the host cell machinery for their replication [31][32]. This is accomplished through host-virus PPIs, which may involve the binding to host domains through viral SLiMs [6]. Viral SLiMs may engage in higher affinity binding with the target protein than the host SLiM, thereby outcompeting endogenous interactions, or may outcompete the host by a high concentration of viral proteins at a later stage of infection [6]. One example of SLiM-mediated viral hijacking is the PTAP motif in HIV-1 gag polyprotein, which engages in interaction with the UEV domain of human TSG101, thereby recruiting the ESCRT(endosomal sorting complexes required for transport)-I complex for viral budding [33][34][35].

Viruses have small and compact genomes, for example, the SARS-CoV-2 genome is 0.029903 megabasepair long (SARS-CoV-2 reference genome ASM985889v3 [36]) coding for only 12 proteins (29 after processing [37]) while the human genome has 3298.43 megabasepairs (human reference genome GRCh38.p14 [38]) coding for more than 20 000 proteins. The use of SLiMs is an elegant solution requiring only a small part of the proteome, and the gain of motif can be a result of one or a few, *ex nihilo* point mutations [39]. Viruses have a high mutation rate, as exemplified by the 2175 different SARS-CoV-2 lineages identified since the start of the pandemic (from cov-lineages.org) [40]. Thus, viral SLiMs may evolve quickly.

As presented above SLiMs often play an important role in viral infection. Therefore, discovering and studying host-pathogen motif-mediated interactions are valuable for developing novel therapies.

Viruses of relevance for this thesis

As viruses often employ SLiMs and are generally reliant on hijacking host protein-protein interactions, I investigated virus-host interactions. I also explored the shift of the PPIs upon infection. Here I briefly describe different viruses relevant for my research.

Betacoronavirus. SARS-CoV-2 is the virus responsible for the ongoing Covid-19 pandemic.

Alphacoronavirus. Human coronavirus 229E (HCoV-229E) causes respiratory symptoms like common cold, or more serous conditions [41].

Henipavirus. The Hendra virus (HeV) is bat-borne, affecting humans and horses mainly in Australia and neighbouring regions [42].

Flavivirus. TBEV infections are most common in Eastern Europe and Russia, yet the Southern part of Sweden is also affected [43]. The mortality rate

in the case of TBEV is between 1-20% (depending on the subtype), however, up to 46% of patients live with long-term health effects [44]. Langat virus (LGTV) is closely related to TBEV with a high sequence match, but it is less pathogenic and not associated with human disease [43][45].

Methods of capturing, validating, and characterizing protein-protein interactions

Proteome-wide discovery of SLiMs by phage display

Phages (or bacteriophages) are viruses that infect and replicate in bacteria. Using genetic editing, peptides and proteins can be displayed fused to the outer coat proteins of the phage, generating a sort of affinity particle. Phage display was first developed by George Smith in 1985 [46], focusing on high-affinity interactions. His discovery was awarded the Nobel Prize in Chemistry in 2018, also recognizing the pioneering work of Sir Gregory Winter [47].

A commonly used phage for phage display is the filamentous M13 phage. The most widely used coat proteins of this rod-shaped phage are the minor coat protein p3 and the major coat protein p8. P3 is displayed five times on the shorter end of the phage, while there are around 2700 copies of the main coat protein p8 that makes up the capsid [48]. When p3 is used for display, between one and five copies of the fused sequence are presented on the phage depending on the system used, therefore this approach is more suitable for identifying strong binders. Employing the p8 protein, provides a highly multivalent display, as around one-third of the 2700 copies of p8 protein will display the fusion sequence [49]. The multivalent display allows for the capture of lower affinity interactions. The libraries using this technique can reach the size of 10^9-10^{10} individual peptide or protein sequences resulting in extremely high throughput with a simple and cost-effective method [50].

Over the last decade proteomic peptide-phage display (ProP-PD) [51][52][53], has been developed into a mature approach for large-scale identification of SLiM-based interactions. In ProP-PD) libraries are designed based on a certain part of the proteome (e.g. the intrinsically disordered region of the human proteome) [53]. The library is then used in selection against immobilized full-length proteins or protein domains. Non-binding phages are washed off, and the bound ones are eluted and proliferated by bacteria. The process is repeated 3-5 times, and the sequence of the enriched displayed peptides are revealed using next-generation sequencing. The ProP-PD approach provides an efficient way of proteome-wide screening of host-host or host-virus SLiM-based interactions [51][52][53].

A limiting factor is that the bait proteins have to be individually expressed and purified and that the setup does not resemble the real biological environment where the interactions would happen.

Using this high-throughput approach provides a high number of candidate interactions that most often require further validation. This is typically done by affinity determinations. Additional validation using orthogonal methods can be done with for example affinity purification and co-localization assays (such as proximity ligation assay (PLA)). It is also possible to identify key residues in a peptide when the results from ProP-PD do not provide sufficiently clear enrichment. Experimental methods allowing to carry out these validation and discovery steps are described below.

Methods for affinity determinations of SLiM-based interactions

There are several methods to determine affinities of protein-protein interactions, such as fluorescent polarization (FP) [54][55], isothermal calorimetry (ITC) [56], surface plasmon resonance (SPR) [57], MRBLE-pep (microspheres with ratiometric barcode lanthenide encoding coupled to peptides) [20], and holdup assay [58] and its native version where interaction are measured in cell extract [59].

Here I present the background of FP, which was used for the results included in Papers I-III.

Fluorescent polarization

FP measurements are often used to determine the dissociation constant (K_D) of protein-peptide interactions [60][61]. The technique uses polarized light to excite a fluorescent analyte and the emitted light is recorded. When a small, fluorophore-labelled molecule is in solution it rotates fast and if it is excited with polarized light, the emitted light will be depolarized. However, if this small molecule interacts with e.g. a bulky protein the rotation slows down, and the emitted polarized light stays polarized (Figure 2). By analysing the emitted light from solutions where different concentrations of proteins are added to the labelled peptide the K_D can be determined. By changing the setup, competition experiments can be carried out as well, to determine affinities for unlabelled peptides.

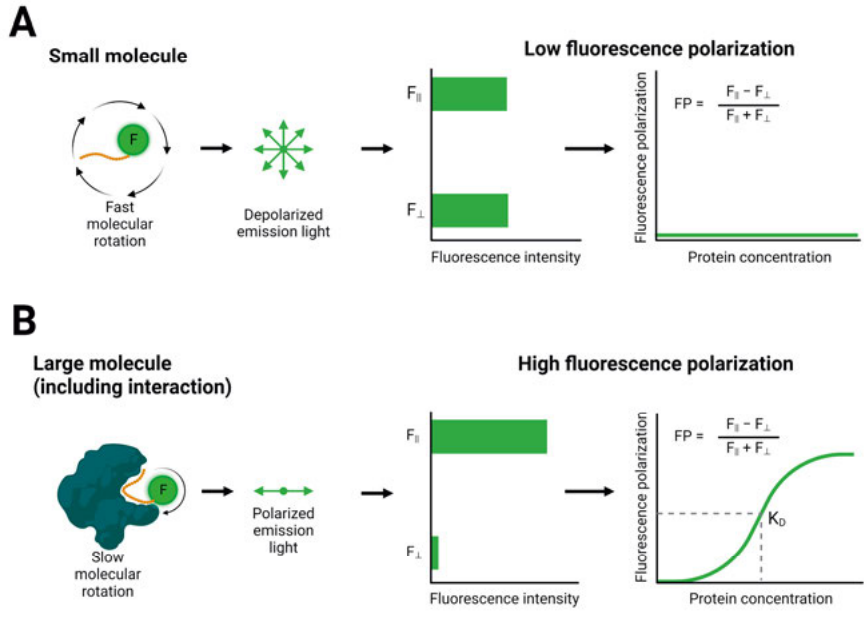


Figure 2: FP assay. A) A small, fluorescently labeled molecule quickly rotates in solution. After excitation with polarized light, the fluorescent tag will emit depolarized light because of the fast rotation. B) The small, fluorescently labelled molecule engages in interaction with a larger molecule, the rotation of the interaction complex is slower, than that of the small molecule. Therefore, the emitted light following excitation will be polarized. Adapted from "Fluorescence Polarization Assay", by BioRender.com (2020). Retrieved from https://app.biorender.com/biorender-templates

Validation of binding and validation of key residues using peptide arrays

It is sometimes needed to validate the binding of predicted SLiMs or to identify key residues for binding in a given peptide sequence. For this purpose, peptide arrays can be very useful. Array-based techniques involve the screening of a single prey protein or even whole cell lysate against immobilized putative interaction partners – such as peptides [62].These bait molecules can be immobilized or synthesized on cellulose membranes (e.g. with the SPOT-synthesis technique) [62] or on a chip-like glass microarray [63].

Alanine scanning (where one by one each residue in the peptide is changed to alanine) is often used to systematically map the effects of truncation mutations along a peptide sequence (see for example [64][65][66]). When the array is incubated with the interacting protein, the interaction forms between the protein and those peptides where the key residues are not mutated. The non-binding proteins are removed and the detection of protein binding is carried

out with a protein-specific, HRP-conjugates antibody using chemiluminescence (Figure 3). This results in the direct identification of the key binding residues in the peptide. A drawback of this approach, depending on the bait synthesis, is that the synthesized peptides cannot be purified, therefore non-specific and unwanted interactions might arise.

In **Paper III** cellulose-based SPOT arrays were used to elucidate key residues (motif) in peptide sequences previously shown by ProP-PD to interact with SARS-CoV-2 proteins.

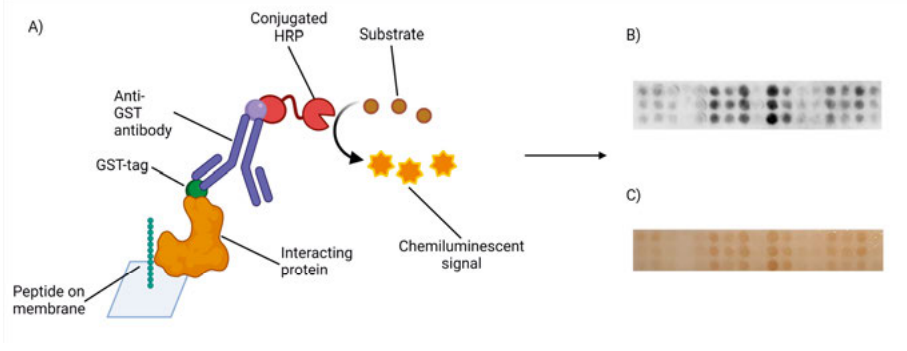


Figure 3: The process and readout of the SPOT-arrays. A) Scheme of the GST-tagged protein of interest binding to the peptide on a cellulose membrane. The HRP-conjugated antibody recognizes the GST-tag and a chemiluminescent signal is produced by the HRP after adding the substrate. B) Chemiluminescent readout of the membrane. C) Naked eye visualization of the membrane. Created with BioRender.com

Proximity-based methods using antibodies to validate co-localization

Proximity ligation assay (PLA)

To validate protein-protein interactions based on proximity, the proximity ligation assay (PLA) can be used [67][68]. In this method, protein-specific antibodies are used for the two interaction partners, then secondary antibodies with special oligonucleotide conjugates are added. If the two proteins are in close enough proximity (<40 nm), the two conjugated oligonucleotides can hybridize with an added connector. These then are ligated and the resulting DNA molecule can be amplified and fluorescently visualized using detection oligos.

Affinity-capture methods for PPI validation

There are multiple approaches to capturing the interactome of a protein of interest (bait) using affinity matrices against the bait or a specific tag, as detailed in the following section.

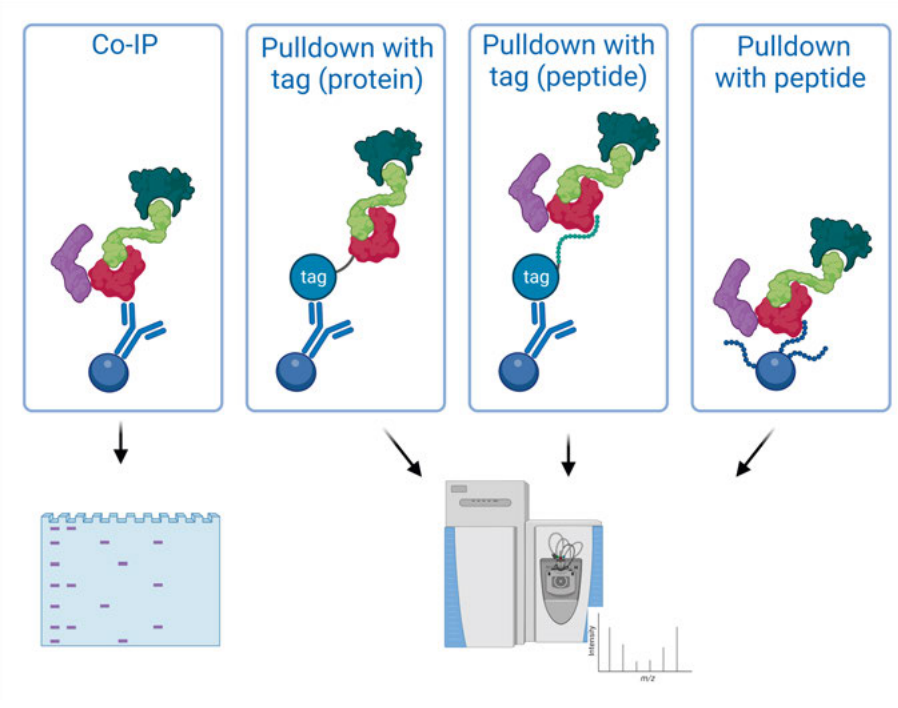


Figure 4: Different approaches to affinity-based pulldowns and common techniques used for protein identification downstream. First panel: co-IP using a specific antibody against the protein of interest (POI) and its interaction partners. Readout of pulled-down proteins is done by Western blot. Second panel: a tagged construct of the POI is pulled down with its interaction partners using a tag-specific antibody. Third panel: The fusion construct of tag and peptide is pulled down with a tag-specific antibody. Fourth panel: immobilized peptides are pulling down the POI and its interactors. In the last three cases, it is common to use mass spectrometry to identify the pulled-down proteins. Created with BioRender.com

Co-IP

In co-immunoprecipitation (co-IP) (Figure 4, first panel) an antibody (which can be immobilized on the surface of a bead) is used against the bait protein in e.g. a cell lysate, which allows the enrichment of binary interactors as well as other proteins that form complexes or engage in interaction with these interactors [69]. In this approach, a specific antibody is needed for each bait protein, which can quickly become costly with a higher number or baits. The

readout is often done by Western blot, where again, specific antibodies are needed and the results depend on their specificity.

Affinity purification for mass spectrometry analysis

In affinity purification coupled to mass spectrometry (AP-MS) approaches a variety of affinity matrices may be used, for example, an antibody against a genetically fused protein tag or synthesized peptides as potential interactors [70]. When using AP-MS for low affinity and transient interactions it is crucial to apply gentle washing steps after capturing the interactions and before eluting the proteins for further analysis, otherwise, these interactions can be easily lost. A great advantage of the mass spectrometry-based techniques is that they are unbiased, meaning that no previous knowledge is needed about the sample (e.g. what proteins are expected) for the detection and identification of proteins, it is only needed to use the correct database. This makes MS-based techniques a great discovery tool and allows for further validation of the interactions. However, this way we cannot gain direct information on binary interactions, as whole protein complexes and interactors of interactors can be pulled down by a given bait [71]. Large-scale studies to discover the interactome have been carried out, such as the BioPlex, where the most recent version (BioPlex 3.0, [72]) contains 118,162 interactions generated using 10,128 C-terminal FLAG-HA-tagged proteins as baits. Similarly, Hein et al. [73] used 1,125 GFP-tagged proteins as baits to discover 28,500 interactions.

As quantification methods and instruments have been developed, it is possible to aim for affinity enrichment (AE) rather than affinity purification (AP). In this approach, the goal is not to remove all non-specific binders but to qualitatively and quantitatively enrich real interactors. By quantifying protein levels between the sample and control, the enriched proteins can be identified [74].

There are many commonly used fusion tags for pulldowns, such as short epitope tags like FLAG [75], HA (human influenza hemagglutinin) [76], or c-myc [77], whereas larger molecules are also employed, such as GFP (green fluorescent protein) [78] or GST (glutathione S-transferase) [79]. When these tags are fused to a protein or peptide, many parameters should be considered, for example, if the tag should be on the N- or C-terminus, what expression level of the construct should be reached, and so on. There is also a risk that the chosen protein or peptide sequence is toxic for the cell where it is expressed leading to cell death. In this case, an inducible expression system may be useful. When the aim is to investigate SLiM-based interactions this kind of setup allows for several repeats of the peptide sequence (as in **Paper II**), increasing the local concentration of the motif and providing opportunities for rebinding, which may increase the chance of capturing lower affinity interactions [4]. The genetic approaches allow for large flexibility in terms of peptide or protein sequence and tag options. However, the know-how of DNA cloning

and access to a cell culture laboratory is needed. The setup of different constructs and producing cells for harvest can also be time-consuming. An alternative approach can then be to use biotinylated synthetic peptides for pulldowns, as extensively discussed in **Paper I**.

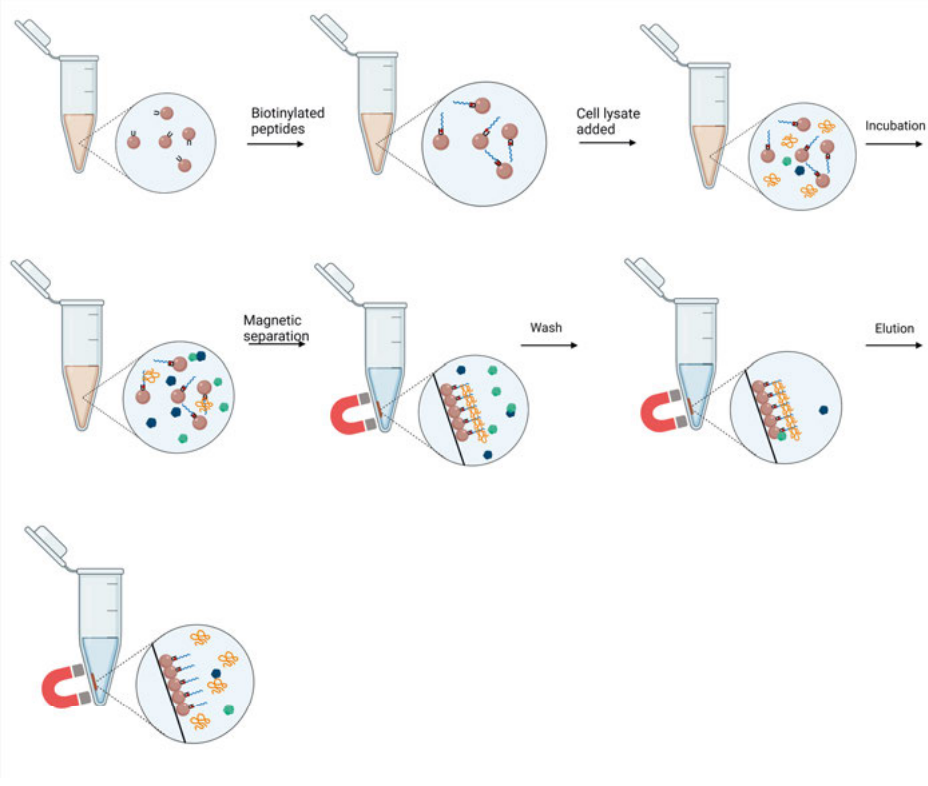


Figure 5: Step-by-step pulldown process using immobilized peptide baits on magnetic beads, as in **Paper I**. Biotinylated peptides are immobilized on streptavidin coated magnetic beads, then they are incubated with cell lysate. The proteins interacting with the peptide engage in interaction, the rest is washed away. Finally, the interacting proteins are eluted. Created with BioRender.com

Protein Interaction Screen on Peptide Matrix (PRISMA)

PRISMA (Protein Interaction Screen on Peptide Matrix) [80] is a peptide array-based approach combined with MS. Peptides are synthesized on a cellulose membrane, which is then incubated with cell lysate. The different peptide spots are then punched out and digested separately. This technique is comparable to the peptide pulldowns as here it is also the low-affinity interaction between the SLiMs and proteins that is to be preserved. However, the throughput is higher and the peptide density is bigger as well. PRISMA is less expensive and requires less hands-on experiment time than the peptide pulldowns.

The method is, however, limited to a maximum 20-25 amino acid residues per peptide [81] which greatly narrows the possibility of having multiple repeats of a peptide or SLiM sequence and can mean the removal of important flanking regions. The experimental steps more easily allow for SILAC (stable isotope labeling by/with amino acids in cell culture) labelling in this case compared to peptide pulldowns, and that is what was used in the first several publications with this method [80][82][83], however, the recently developed universal method employs LFQ (label-free quantification) [84]. PRISMA was used in **Paper I** with both SILAC and LFQ quantification.

Proximity biotin labelling

In the proximity biotin labelling approach the bait protein is expressed as a fusion protein together with a specialized enzyme. Most commonly used are the different versions of the methods called BioID, using the enzyme BirA*(a mutant biotin protein ligase) [85], and APEX, using a mutant ascorbate peroxidase [86]. Both of these approaches have gone through a series of development, reducing their size and decreasing the time needed for labelling. The currently most widely used version of APEX and BioID (APEX2 [87] and miniTurboID [88], respectively) work similarly, with a labelling time of approximately 1 to 10 minutes, labelling the residues tyrosine, tryptophan, histidine, cysteine or lysine.

Crosslinking

Another approach to identifying interacting proteins (or proteins that are in very close proximity to each other) is crosslinking, generating covalent bonds between interacting proteins, which can lead to the preservation of low affinity and transient interactions [89][90].

Crosslinking is usually achieved either by non-specific crosslinking agents, such as formaldehyde [91], or homobifunctional N-hydroxysuccinimide esters which can link lysine residues [92].

After enzymatic digestion this method results in non-customary peptides, therefore special data analysis tools are needed [93].

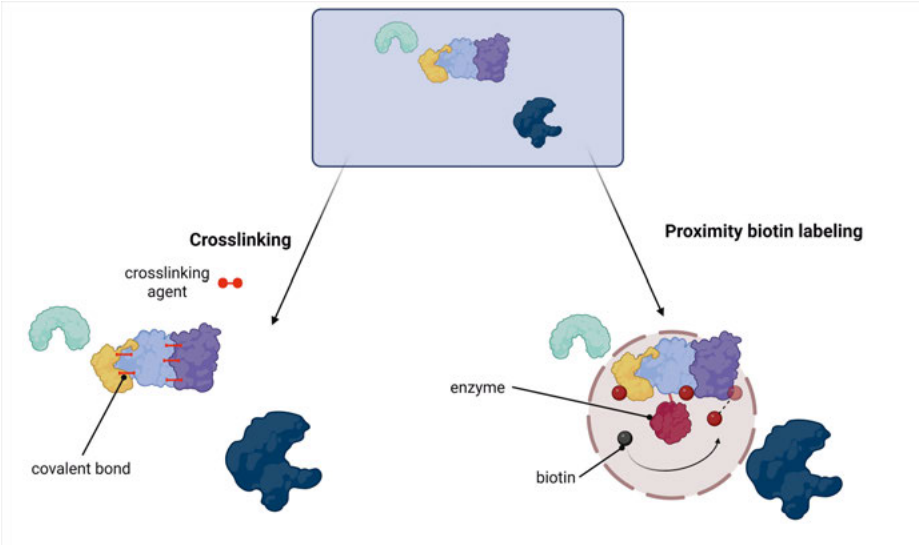


Figure 6: Pulldown approaches employing crosslinking and proximity labelling to ensure capturing possible interactors. In the crosslinking approach, covalent bonds are formed between interacting proteins. In proximity biotin labelling, the proteins in close proximity are labelled with biotin. Created with BioRender.com

Proteomics with mass spectrometry

The development of protein analysis by MS has led to outstanding additions to the PPI investigation toolbox. AP-MS has become a common scalable approach often employed for high-throughput PPI screening [72][73], while newer, related approaches (e.g. proximity biotin labelling, crosslinking) also rely on mass spectrometry detection. MS provides a unique, unbiased approach that is highly valuable for PPI validation and discovery. In this chapter, I present the basis of bottom-up protein analysis as well as the data analysis tools needed to evaluate the data and propose PPIs.

Approaches to protein mass spectrometry

The main approaches to analyse proteins using mass spectrometry are top-down [94] and bottom-up [95]. In the top-down approach, whole proteins are ionized and analysed, while in the bottom-up approach the proteins are enzymatically digested and then the resulting peptides are analysed to provide information about the original proteins. During my work, I focused on the bottom-up analysis, (**Papers I, II, and IV**). The general sample preparation and analysis workflow of bottom-up proteomics is shown in Figure 7 and detailed in the following section.

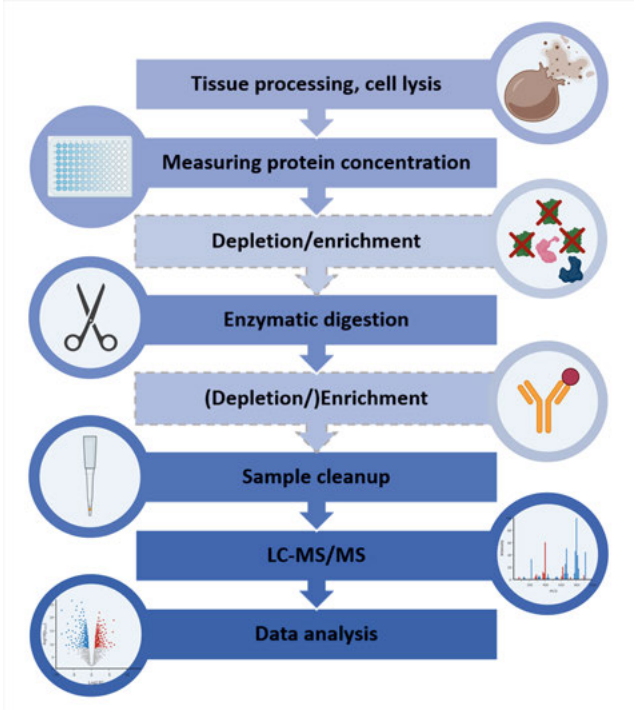


Figure 7: General workflow of the bottom-up proteomic analysis. Created with BioRender.com

Sample preparation

In many cases, the first step is to extract and solubilize the proteins from the original sample (e.g. tissue, body fluid, cells). For cell lysis, several approaches are available, including lysis buffers with detergents or chaotropic agents [96], freeze-thawing cycles, mechanic lysis [97], and a combination of these. Following this, the total protein concentration is determined, which allows to use of appropriate amounts and also to normalize samples. Different colorimetric assays are available for different sample compositions, such as Bradford assay [98] and its variations. In the next, optional, step either specific, often high abundant proteins are removed by depletion (e.g. albumin) [99] or low abundant analytes can be enriched (e.g. posttranslational modifications [100][101]). Depending on the research question, this step can be carried out after digestion at the peptide level as well. In a way, the affinity-based pulldowns can be viewed as enrichment which is carried out at the protein level.

Before the enzymatic digestion, the proteins are treated so that most of the possible cleavage sites would be accessible for the protease in use. For this, often detergents or chaotropic agents are used to disrupt the three-dimensional structure of proteins. In addition, the disulphide bonds are usually reduced using dithiothreitol (DTT) and the generated thiol groups are alkylated with iodoacetamide (IAA) to prevent the re-forming of disulphide bonds.

After reduction and alkylation of the disulphide bonds the proteins are digested into peptides, most often using trypsin, for it cleaves reliably and predictably after lysine and arginine residues [102]. Sometimes multiple-step digestions or protease mixtures including LysC are used to achieve more complete digestion [103]. Subsequently, the sample is cleaned mainly to remove excess salts that would hinder the MS analysis [104]. The commonly used methods are mostly based on solid phase extraction, where the peptides are bound on a stationary phase, salts are removed and the peptides are eluted. Clean-up methods for bottom-up proteomics include commercially available options, such as ZipTips and spin columns, as well as in-house produced StageTips [105]. The final steps (LC-MS/MS and data analysis) are detailed below.

Separation by liquid chromatography

Because of the complexity of most bottom-up proteomics samples, which contain a high number of different peptides, a separation step is needed before the mass spectrometry analysis. Most commonly reversed-phase liquid chromatography is applied. In liquid chromatography, the sample is dissolved in a liquid and carried through a column filled with a solid stationary phase by a pumped liquid mobile phase. The different analyte molecules (here peptides) are distributed between the mobile and the stationary phase and reach the end of the column at different times, based on the analyte's affinity to the stationary phase. In reversed-phase chromatography the mobile phase is polar, usually a mixture of buffered water and an organic solvent. In proteomics, water and acetonitrile with 0.1% formic acid is commonly used. The stationary phase is more hydrophobic than the mobile phase, very often consisting of silica gel particles with surface modification. A commonly used stationary phase category is termed C18 after the octadecyl groups bound to the silica gel surface [106].

In many applications, as in proteomics, gradient elution is often applied. This means that the composition of the mobile phase is changing over time (commonly the ratio of the organic solvent is increased), favouring different molecules to elute. Generally, the more complex the sample, the longer the gradient, 30-150 minute gradients are routinely used for bottom-up proteomics. To protect the analytical column often a smaller protective pre-column is used upfront.

While common HPLC setups work with flow rates in the ml/min range and column diameters ranging up to around a centimeter, nanoLC systems commonly used in proteomics require only 50-500 nL/min flowrates and the analytical column is thin, usually smaller than 100 μm. This system allows for using small sample amounts that would not get too diluted and is capable of producing even smaller droplets in the ion source [107][108]. A delicate system, however, requires constant maintenance.

Mass spectrometry: ion source and detector

Most mass spectrometers consist of three distinct parts: the ion source, the mass analyser, and the detector. There are different solutions suitable for different analytes and different research questions, but here what is described is the instrumental setup used in **Papers I, II, and IV**, which is a Thermo Scientific QExactive Plus Orbitrap instrument.

The analyte molecules are ionized in the nano electrospray ionisation (nanoESI) ion source, namely Nanospray Flex™ and EASY-Spray™ ion sources. As the molecules exit the analytical column through an emitter, small droplets are formed. The droplets are generated because of the electric field applied (1-3 kV), which provides extra charges to the liquid in the emitter, and these charges will be situated on the surface of the fluid. Because of the electric field, a Taylor cone is formed at the tip of the emitter, and small droplets are generated. The droplets are charged and lose solvent as it evaporates while the droplet is nearing the MS inlet. At the end charged gas-phase ions remain and enter the instrument [109][110]. The exact mechanism of this process is still partly unknown and there are three main theories: the ion evaporation model (IEM), charge residue model (CRM), and chain ejection model (CEM) [111]. The ESI is most often operated in the positive ion mode for MS-based bottom-up proteomics because the tryptic peptides most often have basic residues that can accept an extra charge. Ions then enter the inside of the MS, which is under a high vacuum to prevent the ions from colliding with gas molecules. The ion beam is focused and non-charged molecules are removed, then the ions are moved into the C-trap, where they are compressed into a small packet and injected into the mass analyser, the Orbitrap. The Orbitrap was first introduced in 2000 [112] and has since revolutionized proteomics, providing exceptional resolution and mass accuracy. This mass analyser consists of two electrodes, an outer barrel-shaped one and a spindle-shaped electrode inside. The ions are injected and start to spin around the spindle electrode, while also oscillating perpendicularly to the axis of the spindle. The image current of the ions is recorded and converted into a mass spectrum using Fourier transformation [112].

For MS/MS the ions are further fragmented in and the fragment ions are recorded. In **Papers I, II, and IV** the fragmentation happened in a separate cell using higher energy collisional dissociation (HCD) [113]. Other commonly used fragmentation methods include collisionally induced dissociation (CID), electron transfer dissociation (ETD) [114] and electron capture dissociation (ECD) [115].

Quantification

Mass spectrometry is not inherently quantitative – the recorded ion intensities are influenced by various factors e.g. the structure of the molecule and mobile phase composition. For this reason, different approaches have been developed to allow quantification [116][117].

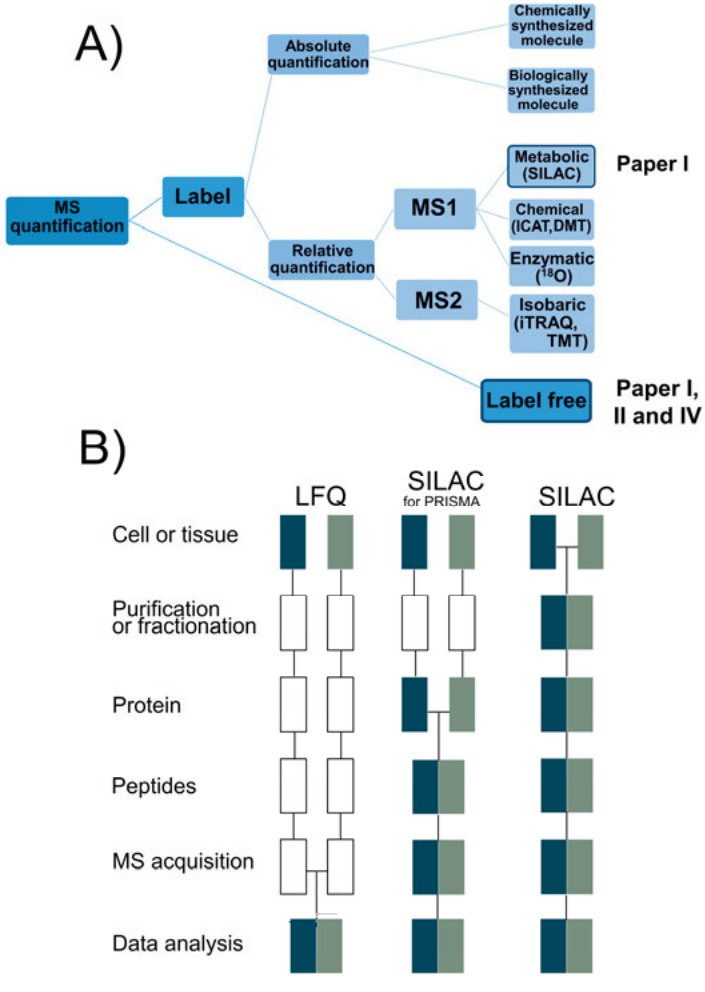


Figure 8: Quantitative proteomics. A) Commonly used quantification approaches in MS-based proteomics. Adapted from Zhang et al.[118] ICAT (isotope-coded affinity tag), DMT (dimethyl labeling), iTRAQ (isobaric tag for relative and absolute quantitation), TMT (tandem mass tag) B) Sample preparation for different techniques, showing when the samples are mixed or processed together. The empty rectangles indicate possible experimental variance at each step.

For absolute quantification, when we can tell e.g. the concentration of a protein or peptide in the sample, specially (chemically or biologically) synthesized standards are needed, which are used as internal standards [119]. While this is a time-consuming and expensive method, absolute quantification is often not needed.

In many cases, the question is more aimed toward how the sensitive balance of protein abundances has changed after for example a treatment. In these cases, relative quantification can be used. This provides fold-change information and can highlight, for example, altered pathways.

There are many different methods for relative quantification, which differ mostly in three parameters: when the labelling is applied, how many different groups can be formed by differential labelling (sample multiplexing), and which step of the MS analysis is the quantification done (MS1 or MS2).

Label-free quantification

Label-free quantification (LFQ) [120] is an easy and cost-effective method, where simply no labels are used. The samples are processed in parallel and analysed separately. Comparative quantification is done during data analysis. For this technique, usually many replicates are used per condition, because the separate sample handling can introduce additional variance at every step. With this technique, a large number of different groups can be compared. There are two basic approaches to label-free quantification, based on spectral counts of ions or peptide-ion intensities, out of which the latter is preferable when the mass resolution and spectrum quality allows for it [121]. LFQ was used in **Papers I, II, and IV**.

Stable isotope labelling by amino acids in cell culture

Stable isotope labelling by amino acids in cell culture (SILAC) [122] is a metabolic labelling technique, where cells are cultured under special conditions. For one group, the cell culture media contains heavy labelled lysine and arginine (or other) amino acids, and after five-six cell doublings, virtually all lysine and arginine residues in the cells are heavy labelled. This allows for early mixing of the samples, and the sample preparation is done on this combined sample, avoiding the introduction of different variance for the different groups. Here usually fewer replicates are used as the results have lower variance. Disadvantages of this technique include the limited multiplexing options (usually 2-3 groups, heavy/light/medium) and that it is not applicable for commonly used samples such as human body fluids. SILAC was used in **Paper I**.

Data analysis

Database search

There are many different tools for the analysis of the great amount of data acquired from a bottom-up proteomics experiment. In **Papers I, II, and IV** the software MaxQuant [123] was used. This tool is freely available and commonly used in the field of quantitative proteomics. Peptides are identified using the high-resolution peptide mass and the fragmentation pattern in MS/MS, and these peptides are then compared to an *in silico* digested protein database consisting of FASTA entries. A false discovery rate is used to filter out false positive hits. It includes a list of common contaminants (for example keratin) and a reverse decoy database, where the sequences of the original database are reversed. Before the search, different parameters, determining the strictness of the search, have to be set, for example, the number of minimum identified peptides, the number of minimum identified unique peptides, peptide length, etc. Also, essential for quantitative investigations, the type of quantification and the used labels (if not label-free quantification) have to be set. Here, consideration has to be made about the nature of the proteins in the database. Viral proteins are often short, therefore, somewhat looser search conditions might be applied, e.g. shorter minimum length of peptides, and fewer minimum identified peptides per protein.

In the case of label-free quantification, MaxQuant uses the intensity based on the peptide ion and MS/MS fragment intensity with MaxLFQ [121].

Another proteomics data analysis tool is Proteome Discoverer (Thermo Fisher Scientific), which was just recently updated to include quantification. Version 1.4.0.288 was used in **Paper IV** to better visualize identified peptides and their fragment ions.

Statistical analysis

After acquiring the quantitative intensity values from MaxQuant, the Perseus software provides a versatile surface for data analysis and statistics [124], amongst others. After data filtering, most commonly t-tests or an ANOVA analysis is carried out to see which proteins are differently expressed in the different groups. Volcano plots provide an intuitive visualization of the results. Perseus was used in **Papers I, II, and IV.**

Other software tools or coding platforms can be also used, such as Statistica MSQuant [125], R and Matlab.

Missing values and imputation

There are different approaches to handling missing values for statistical tests. In many experimental setups, a missing value equals an experimental mistake, for example, a dropped Eppendorf tube. However, in the case of affinity enrichment, missing values are natural and can be desirable: when a protein can

be found in all of the samples, but none of the controls (giving a 0 intensity, which is later transformed into NaN (not a number)– a missing value), it means that the protein is greatly enriched in the sample compared to the control. To be able to include these values in the statistical analysis, the missing values have to be replaced with a numerical value in the process of imputation [126].

Scoring and evaluating results

Because of the nature of pulldown experiments, there is a special focus on evaluating the results to remove false positive hits. One strategy aims at discarding protein hits commonly identified in pulldown experiments, that might be specific binders of e.g. the affinity matrix, but not the bait protein. The contaminant repository for affinity purification (CRAPome) database consists of protein hits complied from 360 experiments, collecting data also about the protocol, cell line and affinity matrix used, and so on [127]. By selecting the actual experimental conditions, a list of "frequent flyers" is easily attainable.

Also available are tools for scoring the affinity purification hits, with a higher score meaning that the interaction is more likely to be biologically relevant. The most known scoring algorithms are SAINT (significance analysis of interactome) [128][129], CompPASS (Comparative Proteomic Analysis Software Suite) [130], and MiST (Mass spectrometry interaction Statistics) [131]. In **Paper I**, SAINTexpress, a further improved version of SAINT was used. This software provides confidence scores using label-free quantification intensity input from samples and negative controls.

Network analysis

After the enriched proteins are identified, different tools are at hand to provide a deeper understanding of interactions. BioGrid [132] and IntAct [133] are both curated databases including around 866,000 and 750,000 human interactions, respectively. The HIPPIE (Human Integrated Protein-Protein Interaction rEference) database [134] contains only experimentally validated and confidence-scored interactions, while STRING [135][136] consists of both known and predicted interactions, providing easy network visualization and enrichment analysis. DAVID [137] allows for a wide range of functional annotations. Interactors listed in BioGrid were used for comparison with data generated in **Paper I**.

Present investigations

As methods and instruments improve, we gain a more detailed insight into the complex PPI networks and understand the role specific interactions play in cellular processes. Motif-based interactions, however, require special tools for discovery and validation, as these interactions tend to be transient and have a low-to-mid micromolar affinity [4][53]. Often neglected, only a small portion of this kind of interaction is known today, despite their role in normal cell function (e.g. signalling [138]) and disease (viral host motif mimicking) [27][6].

It is clear, therefore, that tailored techniques are needed. One of the gold standard methods for PPI discovery is AP-MS, in which a tagged protein is used to enrich its interaction partners from a protein mixture as described in the previous section. This, however, can be challenging in the case of low affinity and transient interactions, as they are easily removed during washing steps, which are normally used to eliminate non-binding proteins.

During my PhD research, I focused on developing and evaluating techniques for SLiM-based interaction capture for MS analysis (**Paper I**), investigating SLiM-based interactions between viral and human proteins (**Papers II and III**), and elucidating the change of interactome upon viral infection (**Paper IV**)

I. Synthetic peptide-based methods for SLiM-mediated interactions

In **Paper I** I used immobilized peptides for pulling down interactors from a cell lysate. One or two repeats of a motif-containing peptide were used, either biotinylated and immobilized on streptavidin-coated magnetic beads, or on cellulose membranes. The enriched proteins were determined with MS using label-free quantification. The goal of these experiments was to determine which capture method is preferable and also to what extent the methods are suitable for SLiM-based pulldowns and if there is a generic protocol applicable.

II. SLiM-based interactions between viral and human proteins

To explore how RNA viruses hijack human cells using SLiM-based interactions, large-scale screening of RNA-virus peptides against human domain baits was carried out using ProP-PD in **Paper II**. Peptides from different viruses were found to interact with the human protein PABP (poly(A)-binding protein), which plays a crucial role in translation initiation. To validate this interaction and discover the wider interactome of the peptides, GFP-tagged viral peptide repeats were used for pulldowns from uninfected or virus-infected cell lysates and analysed by AP-MS.
In **Paper III**, a reversed approach was used. To determine how SARS-CoV-2 protein might interact with human SLiMs, a human disorderome phage library [53] was screened against SARS-CoV-2 protein domains. As only a small number of hits were identified, the challenge was to establish if the peptides contained defined binding motifs. Here I used detailed SPOT peptide arrays with alanine scanning to determine the key residues in each peptide.

III. Changes in the interactome upon viral infection

Upon TBEV and LGTV infections, the cellular localization of the nuclear pore complex protein NUP153 (nucleoporin 153) changes and possibly plays a role in viral replication. To understand how the interactions change in the case of infection, in **Paper IV** EGFP (enhanced GFP)-tagged full-length NUP153 was used for pulldowns from uninfected, TBEV- or LGTV-infected cell lysates, and the enriched proteins were analysed by MS.

I. Synthetic peptide pulldowns (Paper I)

For a large part of my PhD research, I worked with synthetic SLiM-containing peptides to pull down protein interactors. The goal was to develop a pulldown method more tailored toward SLiM-based interactions. A relatively quick and simple method for SLiM-based interaction validation is needed, as more and more information is available through high-throughput methods, such as ProP-PD. The use of synthetic peptides provides a good opportunity to prove that the interaction is based only on the given motif, and not other parts of the protein. If the interaction motif is known, the key residues can be mutated into alanine (or glycine) to disrupt the interaction.

To explore if it was possible to robustly pull down SLiM-based interactors with varying affinity I used biotinylated peptide pulldown. I further compared the PRISMA assay with the peptide pulldown using the same interaction motifs, to explore if this higher throughput method could be used to reliably pull down SLiM-binding proteins. The PRISMA assay was introduced in 2017

[80] and presented a simple, cellulose membrane-based approach that could be used for the goals presented above.

In both cases, the identification and relative quantification of enriched (and possibly interacting) proteins was done by MS, to explore if the expected binders were pulldown, and to explore if also other proteins could bind to the peptides used. This way additional interactions could be elucidated.

From the beginning, it was clear that the low-affinity interactions require special handling. Below I describe how I attempted to accommodate this, what challenges I met during the process, and my conclusion from the analysis.

Parameters to consider and method development

Several parameters affect PPIs, such as local protein (and motif) concentration and the chemical environment (detergents, salts, pH, etc.). In the case of SLiM-based interactions, the flanking regions and disorder propensity may also affect the affinity of the interactions [21]. There are several steps in the pulldown workflow, therefore, that need to be optimized for higher success rates. Ideally, these parameters would be altered for each specific interaction. However, in my project, the goal was to find mostly universally applicable parameters. It should also be added, that these parameters are rarely discussed in detail in publications, even though they can critically affect the experimental outcome.

The first thing to consider is the used peptide sequence itself. Even though SLiMs usually contain only a few key residues, flanking regions can be also of importance. This should be considered when selecting the exact peptide sequence. In my case the source of most of the peptide sequences was phage display, where phage libraries displaying 16 AA long peptides were used [53], therefore that became my choice of length too. It is in line with the length used in similar approaches [139][140][141] (however, the double repeat sequence is longer than usually).

Another crucial component is the cell type chosen and the lysis method applied to obtain proteins. Choosing the cell type can be made based on the targeted interactions in question – maybe they are more relevant in a certain tissue. It is also important that the prey protein in question is expressed in the cell line used. However, it is hard to gain reliable, protein-level expression information for many cell lines and many proteins. RNA expression level available from the Human Protein Atlas (proteinatlas.org) [142] can give a clue but is not clearly correlating with protein expression. For this project I first started working with HEK293 cells, later switching to and carrying out all experiments with HaCat cells. HaCats were slightly easier to handle and grow in cell culture and also produced better results with the first tested candidate interaction (involving the PTAP motif) (**Paper I).**

During the lysis, it is also important to choose a method, which does not denature or otherwise degrade the proteins. Otherwise, the proteins might lose their ability to engage in their natural interactions. Here I employed a simple, minimally denaturing lysis buffer consisting of PBS, 1% NP-40 substitute (non-ionic detergent), and protease inhibitors, and omitted harsh physical lysis methods such as sonication.

Choosing the concentration level of proteins is also important, as with very high concentrations false positive hits are more likely to be generated, and it also requires a lot of resources. Here I aimed for concentrations similar to those reported in the literature: 0.8 mg/ml total protein in peptide-pulldowns [143][144][145][146], and 4 mg/ml in the case of PRISMA [84]. As for the concentration of the peptide baits, on the PRISMA membrane, it was around 7,000 times higher than on the magnetic beads (according to manufacturer information) for single repeat peptides. To increase the local concentration of the SLiMs, double repeats of the peptide sequences were also tested.

It is also important to consider the amount of detergent used during the pulldown and subsequent washing steps. Higher levels of detergent in the solvent can negatively affect interactions, so in the case of low-affinity interactions, it should be kept at a minimum. For this reason, the original lysate was diluted to the preferred concentration with lysis buffer without detergent.

One of the most crucial steps is the removal of unbound proteins after the pulldown. Insufficient washes result in a lot of background, non-interacting proteins. To be careful, these steps were carried out as fast as possible, with no detergent and on ice, to minimize complex dissociation.

Furthermore, there are different approaches to elute the interacting proteins as well. In the beginning, I used desthiobiotinylated peptides that could be eluted from the streptavidin beads using biotin [147][148]. However, this method resulted in a lot of biotin-related contaminants in the mass spectra. Later, using biotinylated peptides, glycine-, urea- and SDS-based elution buffers were also tested to elute proteins from the peptides, as the biotin-streptavidin interaction is very hard to disrupt. The glycine-based elution resulted in the highest amount of eluted proteins. This method is more gentle to the proteins, however, a very acidic condition (pH 2.5) is required, so the eluate should be immediately neutralized with a suitable, high pH buffer. It is also possible to omit the elution step and move on with “on-bead” digestion. However, I decided to implement an elution step to avoid a high number of streptavidin-related peptides after digestion.

In the case of the PRISMA, there is no need for elution, as the peptide spots with the proteins interacting are punched out, and the digestion is carried out directly “from” the cellulose membrane. In this case, I followed the recommendations of the developers of the method [84] and used LysC in combination with trypsin, to aid the digestion.

Usually, the amount of enzyme used for protein digestion is determined based on the total protein amount in the sample. However, here there was no

straightforward way to determine the concentration and it was estimated to be different in different samples by any means. Therefore, a small amount of enzyme was used, expecting 10 µg or less total protein in the samples.

Limiting factors

When working with synthetic peptides the first issue might arise right at the synthesis. Certain sequences are just not possible to synthesize, or cannot be purified to the extent needed. This we also experienced, most often with control peptide sequences that contain more alanine residues. The length of the peptide is also a key element, as longer sequences are harder to synthesize, with common peptide synthesis methods being limited to ca. 50 amino acid residues, although many shorter sequences may be difficult to synthesize [149]. This in turn affects the number of peptide repeats that can be used in tandem. In the case of PRISMA, the length is limited to 20-25 amino acids, so this is even more pronounced, and can make truncating the flanking regions necessary. For this approach, the peptides are synthesized directly on the membrane and no purification of the peptides is possible, which can result in noisier data.

The peptide spots on the PRISMA membrane are quite uneven (see Figure 9), which can influence the quantitative comparison. In comparison, even though some magnetic beads were lost in the sample preparation steps in the case of the biotinylated peptide pulldown, that error affected all samples similarly.

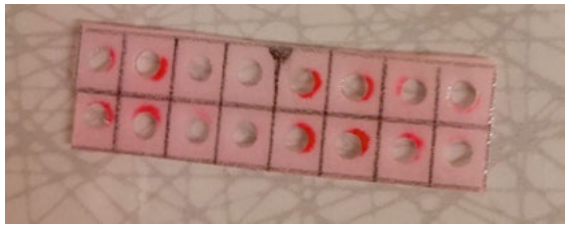


Figure 9: PRISMA membrane after punching out each spot, Ponceau S stained.

It should also be mentioned, that buying purified synthetic peptides is much more costly than the PRISMA membranes, but it does provide enough material for several experiments.

In comparison to PRISMA, biotin-peptide pulldowns need a lot more hands-on time as well and include more steps. For the latter reason, this setup is not suitable for early labelling quantification methods, such as SILAC, as the samples could only be mixed after the elution step.

Another thing to consider is the expression level of the individual proteins. This can depend on the cell type, cell cycle stage, or external effects that could put the cells under stress. For this, a great example is the peptide pulldown done with a peptide from human TP53 that is known to bind with MDM2 SWIB domain, which is explained below.

Results and comparison

The biotinylated peptide pulldown was carried out using 8 different peptides (both as single and double repeats), which were known to interact with 3 different human peptide binding domains with a wide range of affinities (K_D between 15.3 μM and 0.0014 μM). Enriched proteins were detected and quantified using label-free MS, then the results from peptide and control were compared using t-test and SAINT scoring. For one interaction, that consistently produced good results in the peptide-pulldowns, the same experiments were repeated using PRISMA membranes and SILAC as well as label-free quantification.

Evaluating the results in groups based on the target proteins used revealed that the biotinylated peptide-pulldown is an applicable tool to validate some SLiM-based interactions, but also highlighted some of the parameters that can affect this kind of experiment, and that the approach is not applicable to all interactions.

Peptide pulldown of interactions with various affinities

The TGE motif (more completely [DNS]x[DES][TNS]GE) can be found in several human proteins and binds to the Kelch domain of KEAP1 [150][151][152][153]. The motif is classified as a degron, for it is important in ubiquitin-dependent protein degradation. When using TGE-motif-containing peptides for pulldowns, they effectively pulled the known interactor KEAP1 and several known interactors of the bait protein or KEAP1 (Figure 10). Importantly, the interaction was preserved during the pulldown over a wide range of binding affinities (K_D between 5.7 and 0.0014 μM). The results showed a lower number of enriched proteins using the single repeat peptide versus the double repeat, which indicates that while single repeat peptides might give cleaner results of binary interactions, double repeats can provide additional information on potential interaction networks.

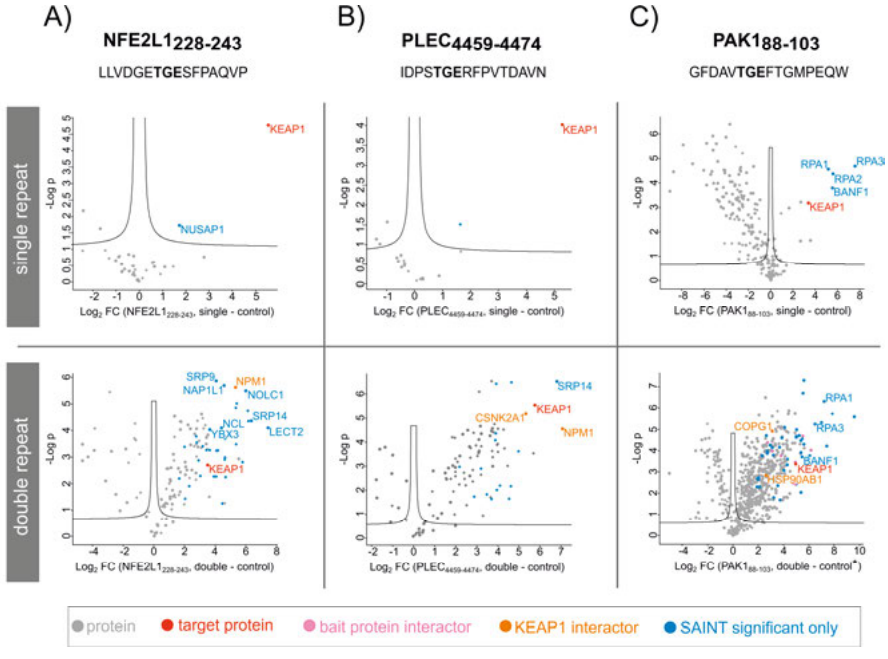


Figure 10: Volcano plots showing enriched proteins after pulldown with TGE motif-containing peptides. A) NFE2L1$_{228-243}$ peptide, single (top panel) and double (bottom panel) repeats B) PLEC$_{4459-4474}$ peptide, single (top panel) and double (bottom panel) repeats C) PAK1$_{88-103}$ peptide, single (top panel) and double (bottom panel) repeats. Significance determined by permutation-based FDR<0.05 (250 permutations), S_0:0.1, SAINT significant: SAINT score >0.85. From Paper I.

The use of single or double peptide repeats matter

Employing double repeat peptides proved to be essential in the case of the key interaction between TSG101 UEV domain and the PTAP motif-containing peptide from the GAG polyprotein of HIV-1 and HIV-2. In this case, interestingly, the known interactor TSG101 was only pulled down with the double repeat peptide from HIV-1 GAG (Figure 11). This highlights the possible beneficial effect of displaying multiple motif repeats. On the other hand, the results underline that in a pulldown, there are parameters affecting the results other than affinity, as seen in the case of HIV-2 GAG peptide, where despite higher affinity TSG101 was not pulled down. The reason for this is not evident, the interaction can be affected by the flanking regions, competing interactions, or there might be reasons related to peptide synthesis or degradation.

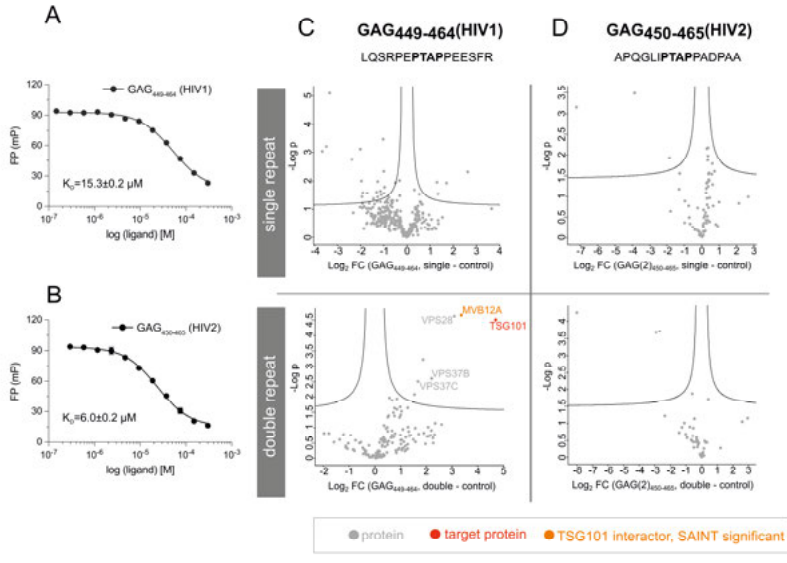


Figure 11: A-B) Affinity measurements using fluorescence polarization displacement against TSG101 UEV domain (error bars indicating one standard deviation). K_D variance is shown with one SEM (standard error of the mean) A) $GAG_{449-464}$ peptide (HIV-1) B) $GAG2_{450-465}$ peptide (HIV-2) C-D) Volcano plots showing enriched proteins after pulldown with TGE motif-containing peptides. C) $GAG_{449-464}$ peptide, single (top panel) and double (bottom panel) repeats D) $GAG2_{450-465}$ peptide, single (top panel) and double (bottom panel) repeats. Significance determined by permutation-based FDR<0.05 (250 permutations), S_0:0.1, SAINT significant: SAINT score >0.85. From Paper I.

The effect of protein abundance on pulldowns

Three peptides were tested to pull down MDM2, originating from the human proteins TP53, RNF115, and NUMB and known to interact with the MDM2 SWIB domain [154][155][53][156] . All three peptides (and their double repeat versions) failed to enrich MDM2, despite for example the TP53-MDM2 interaction being very well-established. As MDM2 is low abundant in cells that are not under stress [157], the low local concentration of this protein was likely the challenging factor. To this end, MDM2 SWIB domain-spiked lysates were also tested, where MDM2 was identified and enriched using the TP53 peptide. Thus, false negative results may be obtained from peptide-pulldowns when aiming to validate interactions with low-abundant, or temporarily expressed proteins.

Comparison of PRISMA-pulldowns and biotinylated peptide-pulldowns

The PRISMA method offers an economical way of testing peptide-based interactions with much less hands-on time. To test how it measures up against

the biotinylated peptide-pulldown in the case of the SLiM-based interactions, partly for what it developed, the same peptide sequences were used (double repeats had to be truncated because the maximum peptide length was 20AA). For benchmarking we used the KEAP1 Kelch binding $PLEC_{4459-4474}$ peptide, which performed excellently in peptide pulldowns. First, we used SILAC-labeled cell lysates, as in nearly all early PRISMA publications [80][82][83], but failed to identify the KEAP1 in the pulldown. Then the experiments were repeated with an LFQ quantification approach, which has a higher dynamic range [158]. KEAP1 was identified only with the double repeat peptide and with only a few peptides showing up in the mass spectra. We also included $GLUT1_{478-491}$ and $SOS1_{1146-1159}$ positive control sequences in our testing, which resulted in enriched proteins considerably overlapping previous results [84], demonstrating that the method was applied correctly.

Noting that the motif in the $PLEC_{4459-4474}$ peptide is compact (TGE), using the double repeat approach might be a valid approach using PRISMA for this particular motif, but likely not for longer motifs. These points highlight the moderate power of this approach compared to the peptide pulldown.

Conclusion and outlook

As the number of identified and predicted SLiMs grow, new and improved methods are needed to validate their interactions. The methods presented here (biotinylated peptide pulldown and PRISMA) each have their limitations, but they are valid approaches for at least some motifs. Further studies using the PRISMA approach would be of interest, in terms of what types of SLiM-based interactions it can capture and which interactions it fails to identify. When compared, PRISMA seems to be inferior to peptide pulldowns, as maybe can be expected for a higher throughput approach. A disadvantage of the peptide pulldowns is the tedious sample preparation, but this can be overcome by readily available automation options using the same magnetic beads as employed in our study.

II. SLiM-based interactions between viral and human proteins (Papers II and III)

While there are many crucial SLiM-based interactions in healthy human cells, viruses also often employ SLiMs to hijack the host cell machinery. For this reason, we turned our attention to host-virus PPIs. We explored the topic from a broader perspective (**Paper II**) and with a SARS-CoV-2-focused approach (**Paper III**).

RNA-virus motifs interacting with human proteins – large-scale screening

To investigate how viruses employ SLiMs across viral species, a peptide phage library tiling the intrinsically disordered regions of 229 RNA viruses was generated and screened against 139 human protein domains [23]. In total, 1,712 possible interactions were discovered involving 1,285 medium/high confidence peptides. It was apparent, that most types of viruses use the SLiM-mimicking strategy and the interactions are often with human proteins from specific pathways. Compelling examples were selected for further validation.

Validations

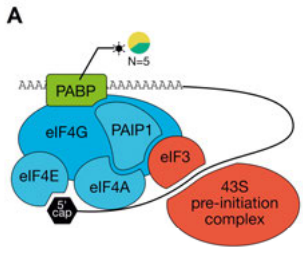
The ESCRT complexes are commonly targeted by viruses for essential functions such as viral budding [159]. We found 64 different peptides interacting with four different ESCRT proteins (TSG101, ALIX, NEDD4, and CEP55). Affinity measurements carried out in this study revealed that the affinities between viral SLiMs and human proteins are often not higher than between the original human SLiM-human protein interactions, contrary to the previous notion [6].

Secondly, the clathrin-mediated endocytosis machinery also appeared as a common target of viral interference through SLiM –based interactions, with 172 viral peptides targeting clathrin or several of the adaptor proteins. Here, several interactions were confirmed via affinity measurements, pulldowns with Western blots, PLA, and the clathrin N-terminal domain (CLTC-NTD) was co-crystallized with two different viral peptides.

Viruses targeting translation

Viruses need to get access to the translational machinery to get their own proteins expressed. In our screen, five different viral peptides were found to bind to PABPC1 PABC (Polyadenylate-binding protein 1 C-terminal domain). PABPC1 plays a crucial role in translation initiation, where its PABC domain interacts with PAIP1 or with the repressor PAIP2 while the RNA-binding domains (RRMs) interact with the poly(A) tail of mRNA (messenger RNA)[160][161].

The identified viral peptides contained similar recognition motifs to the endogenous proteins and bound with similar or lower affinities as those (Figure 12B).



B

Peptide sequence	Protein	K_D (μM)
108-KSN**LNP**N**A**KE**F**V**P**GVK-123	PAIP2 (Human)	0.16
125-MSK**LSV**N**A**PE**F**Y**P**SGY-140	PAIP1 (Human)	1.6
183-SAG**LNP**A**A**VP**F**V**P**KNQ-198	P/V/C (HeV)	2.0
183-TTG**LNP**T**A**VP**F**T**L**RNL-198	P/V/C (NiV)	15
351-HPL**LNP**S**A**LE**F**N**P**SQT-366	N (HCoV 229E)	23
2-NSM**LNP**N**A**VP**F**Q**P**SPQ-17	N (BeV)	2.7

Figure 12: A) "Closed loop" structure including PABP, which is needed for translation initiation. N=number of different interactions. The pie chart indicates the share of different viral types, yellow: (-)ssRNA (single stranded RNA) green: (+)ssRNA enveloped. B) Human and viral peptides interacting with PABPC1 PABC, with the key residues shown in bold. Affinities are shown as determined by competitive FP. Adapted from Paper II.

The interactions were further confirmed with pulldown experiments with Western blot detection, and the HCoV229E $N_{351\text{-}166}$ peptide was successfully co-crystallized with the PABC domain (PDB id: 7BN3). The crystal structure reinforced what was indicated before, the viral peptide occupies the same binding pocket as the endogenous proteins, with conserved amino acid residues in positions 1, 2, 3, 5, 8, and 10 (L[NS][PV]xAxxFx[PL]).

Because of their interference with a key process such as translation and their similar binding as the translation repressor PAIP2, the viral peptides were investigated as possible inhibitors of viral replication. To this end, lentiviral constructs expressing EGFP-4x($P(HeV)_{183\text{-}198}$), also referred to as EGFP-PABPi, and a similar control with mutated motifs was generated, which proved to inhibit infection in the case of almost all RNA viruses tested (Figure 13).

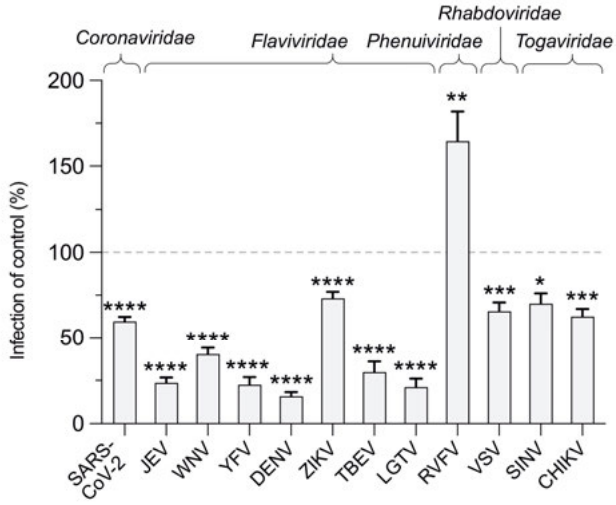


Figure 13: Antiviral effect measured in case of different viruses. The cells were first transduced with lentiviral constructs of EGFP-PABPi or EGFP-PABPi mut, then infected with the specific virus. The percentage of infected cells was compared. Adapted from Paper II.

Affinity purification-mass spectrometry

For this study, I had the opportunity to do a more classic AP-MS experiment using the above-mentioned lentiviral EGFP-PABPi construct and its respective control. These constructs were expressed in and used for pulldown from uninfected HEK293 (transiently transfected), TBEV-infected VeroB4, and SARS-CoV-2-infected VeroE6 cells (stably expressing).

With this method, additional parameters have to be considered. The use of a larger tag, such as EGFP (26.9 kDa) [162] might disturb the function of the fused peptides. However, previously, using YFP, it was shown that this approach was viable [163] and it also worked well in our hands. A great advantage here is that there is no strict limit on the length of the fusion product and therefore no limit on the number of peptide repeats used. Here we used 4 peptide repeats, as it has been successfully used in other studies.

For these pulldowns, I employed magnetic beads with immobilized anti-GFP nanobodies (Chromotek) with glycine-HCl, pH 2.5 for elution, and then similar sample processing as in the biotinylated peptide pulldowns.

In the case of the uninfected HEK293 lysate, the pulldown resulted in high enrichment of PABPC1 and its homolog, PABPC4, as expected. Almost all of the other enriched proteins were ribosomal or RNA-related proteins (Figure 14A).

When testing Vero cells one of the challenges arose from the fact that it is not a human cell line, but originating from African green monkey (*Chlorocebus*), therefore the proteome is not complete in the database, which can lead to losing some valuable hits.

Carrying out the pulldown from SARS-CoV-2 infected cell lysate resulted in a similar profile (Figure 14B). The only identified viral protein was N (Nucleocapsid), and its previously [164] identified human interactor, G3BP1 was enriched as well. These proteins may have been pulled down indirectly through RNA-mediated interactions.

The TBEV-infected lysate resulted in fewer significant hits, but the trends were similar to those above (Figure 14C).

The results thus confirmed that EGFP-PABPi interacts with PABPC1, and revealed that it also binds to the homologous protein PABPC4.

The genetically fused EGFP-peptide approach thus appears to robustly pull-down peptide binders.

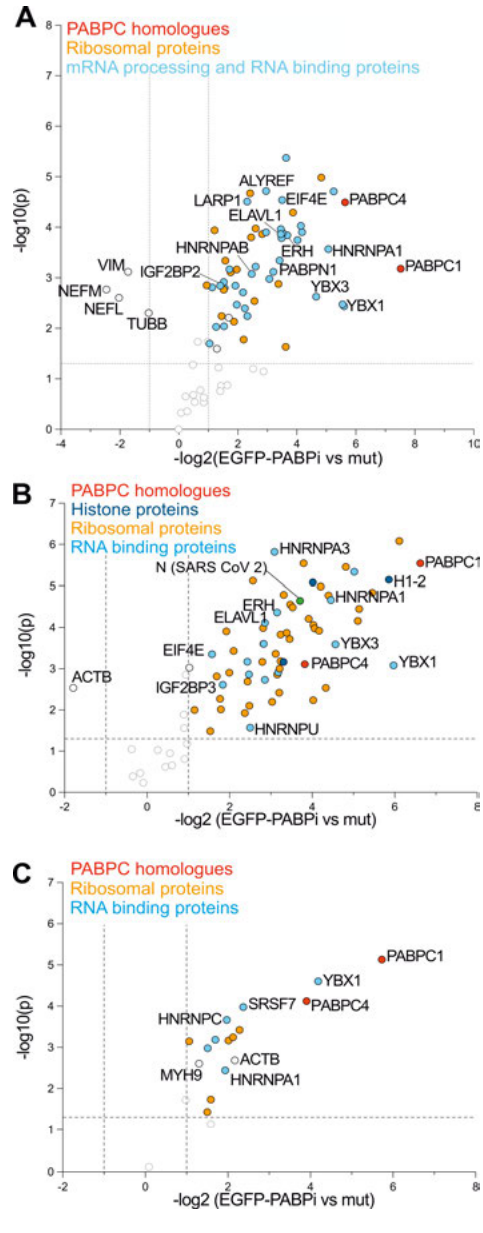


Figure 14: Mass spectrometry analysis of differential expression in lentivirus transduced cells expressing EGFP-PABPi or EGFP-PABPi mut . A) non-infected HEK293 cells B) SARS-CoV-2 infected VeroE6 cells C) TBEV infected VeroB4 cells. Adapted from Paper II.

One drawback of the approach is that the peptide sequences may be toxic to the cell, which may limit cell viability and growth. Furthermore, the expressed peptides might degrade in the cellular environment, which is difficult to control for.

Human SLiMs that interact with SARS-COV-2 viral proteins

Viruses usually have small genomes that code only for a dozen or two proteins. In SARS-CoV-2, for example, 14 open reading frames (ORFs) code for 29 proteins in total [37]. While we usually focus on SLiM-discovery in viral proteins [164] here an opposite approach was used, where 25 SARS-CoV-2 proteins or protein domains were screened using a human disorderome peptide phage library including around 1 million peptides [53]. 281 medium/high confidence human peptide interactions were found involved with nine different viral proteins (Figure 15A).

Some viral proteins were involved in a high number of interactions with human peptides, such as Nsp9 (118 interactions), Nsp1 (47 interactions), and Nsp5 Mpro (inactivated) (32 interactions). Some of these binding domains might be exploited as sites for inhibiting viral infection with peptide-like molecules

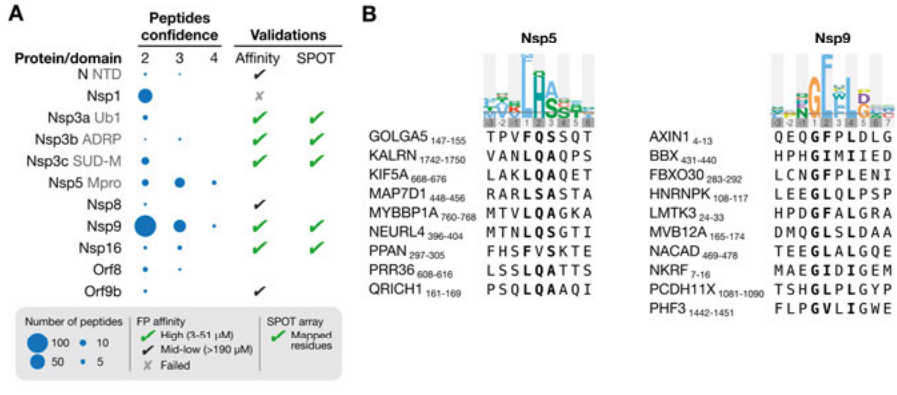


Figure 15: A) Interacting peptides found for different SARS-CoV-2 proteins/domains, and whether affinity measurements and SPOT-array validations were carried out. B) Selected peptide hits found to interact with Nsp9, and the enriched interaction motif, key residues appear larger. Adapted from Paper III.

Motif discovery and methods

The more peptide hits that are identified as interactors of a given protein domain, the higher the chance to identify a consensus interaction motif. For this, peptides are aligned to reveal conserved residues. In the case of Nsp5 and

Nsp9 it was possible to define the consensus motifs: [FLM][HQ][AS] and G[FL]xL[GDP] (Figure 15B), respectively.

When only a few peptides are found, determining the consensus motif might be more challenging.

Here, I employed alanine scanning using SPOT arrays in order to identify and confirm key binding residues. This method is more straightforward and cost-effective than other methods used by our consortium, such as FP (high number of peptides to order), or AP-MS (many constructs/synthetic peptides would be needed and extensive instrument time).

To highlight two different ways the SPOT arrays were useful in this study, I will present the cases of Nsp9 and Nsp3 ADRP.

For Nsp9 the established motif based on ProP-PD was G[FL]xL[GDP]. By subjecting two peptides ($LMTK3_{22-36}$ and $NKRF_{8-23}$) to alanine-scanning, the motif could be further refined.

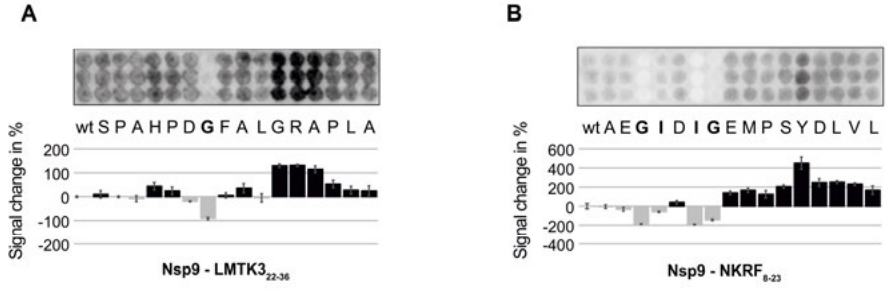


Figure 16: SPOT array results after incubation with purified GST-6xHis-Nsp9. A) $LMTK3_{22-36}$ peptide alanine scanning, triplicates. B) $NKRF_{8-23}$ peptide alanine scanning, triplicates. Adapted from Paper III.

While the analysis of the $LMTK3_{22-36}$ peptide confirmed the importance of the first G and the following 4 residues also conformed with the identified motif (Figure 15), mutating the second G to A resulted in a higher signal than wt. In the case of $NKRF_{8-23}$, the two consensus Gs were found to be crucial for the interaction, as expected from the consensus motif (Figure 17). The importance of the first I (isoleucine) also agreed with the determined motif. The second I was not included in the original motif, therefore the motif was adjusted to GΦxΦ[GDP], where Φ is a hydrophobic residue. Interestingly, this motif can be found in the C-terminal of Nsp9 and is known to promote dimer forming [165]. Via HSQC NMR experiments it was determined that the Nsp9-binding peptides interfere with the dimer formation of Nsp9.

Contrary to Nsp9, there were only four peptides found to interact with Nsp3 ADRP domain (Table 1), two of which were overlapping (MBOAT1). Thus, it was not possible to directly suggest a consensus binding motif.

Table 1:Peptides identified in ProP-PD to interact with Nsp3 ADRP domain

Protein	Sequence	K_D
MBOAT1	16-GSTYLHPLSELLGIPL-31	100 μM
MBOAT1	18-TYLHPLSELLGIPLDQ-33	-
AZIN2	1-MAGYLSESDFVMVEEG-16	-
PARP10	700-DGGTDGKAQLVVHSAF-715	310 μM

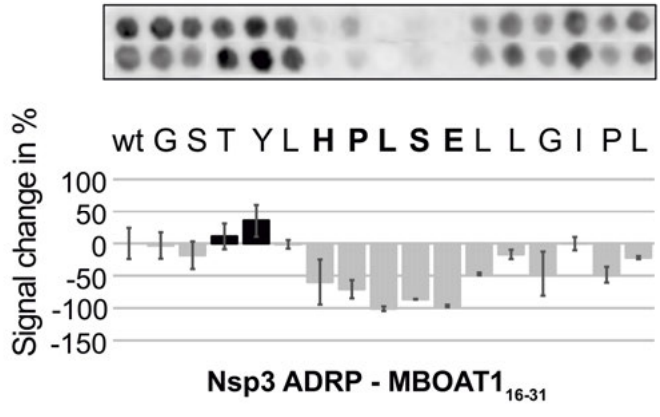


Figure 17: SPOT array results after incubation with purified GST-6xHis-Nsp3 ADRP using $MBOAT1_{16-31}$ peptide alanine scanning, duplicates. Adapted from Paper III.

The SPOT array analysis of $MBOAT1_{16-31}$ revealed that a four amino acid stretch (HPLSE) is critical for binding to this peptide (Figure 17). Notably, the AZIN2 peptide contains a similar YLSE segment, suggesting that the motif may be defined as ΦLSE. The $PARP10_{700-715}$ peptide was also tested, however, possibly due to its lower affinity, no interaction was observed on the membrane. The results demonstrate the intuitive and informative way SPOT arrays with alanine scanning can be used in situations similar to this.

SARS-CoV-2 domain binding peptides as antivirals

We further showed that several of the identified peptides have an inhibitory effect on SARS-CoV-2 infection. Using EGFP-4xpeptide lentiviral constructs proved to decrease the infection rate in VeroE6 cells (Figure 18A). This effect was also confirmed in the case of $NOTCH4_{1605-1620}$ and $DYRK1B_{394-409}$ fused to the HIV Tat-derived sequence (YGRKKRRQRRRGSG) and used as cell-

penetrating peptides (Figure 18B), showing that the peptides have potential as starting molecules for design for SARS-CoV-2 inhibitors.

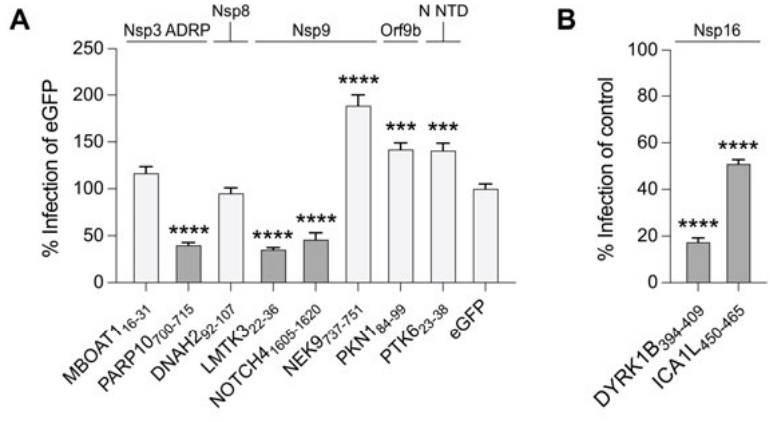


Figure 18:Antiviral effect of lentiviral constructs including different SARS-CoV-2 protein-targeting peptides in SARS-CoV-2 infection, compared to A) EGFP or B) constructs with mutated peptide sequence.
*p<0.05,**p<0.01,***p<0.001,and****p<0.0001 showing significance determined by unpaired t-test. . Adapted from Paper III.

Conclusion and outlook

The two studies outlined above emphasize the value of large-scale ProP-PD screening for understanding host-viral interactions and for potential inhibitor discovery. Our work also shows how valuable information might be omitted when using more traditional interaction discovery approaches, such as AP-MS [37][166][167][168], and the need for specific methods for finding and validating SLiM-based interactions. An AP-MS approach tailored to SLiMs (using peptide repeats and suitable controls) can provide an additional layer of information. The simple method of SPOT array was helpful to identify and refine the interaction motifs.

In both **Paper II** and **Paper III** we have identified candidate peptides that interfere with SARS-CoV-2 infection and can be employed as the starting point for antiviral development.

III. Changes in the interactome upon viral infection (Paper IV)

During viral infection, the viruses rely on hijacking interactions of the host cell at nearly every step. This can result in immense changes in the interactomes, localization as well as expression, and function of host proteins.

One such example is NUP153, one of the members of the nuclear pore complex (NPC). The NPC serves as a gate between the nucleus and the cytoplasm, allowing for the transfer of proteins and RNA [169]. Many members of the NPC, such as NUP153, are highly disordered proteins, serving as "molecular filters". Viruses rely on interactions with NPC proteins as a way of entering the nucleus [170]. It was shown that multiple viruses target NUP153 and engage its disordered FG-repeats for interaction with viral proteins [171][172]. A great example of this is the HIV-1 CA (capsid) protein, which has been shown to bind to a NUP153 FG-repeat via its intermolecular N-terminal/C-terminal interface [173][174].

NUP153 in flavivirus infection

To investigate the role NUP153 plays in flavivirus infection, mock-infected, TBEV-and LGTV-infected cells were explored using different methods, such as immunofluorescence, pulldown and Western blot, and so on.

Immunofluorescence and co-immunoprecipitation showed that NUP153 is upregulated in infected cells, and it is recruited to the viral replication site in the endoplasmic reticulum (ER) by viral proteins NS3 and NS5, where it binds viral RNA through its RNA-binding domain (Figure 19A). Partially knocking down NUP153 using CRISPR/Cas9 resulted in a significant decrease in LGTV infection, highlighting NUP153 as a pro-viral factor (Figure 19B-C).

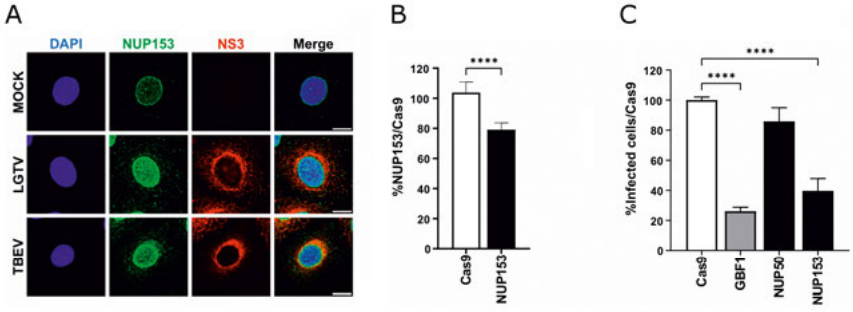


Figure 19: A) Immunofluorescence recorded from A549 cells, mock-infected or infected with LGTV or TBEV (MOI (multiplicity of infection)10). Shown in green: anti-NUP153 antibodies, red: anti-NS3. The scale bar is10 µm. B) % of NUP153 detected after transient CRISPR-Cas9 knock-down of NUP153 in HEK293T cells, measured using GFP-conjugated anti-NUP153 antibodies and FACS analysis. Cas9 is used as control. Error bars show ±SEM. ****, $p<0.0001$, calculated with t-test. C) Infection rates (LGTV, MOI1 for 48 h) after knock-down using FACS analysis and antibody against the viral E protein. GBF1 is positive control, NUP50 is unrelated NPC control. Error bars show ±SEM. ****, $p<0.0001$, calculated using ANOVA. Adapted from Paper IV.

AP-MS of NUP153 in mock-infected and flavivirus-infected cells

In order to gain insight into the potential changes in the interactome of NUP153 upon infection, I carried out AP-MS experiments with full-length, N-terminally EGFP-tagged NUP153 or only EGFP as control. HEK293T cells were transiently transfected to express these constructs and then either mock infected or infected with TBEV or LGTV. Subsequently, pulldowns using anti-GFP nanobodies on magnetic beads were carried out similarly as described in Paper II.

Comparing the enriched proteins from TBEV- and mock-infected samples I found that actin-related proteins, which are mostly located in the cytoplasm were enriched in the NUP153 pulldown from TBEV-infected cells. I also found several ribosomal and RNA-processing proteins, which aligns with the possible function of NUP153 in viral replication. HSPA5, a protein responsible for correct protein folding in the ER lumen, was also enriched from infected cell lysates (Figure 20).

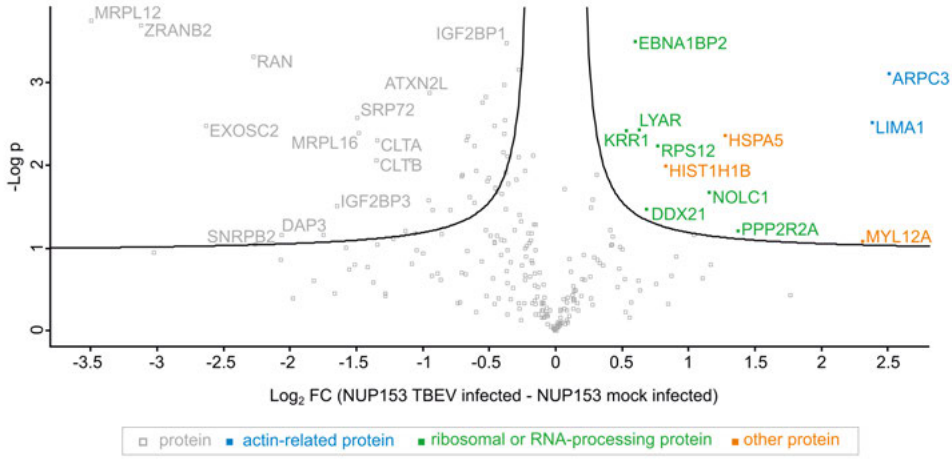


Figure 20: Quantitative MS results after pulldown using EGFP-tagged NUP153 from TBEV-infected or mock-infected HEK293 cell lysate in triplicates. After filtering, the significantly enriched proteins were determined with a t-test using permutation-based FDR (250 permutations, FDR<0.05, S0:0.1). Adapted from Paper IV.

We also searched for interactions between viral proteins and NUP153. Using the regular search settings, no viral proteins were found in the TBEV-infected samples. A special targeted viral data search was carried out, which indicated the presence of viral protein C, NS5, and NS3 in the TBEV-infected NUP153 pulldown samples. NS5 also showed up with 2 unique peptides in the TBEV-infected EGFP control samples. The LGTV-infected samples provided similar results in terms of pulldown and change in the interactome. The results were further validated by other experiments, suggesting a potential mechanism for translocation of NUP153 from the NPC.

Conclusion and outlook

In this study, we showed that NUP153 plays an important pro-viral role in TBEV and LGTV infections, where its expression is upregulated and it partially changes localization from the NPC to the cytoplasm and ER, to the viral replication site. The AP-MS results are supporting this, which demonstrate interactions with proteins located outside of the nucleus and many of them are involved in RNA processing.

Further studies of the mechanism of tick-borne flaviviruses may lead to the development of antivirals against these viruses, which we lack today.

Concluding remarks and future perspectives

SLiM-based protein-protein interactions are crucial for cellular functions and common amongst host-viral interactions. Despite their importance, this category of interactions is not well-studied and underrepresented in the literature. By establishing methods for SLiM-based interaction discovery and validation we can get a deeper understanding or interaction networks in health and disease.

The methods I evaluated in Paper I, biotinylated peptide-pulldown and PRISMA were to test their suitability to carry out pulldowns using SLiM-containing synthetic peptides. The peptide pulldown proved to be working over a range of affinities. However, the method is not very robust and in certain cases double repeat peptide sequences were needed to increase the local concentration of available motifs. The protein concentration in the cell lysate was also a crucial parameter, as we observed that the well-studied TP53-MDM2 interaction could only be captured when the MDM2 SWIB domain was spiked in the cell lysate. PRISMA, on the other hand, was not found to be a reliable method in our hands. Clearly, there is room for improvements in terms of how to capture SLiM-based interactions for MS analysis.

To further investigate methods for SLiM-pulldown, I also employed GFP-pulldown using N-terminally EGFP-tagged peptide repeats to great effect (Paper II). A peptide sequence from a Hendra virus protein, identified by ProP-PD, successfully pulled down the PABP homologues and other relevant proteins from cell lysate. This approach provided a setup with virtually no limit in the number of peptide sequence repeats.

In another approach, I used SPOT arrays with alanine scanning to identify key residues within a peptide sequence (Paper III). This is of high importance in cases where only a handful of peptides are enriched or known. A drawback of SPOT arrays is that they can only be used for one protein, and extensive use can become fairly costly. Emerging deep mutational scanning approaches may turn out to be more cost efficient and informative in a future perspective.

Lastly, I utilized the GFP-pulldown method with a full length, highly disordered protein, NUP153 (paper IV). This protein has been shown to interact with several viral proteins and changes its localization to the viral replication center in the case of Flavivirus infection. Comparing pulldowns from mock-, TBEV- and Langat-infected lysates, I noted shifts in the interactome of

NUP153. Clearly, much remains to be understood in terms of how viruses perturb the host cell interactome upon infection.

By performing this research, I was looking for the possibilities and limits of capturing SLiM-mediated interactions and contributed to a further understanding of viral-host interactions. As we learn more about protein disorder and SLiMs, both by high throughput experimental screens and bioinformatics predictions, establishing reliable and robust validation methods for PPIs become more and more important.

Populärvetenskaplig sammanfattning

Bakgrund
Brukar du fundera över hur människokroppen egentligen fungerar? Det krävs många olika molekyler som arbetar tillsammans. I min forskning har jag tittat på hur proteiner interagerar med varandra och bildar livsnödvändiga nätverk av interaktioner. Dessa interaktioner, eller bindingar, behövs för en rad olika funktioner, till exempel för att förmedla information inom och mellan celler, för att flytta proteiner dit de behövs eller för att märka proteiner som inte längre behövs för destruktion.

Interaktionerna kan ske mellan stora, strukturerade delar av proteiner (domäner) eller mellan en domän och ett litet, 3–10 aminosyror långt interaktionsmotiv. I min forskning har jag fokuserat på den senare kategorin, eftersom denna typ av interaktioner är mindre välstuderad och väldigt viktig för våra cellers funktion. På grund av den begränsade interaktionsytan är dessa interaktioner ofta svaga och dissocierar snabbt, men just detta är också egenskaper som behövs i flera processer i våra celler, exempelvis cellsignalering.

Många interaktioner mellan virusproteiner och människoproteiner är baserade på korta interaktionsmotiv (eng. short linear motifs, SLiMs). Virus har betydligt färre proteiner än oss människor. Medan vi människor har över 20 000 proteiner, har det virus som orsakar Covid-19 bara 29. Viruset kapar istället våra cellers funktion för att kunna föröka sig. Det börjar med ett virusprotein interagerar med ett människoprotein för att ta sig in i en cell.
Med ökad kunskap om vilka proteiner som interagerar kan vi förstå hälsa och ohälsa bättre och utveckla läkemedel mot exempelvis virusinfektioner.

Hur hittar vi interagerande proteiner?
När vi letar efter proteiner som interagerar med varandra, använder vi oftast ett protein som bete, och tillsätter det till en blandning av proteiner för att se vilka proteiner det interagerar med (byte). Tänk på det som fiske; bytet dras till betet, och när de interagerar kan vi ta bort det från de andra proteinerna. Men vi måste vara försiktiga, för om vi drar för hårt så kan vi förlora bytet, särskilt när det gäller svaga interaktioner baserade korta interaktionsmotiv.

I vissa av mina projekt använde jag antingen hela proteiner eller peptider (väldigt korta proteiner) som innehöll korta interaktionsmotiv som bete för att se om jag kunde fiska upp deras byten. I ett annat projekt undersökte jag exakt

vilka aminosyror som behövs i ett interaktionsmotiv genom att byta ut aminosyraresterna i en peptid en efter en.

För att ta reda på vilka proteiner som interagerar med betet använde jag masspektrometri. Det är en teknik som gör det möjligt att identifiera molekyler baserat på massa och kan ge information om de proteiner som finns i ett prov.

Vad har jag kommit fram till?

Peptider som sitter fast på magnetkulor fungerar bättre som beten i våra försök än de som sitter fast på cellulosamembran. Jag prövade olika peptider och kom fram till att även svaga interaktioner fångades upp. Ibland stötte jag dock på utmaningar när det t.ex. inte fanns tillräckligt mycket byte i provet jag försökte fiska i.

Covid-19-pandemin bröt ut inte så långt efter att jag påbörjat min doktorandtid. Eftersom det är vanligt med interaktionsmotivbaserade interaktioner mellan virusprotein och människoprotein bestämde vi oss för att titta på Covid-19 och andra relaterade virus. Vi hittade 1285 peptider från olika virus som interagerade med humana proteiner. Vissa humana proteiner fungerade som byten till flera liknande peptider från olika virus. Ett exempel som vi följde upp visade att vissa virus använder interaktionsmotiv för att ta kontroll över maskineriet som tillverkar nya protein. Det är viktigt, eftersom viruset kan minska produktionen av cellens egna proteiner och/eller använda det för att producera virusprotein.

Vi letade också efter interaktionsmotiv i människoproteiner som interagerar med proteiner från SARS-CoV-19. I vissa fall var vi osäkra på vilka aminosyror som var viktiga för interaktionen. Genom att byta ut aminosyrarester i peptiden en efter en kunde jag se vilka som var de viktigaste – när de byttes ut försvann interaktionen. En intressant upptäckt var att vissa av peptiderna hade antivirala egenskaper mot SARS-CoV-19 infektion, vilket innebär att de skulle kunna utvecklas till läkemedel.

Slutligen har jag studerat förändringar som sker vid infektioner med TBE (tick-borne encephalitis). Vi märkte att ett människoprotein, NUP153 (nuclear pore complex protein 153) byter plats i cellen när den infekteras, och att den interagerar med virusprotein. Jag hittade proteiner typiska för den nya placeringen som interagerade med NUP153 i infekterade prover, samt indikationer på virusproteiner.

Varför är det här viktigt?

Interaktionsmotivbaserade interaktioner kan vara kritiska för såväl normala cellfunktioner som virusinfektioner. Trots det vet vi väldigt lite om dem. Ökad kunskap om hur de fungerar och vilka proteiner som interagerar med varandra gör att vi kan göra mer för att förhindra och bota sjukdomar.

Acknowledgements

First, I would like to thank my supervisor, **Ylva**. Even though we did not choose each other, you guided me through the second half of my PhD and gave me all the support I could ask for. Thank you for welcoming me to your group and making me continue when I was ready to give up.
I am grateful that I could work the first half of my PhD under the guidance of **Sara**. Thank you for giving me the opportunity to do this PhD and for your compassion and always listening to me and helping me. I am so happy that you stayed in my "PhD-life" all along the way!
Jonas, I am thankful to have had you as my co-supervisor and for your input on my research as well as encouragement.
I would also like to thank **Evangelia, Norman, Per, Ola** for great discussions and collaboration in the SSF project – I think we made a great team!
Caroline, I cannot put into words such a huge support you have been to me over this PhD. You taught me most of the biochem lab skills I know, talked me through dark times, and also were there for the good ones ☺ I've found a real friend in you.
Showing the recruiting genius of our PIs, I also met you, **Marie**. Always welcome in your lab, you gave me support in everything – science and life. With you and Caroline by my side I felt like home in Uppsala from the beginning.
Johanna, I am so happy that we grew close, your energy and determination has inspired me a lot! And then there were the cakes…
Catia, my endless supporter. Thank you for the chats, hugs and cat-appointments. I am so lucky to have you around! **Ioanna**, thank you for the fun times and support. **Leo**, thank you for the chats, the jokes, and for organizing some much needed social activities.
To the whole Ivarsson group: **Gustav** and **Ali**, you taught me so much, and were always available for a good chat – or anxious questions. **Leandro**, you have never said no when I asked if you had a minute ☺ You showed me so many cool things, and brought a lot of happiness to this lab. **Emily**, it was so nice to see you gain confidence and then be my support when I needed you. We should do a Larnaca-Budapest exchange ☺ **Susanne**, thank you for the great chats and checking up on me. **Max, Priyanka** – you are super nice additions to the group, I wish you success in the future!
Filip, Debbie, Ravi, thank you for the collaborations, lab chats and all the help you gave me.

To the whole Hungarian team (**Petra, Kinga, Eszter, Gábor**) – thank you for letting me complain in my native tongue, reminisce about university times back in Budapest and sharing my interest in acquiring some Túró Rudi.
Varun, Lucie, Felix, my somewhat new officemates, thank you for tolerating me ☺
When nothing seemed to work, my lovely ex-office mates **Hilde, Malin** and **Claudia** made sure to hold me together and help me through the bumps on the road.
For all the people who were here at the very beginning, thank you: **Marcus, Neil, Kyle, Alicia, Kalliroi, Corinna**
Per, Ingela, Jeff, Kumari, Marit and Jean – thank you for all the things I learned from you, either in a class or over coffee.
Sandy, Mark, your office was always a place where I could discuss, share and learn, and also have fun.
Stacey, Sydney, Weifeng, Aman, Vlad, it is so nice that you joined us, good luck for everything!
Nastia, you are so kind and such a hard worker. I wish you luck for the future!
Helena, Sandra, Ahmed, Gun, Merve, Maxim, Louise, Farshid, Gunnar and everyone else in Chemistry-BMC, I always felt supported and welcome, wherever I went around the department. Thank you!
The MS facility gave me the opportunity to measure my samples. Anna, Alex, Levon, I learned so much from you and had a great time doing it. **Alexander, Sara J.** – thank you for working with us.
David, thank you for bringing joy in the office, reminding me that millennials are now old, and making my beautiful cover art for this thesis.
Luca, thank you for the life-saving phone calls and for our friendship, that started 12 years ago ☺
Balázs, Máté, Bence, Maszi, Zita, Zsuzsi, Andi – I would not be here now without you.
I'd like to also thank **my sister**s, who supported me and arranged everything back home when I couldn't.
Mama, neked is olyan sokat köszönhetek! Mindig jó látni téged mikor haza-megyek, tőled sokkal otthonosabb az az otthon.
Apa, te sajnos már nem lehettél itt velem. Köszönöm a biztatást, a segítséget és hogy megmutattad, hogy a tudomány szórakoztató is lehet.
Anya, ez most ide nem férne ki :D Köszönöm, hogy annyi mindenben segítettél, ott voltál a legboldogabb pillanatokban és akkor is, amikor minden összedőlni látszott. Megmutattad, hogy bármi lehetséges, és tanulni mindig érdemes.
Frog or dog? **Johan**, thank you for taking care of me, feeding me and being there for me in the hardest times. Being highly intelligent and kukuuu ☺ As you said, I'm not good at acknowledgements, but I love you. And thank you for reminding me that, after all, science is nothing but a game.

References

[1] R. J. Edwards and N. Palopoli, "Computational prediction of short linear motifs from protein sequences.," *Methods Mol. Biol.*, vol. 1268, pp. 89–141, 2015.

[2] R. Aasland *et al.*, "Normalization of nomenclature for peptide motifs as ligands of modular protein domains.," *FEBS Lett.*, vol. 513, no. 1, pp. 141–144, Feb. 2002.

[3] N. E. Davey *et al.*, "Attributes of short linear motifs," *Mol. Biosyst.*, vol. 8, no. 1, pp. 268–281, 2012.

[4] Y. Ivarsson and P. Jemth, "Affinity and specificity of motif-based protein–protein interactions," *Curr. Opin. Struct. Biol.*, vol. 54, pp. 26–33, 2019.

[5] K. Van Roey and N. E. Davey, "Motif co-regulation and co-operativity are common mechanisms in transcriptional , regulation," *Cell Commun. Signal.*, vol. 13, no. 45, 2015.

[6] N. E. Davey, G. Travé, and T. J. Gibson, "How viruses hijack cell regulation," *Trends Biochem. Sci.*, vol. 36, no. 3, pp. 159–169, 2011.

[7] T. Hagai, A. Azia, M. M. Babu, and R. Andino, "Use of host-like peptide motifs in viral proteins is a prevalent strategy in host-virus interactions," *Cell Rep.*, vol. 7, no. 5, pp. 1729–1739, 2014.

[8] G. Shuler and T. Hagai, "Rapidly evolving viral motifs mostly target biophysically constrained binding pockets of host proteins," *Cell Rep.*, vol. 40, no. 7, p. 111212, 2022.

[9] P. Tompa, N. E. Davey, T. J. Gibson, and M. M. Babu, "A Million peptide motifs for the molecular biologist," *Mol. Cell*, vol. 55, no. 2, pp. 161–169, 2014.

[10] I. M. A. Nooren and J. M. Thornton, "Diversity of protein-protein interactions," *EMBO J.*, vol. 22, no. 14, pp. 3486–3492, 2003.

[11] S. E. Acuner Ozbabacan, H. B. Engin, A. Gursoy, and O. Keskin, "Transient proteinprotein interactions," *Protein Eng. Des. Sel.*, vol. 24, no. 9, pp. 635–648, 2011.

[12] N. London, B. Raveh, and O. Schueler-Furman, "Druggable protein-protein interactions - from hot spots to hot segments," *Curr. Opin. Chem. Biol.*, vol. 17, no. 6, pp. 952–959, 2013.

[13] F. Da Silva, G. Bret, L. Teixeira, C. F. Gonzalez, and D. Rognan, "Exhaustive Repertoire of Druggable Cavities at Protein-Protein Interfaces of Known Three-Dimensional Structure," *J. Med. Chem.*, vol. 62, no. 21, pp. 9732–9742, 2019.

[14] J. C. Gaines, A. H. Clark, L. Regan, and C. S. O'Hern, "Packing in protein cores," *J. Phys. Condens. Matter*, vol. 29, no. 29, 2017.
[15] M. M. Babu, R. W. Kriwacki, and R. V. Pappu, "Versatility from protein disorder," *Science (80-.).*, vol. 337, no. 6101, pp. 1460–1461, 2012.
[16] S. Fukuchi, K. Hosoda, K. Homma, T. Gojobori, and K. Nishikawa, "Binary classification of protein molecules into intrinsically disordered and ordered segments," *BMC Struct. Biol.*, vol. 11, 2011.
[17] M. E. Oates *et al.*, "D2P2: Database of disordered protein predictions," *Nucleic Acids Res.*, vol. 41, no. D1, pp. 508–516, 2013.
[18] T. J. G. Francesca Diell, Niall Haslam, Claudia Chica, Aidan Budd, Sushama Michael, Nigel P. Brown, Gilles Travé, "Understanding eukaryotic linear motifs and their role in cell signaling and regulation," *Front. Biosci.*, vol. 13, pp. 6580–6603, 2008.
[19] N. E. Davey, D. C. Shields, and R. J. Edwards, "Masking residues using context-specific evolutionary conservation significantly improves short linear motif discovery," *Bioinformatics*, vol. 25, no. 4, pp. 443–450, 2009.
[20] H. Q. Nguyen *et al.*, "Quantitative mapping of protein-peptide affinity landscapes using spectrally encoded beads," *Elife*, vol. 8, pp. 1–28, 2019.
[21] K. Bugge *et al.*, "Interactions by Disorder – A Matter of Context," *Front. Mol. Biosci.*, vol. 7, no. June, pp. 1–16, 2020.
[22] H. Y. J. Fung, M. Birol, and E. Rhoades, "IDPs in macromolecular complexes: the roles of multivalent interactions in diverse assemblies," *Curr. Opin. Struct. Biol.*, vol. 49, pp. 36–43, 2018.
[23] F. Mihalic *et al.*, "Large-scale phage-based screening reveals extensive pan-viral mimicry of host short linear motifs," *bioRxiv*, p. 2022.06.19.496705, Jan. 2022.
[24] N. E. Davey, J. L. Cowan, D. C. Shields, T. J. Gibson, M. J. Coldwell, and R. J. Edwards, "SLiMPrints: Conservation-based discovery of functional motif fingerprints in intrinsically disordered protein regions," *Nucleic Acids Res.*, vol. 40, no. 21, pp. 10628–10641, 2012.
[25] J. Jumper *et al.*, "Highly accurate protein structure prediction with AlphaFold," *Nature*, vol. 596, no. 7873, pp. 583–589, 2021.
[26] M. Akdel *et al.*, "A structural biology community assessment of AlphaFold 2 applications," *bioRxiv*, p. 2021.09.26.461876, 2021.
[27] M. Kumar *et al.*, "The Eukaryotic Linear Motif resource: 2022 release," *Nucleic Acids Res.*, vol. 50, no. D1, pp. D497–D508, 2022.
[28] D. E. Scott, A. R. Bayly, C. Abell, and J. Skidmore, "Small molecules, big targets: Drug discovery faces the protein-protein interaction challenge," *Nat. Rev. Drug Discov.*, vol. 15, no. 8, pp. 533–550, 2016.
[29] M. Konopleva *et al.*, "MDM2 inhibition: an important step forward in cancer therapy," *Leukemia*, vol. 34, no. 11, pp. 2858–2874, 2020.

[30] B. Higgins *et al.*, "Abstract B55: Antitumor activity of the MDM2 antagonist RG7388.," *Mol. Cancer Ther.*, vol. 12, no. 11_Supplement, pp. B55–B55, Nov. 2013.
[31] X. Dai, O. Hakizimana, X. Zhang, A. C. Kaushik, and J. Zhang, "Orchestrated efforts on host network hijacking: Processes governing virus replication," *Virulence*, vol. 11, no. 1, pp. 183–198, 2020.
[32] J. A. Hiscox, "RNA viruses: Hijacking the dynamic nucleolus," *Nat. Rev. Microbiol.*, vol. 5, no. 2, pp. 119–127, 2007.
[33] J. Martin-Serrano, T. Zang, and P. D. Bieniasz, "Role of ESCRT-I in Retroviral Budding," *J. Virol.*, vol. 77, no. 8, pp. 4794–4804, 2003.
[34] J. E. Garrus *et al.*, "Tsg101 and the vacuolar protein sorting pathway are essential for HIV-1 budding," *Cell*, vol. 107, no. 1, pp. 55–65, 2001.
[35] O. Pornillos, S. L. Alam, D. R. Davis, and W. I. Sundquist, "Structure of the tsg101 uev domain in complex with the ptap motif of the hiv-1 p6 protein," *Nat. Struct. Biol.*, vol. 9, no. 11, pp. 812–817, 2002.
[36] "SARS-CoV-2 reference genome ASM985889v3." [Online]. Available: https://www.ncbi.nlm.nih.gov/assembly/GCF_009858895.2/.
[37] D. E. Gordon *et al.*, "A SARS-CoV-2 protein interaction map reveals targets for drug repurposing," *Nature*, vol. 583, no. 7816, pp. 459–468, 2020.
[38] "Human reference genome GRCh38.p14." [Online]. Available: https://www.ncbi.nlm.nih.gov/assembly/GCF_000001405.40/.
[39] N. E. Davey, M. S. Cyert, and A. M. Moses, "Short linear motifs – ex nihilo evolution of protein regulation," *Cell Commun. Signal.*, vol. 13, no. 1, p. 43, 2015.
[40] Á. O. Toole *et al.*, "Tracking the international spread of SARS-CoV-2 lineages B . 1 . 1 . 7 and B . 1 . 351 / 501Y-V2 with grinch [version 2 ; peer review : 3 approved] Network for Genomic Surveillance in South Africa (NGS-SA), Danish Covid-19 Genome Consortium (DCGC), ," pp. 1–19, 2022.
[41] P. V'kovski, A. Kratzel, S. Steiner, H. Stalder, and V. Thiel, "Coronavirus biology and replication: implications for SARS-CoV-2," *Nat. Rev. Microbiol.*, vol. 19, no. 3, pp. 155–170, 2021.
[42] H. E. Field, "Hendra virus ecology and transmission," *Curr. Opin. Virol.*, vol. 16, pp. 120–125, 2016.
[43] R. Lindqvist, A. Upadhyay, and A. K. Överby, "Tick-borne flaviviruses and the type I interferon response," *Viruses*, vol. 10, no. 7, pp. 1–21, 2018.
[44] M. Haglund and G. Günther, "Tick-borne encephalitis - Pathogenesis, clinical course and long-term follow-up," *Vaccine*, vol. 21, no. SUPPL. 1, pp. S11–S18, 2003.
[45] G. Dobler, "Zoonotic tick-borne flaviviruses," *Vet. Microbiol.*, vol. 140, no. 3–4, pp. 221–228, 2010.

[46] G. P. Smith, “Filamentous fusion phage: Novel expression vectors that display cloned antigens on the virion surface,” *Science (80-.).*, vol. 228, no. 4705, pp. 1315–1317, 1985.
[47] J. Mccafferty, A. D. Griffiths, G. Winter, and D. J. Chiswellt, “Phage antibodies: filamentous variable domains,” *Nature*, vol. 348, no. December, pp. 552–554, 1990.
[48] P. Malik *et al.*, “Role of capsid structure and membrane protein processing in determining the size and copy number of peptides displayed on the major coat protein of filamentous bacteriophage,” *J. Mol. Biol.*, vol. 260, no. 1, pp. 9–21, 1996.
[49] S. S. Sidhu, H. B. Lowman, B. C. Cunningham, and J. A. Wells, “Phage display for selection of novel binding peptides,” *Methods Enzymol.*, vol. 328, pp. 333–363, 2000.
[50] S. S. Sidhu, B. Li, Y. Chen, F. A. Fellouse, C. Eigenbrot, and G. Fuh, “Phage-displayed antibody libraries of synthetic heavy chain complementarity determining regions,” *J. Mol. Biol.*, vol. 338, no. 2, pp. 299–310, 2004.
[51] Y. Ivarsson *et al.*, “Large-scale interaction profiling of PDZ domains through proteomic peptide-phage display using human and viral phage peptidomes,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 111, no. 7, pp. 2542–2547, 2014.
[52] N. E. Davey *et al.*, “Discovery of short linear motif-mediated interactions through phage display of intrinsically disordered regions of the human proteome,” *FEBS J.*, vol. 284, no. 3, pp. 485–498, 2017.
[53] C. Benz *et al.*, “Proteome□scale mapping of binding sites in the unstructured regions of the human proteome,” *Mol. Syst. Biol.*, vol. 18, no. 1, pp. 1–27, 2022.
[54] Perrin, Francis, “Polarisation de la lumière de fluorescence. Vie moyenne des molécules dans l’etat excité,” *J. Phys. Radium*, vol. 7, no. 12, pp. 390–401, 1926.
[55] W. J. Checovich, R. E. Bolger, and T. Burke, “Fluorescence polarization — a new tool for cell and molecular biology,” *Nature*, vol. 375, no. 6528, pp. 254–256, 1995.
[56] A. Velazquez-Campoy and E. Freire, “ITC in the post-genomic era...? Priceless,” *Biophys. Chem.*, vol. 115, no. 2-3 SPEC. ISS., pp. 115–124, 2005.
[57] A. Frostell, L. Vinterbäck, and H. Sjöbom, “Protein-Ligand Interactions Using SPR Systems.,” *Methods Mol. Biol.*, vol. 1008, pp. 139–165, 2013.
[58] S. Charbonnier, K. Zanier, M. Masson, and G. Travé, “Capturing protein-protein complexes at equilibrium: The holdup comparative chromatographic retention assay,” *Protein Expr. Purif.*, vol. 50, no. 1, pp. 89–101, 2006.

[59] G. G. Boglarka Zambo, Bastien Morlet, Luc Negroni, Gilles Trave, "Native holdup (nHU) to measure binding affinities from cell extracts," *bioRxiv*, 2022.
[60] J. R. Lundblad, M. Laurance, and R. H. Goodman, "Fluorescence of Protein-DNA Interactions Polarization Analysis and Protein-Protein," *Library (Lond).*, pp. 607–612, 2008.
[61] A. M. Rossi and C. W. Taylor, "Analysis of protein-ligand interactions by fluorescence polarization," *Nat. Protoc.*, vol. 6, no. 3, pp. 365–387, 2011.
[62] A. Q. Emili and G. Cagney, "Large-scale functional analysis using peptide or protein arrays," *Nat. Biotechnol.*, vol. 18, no. 4, pp. 393–397, 2000.
[63] L. C. Szymczak, H. Y. Kuo, and M. Mrksich, "Peptide Arrays: Development and Application," *Anal. Chem.*, vol. 90, no. 1, pp. 266–282, 2018.
[64] G. B. Bolger *et al.*, "Scanning peptide array analyses identify overlapping binding sites for the signalling scaffold proteins, β-arrestin and RACK1, in cAMP-specific phosphodiesterase PDE4D5," *Biochem. J.*, vol. 398, no. 1, pp. 23–36, 2006.
[65] X. Li *et al.*, "A scanning peptide array approach uncovers association sites within the JNK/βarrestin signalling complex," *FEBS Lett.*, vol. 583, no. 20, pp. 3310–3316, 2009.
[66] K. F. Houslay, F. Christian, R. MacLeod, D. R. Adams, M. D. Houslay, and G. S. Baillie, "Identification of a multifunctional docking site on the catalytic unit of phosphodiesterase-4 (PDE4) that is utilised by multiple interaction partners," *Biochem. J.*, vol. 474, no. 4, pp. 597–609, 2017.
[67] O. Söderberg *et al.*, "Characterizing proteins and their interactions in cells and tissues using the in situ proximity ligation assay," *Methods*, vol. 45, no. 3, pp. 227–232, 2008.
[68] I. Weibrecht *et al.*, "Proximity ligation assays: A recent addition to the proteomics toolbox," *Expert Rev. Proteomics*, vol. 7, no. 3, pp. 401–409, 2010.
[69] C. Benz, E. Kassa, E. Tjärnhage, S. B. Lind, and Y. Ivarsson, "Chapter 1 Identification of Cellular Protein–Protein Interactions," in *Inhibitors of Protein–Protein Interactions: Small Molecules{,} Peptides and Macrocycles*, The Royal Society of Chemistry, 2021, pp. 1–39.
[70] W. H. Dunham, M. Mullin, and A. C. Gingras, "Affinity-purification coupled to mass spectrometry: Basic principles and strategies," *Proteomics*, vol. 12, no. 10, pp. 1576–1590, 2012.
[71] A. C. Gingras, M. Gstaiger, B. Raught, and R. Aebersold, "Analysis of protein complexes using mass spectrometry," *Nat. Rev. Mol. Cell Biol.*, vol. 8, no. 8, pp. 645–654, 2007.
[72] E. L. Huttlin *et al.*, "Dual proteome-scale networks reveal cell-specific remodeling of the human interactome," *Cell*, vol. 184, no. 11, pp. 3022-3040.e28, 2021.

[73] M. Y. Hein *et al.*, "A Human Interactome in Three Quantitative Dimensions Organized by Stoichiometries and Abundances," *Cell*, vol. 163, no. 3, pp. 712–723, 2015.
[74] E. C. Keilhauer, M. Y. Hein, and M. Mann, "Accurate Protein Complex Retrieval by Affinity Enrichment Mass Spectrometry (AE-MS) Rather than Affinity Purification Mass Spectrometry (AP-MS)," *Mol. Cell. Proteomics*, vol. 14, no. 1, pp. 120–135, 2015.
[75] T. P. Hopp *et al.*, "A short polypeptide marker sequence useful for recombinant protein identification and purification," *Bio/Technology*, vol. 6, no. 10, pp. 1204–1210, 1988.
[76] J. Field *et al.*, "Purification of a RAS-responsive adenylyl cyclase complex from Saccharomyces cerevisiae by use of an epitope addition method," *Mol. Cell. Biol.*, vol. 8, no. 5, pp. 2159–2165, 1988.
[77] M. C. Hillman *et al.*, "A comprehensive system for protein purification and biochemical analysis based on antibodies to c-myc peptide," *Protein Expr. Purif.*, vol. 23, no. 2, pp. 359–368, 2001.
[78] M. Chalfie, Y. Tu, G. Euskirchen, W. W. Ward, and D. C. Prasher, "Green fluorescent protein as a marker for gene expression," *Science (80-.).*, vol. 263, no. 5148, pp. 802–805, 1994.
[79] D. B. Smith and K. S. Johnson, "Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase," *Gene*, vol. 67, no. 1, pp. 31–40, 1988.
[80] G. Dittmar *et al.*, "PRISMA: Protein Interaction Screen on Peptide Matrix Reveals Interaction Footprints and Modifications- Dependent Interactome of Intrinsically Disordered C/EBPβ," *iScience*, vol. 13, pp. 351–370, 2019.
[81] "JPT website about PRISMA." [Online]. Available: https://www.jpt.com/protein-interaction-screen-on-peptide-matrix-prisma/11.
[82] K. Meyer *et al.*, "Mutations in Disordered Regions Can Cause Disease by Creating Dileucine Motifs," *Cell*, vol. 175, no. 1, pp. 239-253.e17, 2018.
[83] E. Ramberger *et al.*, "PRISMA and BioID disclose a motifs-based interactome of the intrinsically disordered transcription factor C/EBPα," *iScience*, vol. 24, no. 6, p. 102686, 2021.
[84] E. Ramberger *et al.*, "A Universal Peptide Matrix Interactomics Approach to Disclose Motif-Dependent Protein Binding Authors A Universal Peptide Matrix Interactomics Approach to Disclose Motif-Dependent Protein Binding," *Mol Cell Proteomics*, vol. 20, p. 100135, 2021.
[85] K. J. Roux, D. I. Kim, M. Raida, and B. Burke, "A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells," *J. Cell Biol.*, vol. 196, no. 6, pp. 801–810, 2012.

[86] J. D. Martell *et al.*, "Engineered ascorbate peroxidase as a genetically encoded reporter for electron microscopy," *Nat. Biotechnol.*, vol. 30, no. 11, pp. 1143–1148, 2012.
[87] S. S. Lam *et al.*, "Directed evolution of APEX2 for electron microscopy and proximity labeling," *Nat. Methods*, vol. 12, no. 1, pp. 51–54, 2014.
[88] T. C. Branon *et al.*, "Efficient proximity labeling in living cells and organisms with TurboID," *Nat. Biotechnol.*, vol. 36, no. 9, pp. 880–898, 2018.
[89] A. N. Holding, "XL-MS: Protein cross-linking coupled with mass spectrometry," *Methods*, vol. 89, pp. 54–63, 2015.
[90] A. Sinz, "Cross-Linking/Mass Spectrometry for Studying Protein Structures and Protein–Protein Interactions: Where Are We Now and Where Should We Go from Here?," *Angew. Chemie - Int. Ed.*, vol. 57, no. 22, pp. 6390–6396, 2018.
[91] B. W. Sutherland, J. Toews, and J. Kast, "Utility of formaldehyde cross-linking and mass spectrometry in the study of protein-protein interactions," *J. Mass Spectrom.*, vol. 43, no. 6, pp. 699–715, 2008.
[92] A. J. Lomant and G. Fairbanks, "Chemical probes of extended biological structures: Synthesis and properties of the cleavable protein cross-linking reagent [35S]dithiobis(succinimidyl propionate)," *J. Mol. Biol.*, vol. 104, no. 1, pp. 243–261, 1976.
[93] L. Piersimoni, P. L. Kastritis, C. Arlt, and A. Sinz, "Cross-Linking Mass Spectrometry for Investigating Protein Conformations and Protein-Protein Interactions-A Method for All Seasons," *Chem. Rev.*, vol. 122, no. 8, pp. 7500–7531, 2022.
[94] L. Fornelli *et al.*, "Top-down proteomics: Where we are, where we are going?," *J. Proteomics*, vol. 175, pp. 3–4, 2018.
[95] L. C. Gillet, A. Leitner, and R. Aebersold, "Mass Spectrometry Applied to Bottom-Up Proteomics: Entering the High-Throughput Era for Hypothesis Testing," *Annu. Rev. Anal. Chem.*, vol. 9, pp. 449–472, 2016.
[96] A. E. Speers and C. C. Wu, "Proteomics of integral membrane proteins - Theory and application," *Chem. Rev.*, vol. 107, no. 8, pp. 3687–3714, 2007.
[97] S. Goldberg, "Mechanical/physical methods of cell disruption and tissue homogenization.," *Methods Mol. Biol.*, vol. 424, pp. 3–22, 2008.
[98] M. M. Bradford, "A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding.," *Anal. Biochem.*, vol. 72, pp. 248–254, 1976.
[99] P. Y. Lee, J. Osman, T. Y. Low, and R. Jamal, "Plasma/serum proteomics: Depletion strategies for reducing high-abundance proteins for biomarker discovery," *Bioanalysis*, vol. 11, no. 19, pp. 1799–1812, 2019.

[100] S. Doll and A. L. Burlingame, “Mass spectrometry-based detection and assignment of protein posttranslational modifications,” *ACS Chem. Biol.*, vol. 10, no. 1, pp. 63–71, 2015.
[101] M. Mann and O. N. Jensen, “Proteomic analysis of post-translational modifications,” *Nat. Biotechnol.*, vol. 21, pp. 255–261, 2003.
[102] E. Vandermarliere, M. Mueller, and L. Martens, “Getting intimate with trypsin, the leading protease in proteomics,” *Mass Spectrom. Rev.*, vol. 32, no. 6, pp. 453–465, 2013.
[103] J. R. Wiśniewski and M. Mann, “Consecutive proteolytic digestion in an enzyme reactor increases depth of proteomic and phosphoproteomic analysis,” *Anal. Chem.*, vol. 84, no. 6, pp. 2631–2637, 2012.
[104] R. M. Tubaon, P. R. Haddad, and J. P. Quirino, “Sample Clean-up Strategies for ESI Mass Spectrometry Applications in Bottom-up Proteomics: Trends from 2012 to 2016,” *Proteomics*, vol. 17, no. 20, 2017.
[105] J. Rappsilber, M. Mann, and Y. Ishihama, “Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips,” *Nat. Protoc.*, vol. 2, no. 8, pp. 1896–1906, 2007.
[106] S. R. Wilson, T. Vehus, H. S. Berg, and E. Lundanes, “Nano-LC in proteomics: Recent advances and approaches,” *Bioanalysis*, vol. 7, no. 14, pp. 1799–1815, 2015.
[107] M. S. Wilm and M. Mann, “Electrospray and Taylor-Cone theory, Dole’s beam of macromolecules at last?,” *Int. J. Mass Spectrom. Ion Process.*, vol. 136, no. 2–3, pp. 167–180, 1994.
[108] K. L. Sanders and J. L. Edwards, “Nano-liquid chromatography-mass spectrometry and recent applications in omics investigations,” *Anal. Methods*, vol. 12, no. 36, pp. 4404–4417, 2020.
[109] J. B. Fenn, M. Mann, C. K. A. I. Meng, S. F. Wong, and C. M. Whitehouse, “Electrospray Ionization for Mass Spectrometry of Large Biomolecules,” *Science (80-.).*, vol. 246, pp. 64–71, 1989.
[110] P. Kebarle and U. H. Verkcerk, “Electrospray: From Ions in solution to Ions in the gas phase, what we know now,” *Mass Spectrom. Rev.*, vol. 28, no. 6, pp. 898–917, 2009.
[111] E. Aliyari and L. Konermann, “Formation of Gaseous Peptide Ions from Electrospray Droplets: Competition between the Ion Evaporation Mechanism and Charged Residue Mechanism,” *Anal. Chem.*, vol. 94, no. 21, pp. 7713–7721, 2022.
[112] A. Makarov, “Electrostatic axially harmonic orbital trapping: A high-performance technique of mass analysis,” *Anal. Chem.*, vol. 72, no. 6, pp. 1156–1162, 2000.
[113] J. V. Olsen, B. Macek, O. Lange, A. Makarov, S. Horning, and M. Mann, “Higher-energy C-trap dissociation for peptide modification analysis,” *Nat. Methods*, vol. 4, no. 9, pp. 709–712, 2007.

[114] J. E. P. Syka, J. J. Coon, M. J. Schroeder, J. Shabanowitz, and D. F. Hunt, "Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 101, no. 26, pp. 9528–9533, 2004.
[115] R. A. Zubarev, N. L. Kelleher, and F. W. McLafferty, "Electron capture dissociation of multiply charged protein cations. A nonergodic process," *J. Am. Chem. Soc.*, vol. 120, no. 13, pp. 3265–3266, 1998.
[116] S. E. Ong and M. Mann, "Mass Spectrometry–Based Proteomics Turns Quantitative," *Nat. Chem. Biol.*, vol. 1, no. 5, pp. 252–262, 2005.
[117] M. Bantscheff, S. Lemeer, M. M. Savitski, and B. Kuster, "Quantitative mass spectrometry in proteomics: Critical review update from 2007 to the present," *Anal. Bioanal. Chem.*, vol. 404, no. 4, pp. 939–965, 2012.
[118] Y. Zhang, B. R. Fonslow, B. Shan, M.-C. Baek, and J. R. Yates III, "Protein analysis by shotgun proteomics," *Chem. Rev.*, vol. 113, pp. 2343–2394, 2013.
[119] J. Villanueva, M. Carrascal, and J. Abian, "Isotope dilution mass spectrometry for absolute quantification in proteomics: Concepts and strategies," *J. Proteomics*, vol. 96, pp. 184–199, 2014.
[120] S. Tate, B. Larsen, R. Bonner, and A. C. Gingras, "Label-free quantitative proteomics trends for protein-protein interactions," *J. Proteomics*, vol. 81, pp. 91–101, 2013.
[121] J. Cox, M. Y. Hein, C. A. Luber, I. Paron, N. Nagaraj, and M. Mann, "Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ," *Mol. Cell. Proteomics*, vol. 13, no. 9, pp. 2513–2526, 2014.
[122] S.-E. Ong *et al.*, "Stable Isotope Labeling by Amino Acids in Cell Culture, SILAC, as a Simple and Accurate Approach to Expression Proteomics," *Mol. Cell. Proteomics*, vol. 1, no. 5, pp. 376–386, 2002.
[123] J. Cox and M. Mann, "MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification," *Nat. Biotechnol.*, vol. 26, no. 12, pp. 1367–1372, 2008.
[124] S. Tyanova *et al.*, "The Perseus computational platform for comprehensive analysis of (prote)omics data," *Nat. Methods*, vol. 13, no. 9, pp. 731–740, 2016.
[125] P. Mortensen *et al.*, "MSQuant, an open source platform for mass spectrometry-based quantitative proteomics," *J. Proteome Res.*, vol. 9, no. 1, pp. 393–403, 2010.
[126] C. Lazar, L. Gatto, M. Ferro, C. Bruley, and T. Burger, "Accounting for the Multiple Natures of Missing Values in Label-Free Quantitative Proteomics Data Sets to Compare Imputation Strategies," *J. Proteome Res.*, vol. 15, no. 4, pp. 1116–1125, 2016.
[127] D. Mellacheruvu *et al.*, "The CRAPome: A contaminant repository for affinity purification-mass spectrometry data," *Nat. Methods*, vol. 10, no. 8, pp. 730–736, 2013.

[128] H. Choi *et al.*, "SAINT: Probabilistic scoring of affinity purificationg-mass spectrometry data," *Nat. Methods*, vol. 8, no. 1, pp. 70–73, 2011.
[129] G. Teo, G. Liu, J. Zhang, A. I. Nesvizhskii, A. C. Gingras, and H. Choi, "SAINTexpress: Improvements and additional features in Significance Analysis of INTeractome software," *J. Proteomics*, vol. 100, pp. 37–43, 2014.
[130] M. E. Sowa, E. J. Bennett, S. P. Gygi, and J. W. Harper, "Defining the Human Deubiquitinating Enzyme Interaction Landscape," *Cell*, vol. 138, no. 2, pp. 389–403, 2009.
[131] S. Jäger *et al.*, "Global landscape of HIV-human protein complexes," *Nature*, vol. 481, no. 7381, pp. 365–370, 2012.
[132] R. Oughtred *et al.*, "The BioGRID database: A comprehensive biomedical resource of curated protein, genetic, and chemical interactions," *Protein Sci.*, vol. 30, no. 1, pp. 187–200, 2021.
[133] S. Orchard *et al.*, "The MIntAct project - IntAct as a common curation platform for 11 molecular interaction databases," *Nucleic Acids Res.*, vol. 42, no. D1, pp. 358–363, 2014.
[134] G. Alanis-Lobato, M. A. Andrade-Navarro, and M. H. Schaefer, "HIPPIE v2.0: Enhancing meaningfulness and reliability of protein-protein interaction networks," *Nucleic Acids Res.*, vol. 45, no. D1, pp. D408–D414, 2017.
[135] B. Snel, G. Lehmann, P. Bork, and M. A. Huynen, "String: A web-server to retrieve and display the repeatedly occurring neighbourhood of a gene," *Nucleic Acids Res.*, vol. 28, no. 18, pp. 3442–3444, 2000.
[136] D. Szklarczyk *et al.*, "The STRING database in 2021: Customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets," *Nucleic Acids Res.*, vol. 49, no. D1, pp. D605–D612, 2021.
[137] D. W. Huang *et al.*, "DAVID Bioinformatics Resources: Expanded annotation database and novel algorithms to better extract biology from large gene lists," *Nucleic Acids Res.*, vol. 35, no. SUPPL.2, pp. 169–175, 2007.
[138] K. Van Roey, T. J. Gibson, and N. E. Davey, "Motif switches: Decision-making in cell regulation," *Curr. Opin. Struct. Biol.*, vol. 22, no. 3, pp. 378–385, 2012.
[139] M. E. Cockman *et al.*, "Widespread hydroxylation of unstructured lysine-rich protein domains by JMJD6," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 119, no. 32, pp. 1–10, 2022.
[140] S. Hanke and M. Mann, "The Phosphotyrosine Interactome of the Insulin Receptor Family and Its Substrates IRS-1 and IRS-2," *Mol. Cell. Proteomics*, vol. 8, no. 3, pp. 519–534, 2009.
[141] O. Trejo-Cerro, P. Massimi, J. Broniarczyk, M. Myers, and L. Banks, "Repression of Memo1, a Novel Target of Human Papillomavirus Type 16 E7, Increases Cell Proliferation in Cervical Cancer Cells," *J. Virol.*, 2022.

[142] M. Uhlén *et al.*, "Tissue-based map of the human proteome," *Science (80-.).*, vol. 347, no. 6220, 2015.
[143] S. E. Ong, "Unbiased identification of protein-bait interactions using biochemical enrichment and quantitative proteomics," *Cold Spring Harb. Protoc.*, vol. 5, no. 3, pp. 1–7, 2010.
[144] K. Malecek and A. Ruthenburg, *Validation of histone-binding partners by peptide pull-downs and isothermal titration calorimetry*, 1st ed., vol. 512. Elsevier Inc., 2012.
[145] R. Klingberg *et al.*, "Analysis of phosphorylation-dependent protein-protein interactions of histone H3," *ACS Chem. Biol.*, vol. 10, no. 1, pp. 138–145, 2015.
[146] M. Vermeulen, *Identifying chromatin readers using a SILAC-based histone peptide pull-down approach*, 1st ed., vol. 512. Elsevier Inc., 2012.
[147] W. Müller *et al.*, "Attempts to Mimick Docking Processes of the Immune System: Recognition-Induced," *Science (80-.).*, vol. 262, pp. 1706–1708, 1993.
[148] J. D. Hirsch *et al.*, "Easily reversible desthiobiotin binding to streptavidin, avidin, and other biotin-binding proteins: Uses for protein labeling, detection, and isolation," *Anal. Biochem.*, vol. 308, no. 2, pp. 343–357, 2002.
[149] R. Behrendt, P. White, and J. Offer, "Advances in Fmoc solid-phase peptide synthesis," *J. Pept. Sci.*, vol. 22, no. 1, pp. 4–27, 2016.
[150] J. Ma *et al.*, "PALB2 Interacts with KEAP1 To Promote NRF2 Nuclear Accumulation and Function," *Mol. Cell. Biol.*, vol. 32, no. 8, pp. 1506–1517, 2012.
[151] N. D. Camp *et al.*, "Wilms tumor gene on X chromosome (WTX) inhibits degradation of NRF2 protein through competitive binding to KEAP1 protein," *J. Biol. Chem.*, vol. 287, no. 9, pp. 6539–6550, 2012.
[152] D. F. Lee *et al.*, "KEAP1 E3 Ligase-Mediated Downregulation of NF-κB Signaling by Targeting IKKβ," *Mol. Cell*, vol. 36, no. 1, pp. 131–140, 2009.
[153] S. C. Lo and M. Hannink, "PGAM5, a Bcl-XL-interacting protein, is a novel substrate for the redox-regulated Keap1-dependent ubiquitin ligase complex," *J. Biol. Chem.*, vol. 281, no. 49, pp. 37893–37903, 2006.
[154] P. H. Kussie *et al.*, "Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain," *Science (80-.).*, vol. 274, no. 5289, pp. 948–953, 1996.
[155] J. Lin, J. Chen, B. Elenbaas, and A. J. Levine, "Several hydrophobic amino acids in the p53 amino-terminal domain are required for transcriptional activation, binding to mdm-2 and the adenovirus 5 E1B 55-kD protein," *Genes Dev.*, vol. 8, no. 10, pp. 1235–1246, 1994.

[156] T. Juven-Gershon, O. Shifman, T. Unger, A. Elkeles, Y. Haupt, and M. Oren, "The Mdm2 Oncoprotein Interacts with the Cell Fate Regulator Numb," *Mol. Cell. Biol.*, vol. 18, no. 7, pp. 3974–3982, 1998.
[157] X. Wu, J. H. Bayle, D. Olson, and A. J. Levine, "The p53-mdm-2 autoregulatory feedback loop," *GENES Dev.*, vol. 53, no. 7, pp. 1126–1132, 1993.
[158] Z. Li, R. M. Adams, K. Chourey, G. B. Hurst, R. L. Hettich, and C. Pan, "Systematic comparison of label-free, metabolic labeling, and isobaric chemical labeling for quantitative proteomics on LTQ orbitrap velos," *J. Proteome Res.*, vol. 11, no. 3, pp. 1582–1590, 2012.
[159] B. McDonald and J. Martin-Serrano, "No strings attached: The ESCRT machinery in viral budding and cytokinesis," *J. Cell Sci.*, vol. 122, no. 13, pp. 2167–2177, 2009.
[160] R. J. Jackson, C. U. T. Hellen, and T. V. Pestova, "The mechanism of eukaryotic translation initiation and principles of its regulation," *Nat. Rev. Mol. Cell Biol.*, vol. 11, no. 2, pp. 113–127, 2010.
[161] M. C. Derry, A. Yanagiya, Y. Martineau, and N. Sonenberg, "Regulation of poly(A)-binding protein through PABP-interacting proteins," *Cold Spring Harb. Symp. Quant. Biol.*, vol. 71, pp. 537–543, 2006.
[162] B. P. Cormack, R. H. Valdivia, and S. Falkow, "FACS-optimized mutants of the green fluorescent protein (GFP)," *Gene*, vol. 173, no. 1, pp. 33–38, 1996.
[163] T. Kruse *et al.*, "The Ebola Virus Nucleoprotein Recruits the Host PP2A-B56 Phosphatase to Activate Transcriptional Support Activity of VP30," *Mol. Cell*, vol. 69, no. 1, pp. 136-145.e6, 2018.
[164] T. Kruse *et al.*, "Large scale discovery of coronavirus-host factor protein interaction motifs reveals SARS-CoV-2 specific mechanisms and vulnerabilities," *Nat. Commun.*, vol. 12, no. 1, pp. 1–13, 2021.
[165] D. R. Littler, B. S. Gully, R. N. Colson, and J. Rossjohn, "Crystal Structure of the SARS-CoV-2 Non-structural Protein 9, Nsp9," *iScience*, vol. 23, no. 7, p. 101258, 2020.
[166] J. P. Davies, K. M. Almasy, E. F. McDonald, and L. Plate, "Comparative Multiplexed Interactomics of SARS-CoV-2 and Homologous Coronavirus Nonstructural Proteins Identifies Unique and Shared Host-Cell Dependencies," *ACS Infect. Dis.*, vol. 6, no. 12, pp. 3174–3189, 2020.
[167] J. Li *et al.*, "Virus-Host Interactome and Proteomic Survey Reveal Potential Virulence Factors Influencing SARS-CoV-2 Pathogenesis," *Med*, vol. 2, no. 1, pp. 99-112.e7, 2021.
[168] Y. Zhou *et al.*, "A comprehensive SARS-CoV-2–human protein–protein interactome reveals COVID-19 pathobiology and potential host therapeutic targets," *Nat. Biotechnol.*, 2022.

[169] M. Beck and E. Hurt, “The nuclear pore complex: Understanding its function through structural insight,” *Nat. Rev. Mol. Cell Biol.*, vol. 18, no. 2, pp. 73–89, 2017.
[170] S. Cohen, S. Au, and N. Panté, “How viruses access the nucleus,” *Biochim. Biophys. Acta - Mol. Cell Res.*, vol. 1813, no. 9, pp. 1634–1645, 2011.
[171] A. Schmitz *et al.*, “Nucleoporin 153 arrests the nuclear import of hepatitis B virus capsids in the nuclear basket,” *PLoS Pathog.*, vol. 6, no. 1, 2010.
[172] Z. Onder and J. Moroianu, “Nuclear import of cutaneous beta genus HPV8 E7 oncoprotein is mediated by hydrophobic interactions between its zinc-binding domain and FG nucleoporins,” *Virology*, vol. 449, pp. 150–162, 2014.
[173] X. Li, A. Engelman, K. A. Matreyek, and S. S. Yu, “Nucleoporin NUP153 Phenylalanine-Glycine Motifs Engage a Common Binding Pocket within the HIV-1 Capsid Protein to Mediate Lentiviral Infectivity,” vol. 9, no. 10, 2013.
[174] S. Li, S. Patel, J. Yang, and M. Crabtree, “Defining the HIV Capsid Binding Site of Nucleoporin 153,” *Msph. 10.1128/msphere.00310-22*, 2022.

Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 2204

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-486917

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2022