Charles University in Prague

Faculty of Science

Study programme: Biophysical chemistry

Study branch: Physical chemistry

Bc. Jakub Staníček

Characterizing DDI2 protein interaction by solution NMR

Charakterizace interakce proteinu DDI2 pomocí NMR spektroskopie

Diploma thesis

Supervisor: Mgr. Klára Grantz Šašková, Ph.D.

Prague 2019

Prohlašuji, že jsem tuto diplomovou práci vypracoval samostatně pod vedením své školitelky Mgr. Kláry Grantz Šaškové, Ph.D. pokud není uvedeno jinak a že všechny použité prameny jsem řádně citoval. Práce nebyla použita k získání stejného nebo podobného titulu.

(I declare that this thesis has not been previously submitted for any degree, that this is my own work under supervision of Mgr. Klára Grantz Šašková, Ph.D. unless stated otherwise and that all used material has been properly acknowledged.)

In Prague ……

Jakub Staníček

Acknowledgement

I would like to thank all the people who made writing of this thesis possible. Firstly, I am hugely grateful to my family and my dear girlfriend, as it was them who motivated me and supported me during this rather difficult experience.

Further thanks go to the members of the Ddi team, who shared all the joys and the difficulties of scientific work with me during the last three years. Special mention is for Michal Svoboda, who taught me large part of what I know about science, even though he would surely disagree, and who spent many long evenings discussing my results with me.

Big thanks to my supervisor Klára, for taking me to her lab to work on a very interesting project, for introducing me to the scientific work and for helping me during the whole process of putting this thesis together.

Thanks to Jan Konvalinka, as I made most of the work on this thesis in his lab on IOCB, and to all members of his lab for being a great collective to be part of.

Special thanks go to Václav Veverka, for consultation during NMR measurement and for controlling my spectral assignments, and to Pavel Srb, who spent long hours consulting my experiments, fitting the spectra and explaining my endless questions about mysteries of NMR.

Acknowledged institutions:

PŘÍRODOVĚDECKÁ
FAKULTA
Univerzita Karlova

ÚOCHB AV ČR
IOCB PRAGUE

Table of contents

Abstract

Human DDI2 protein is a dimeric aspartic protease that has been recently found to play an important role in DNA damage repair and transcriptional regulation of the proteasome expression. Current insights into the mechanistic details of both functions are still quite limited.

We have previously identified the human RAD23B protein to interact with the DDI2 protein. RAD23B also functions in DNA damage repair as part of the XPC complex that stimulates the nucleotide excision repair activity. Moreover, RAD23B participates as an adaptor protein in the process of protein degradation. Therefore, the interaction of DDI2 and RAD23B might have important implications for both known functions of DDI2.

This work describes the DDI2 and RAD23B interaction on the structural level. Recombinant protein variants of both DDI2 and RAD23B proteins were prepared and the interaction was mapped by the affinity pull-down assay. Protein NMR titrations were further used to explore the interaction.

Key words: ubiquitin-proteasome system, DNA damage repair, proteasome expression regulation, aspartyl protease, DDI2, NMR

Abstrakt

Lidský protein DDI2 je dimerní aspartátová proteasa, o které bylo nedávno zjištěno, že hraje důležitou roli při opravě poškození DNA a transkripční regulaci proteasomální exprese. Současné poznatky o molekulárních mechanismech obou funkcí jsou však stále poměrně omezené.

Již dříve jsme identifikovali lidský protein RAD23B jako interaktant DDI2. RAD23B také funguje při opravě poškození DNA jako součást komplexu XPC, který stimuluje aktivitu opravy excize nukleotidů. Navíc se RAD23B účastní jako adaptorový protein v procesu degradace proteinů. Interakce DDI2 a RAD23B proto může mít důležité důsledky pro obě známé funkce DDI2.

Tato práce popisuje interakci DDI2 a RAD23B na strukturální úrovni. Byly připraveny rekombinantní proteinové varianty DDI2 a RAD23B a interakce byla zmapována pomocí afinitní purifikace. K prozkoumání interakce byly dále použity NMR titrace.

Klíčová slova: ubikvitin-proteasomální systém, oprava poškození DNA, regulace exprese proteasomu, aspartátová proteasa, DDI2, NMR

(In Czech)

Abbreviations

A_{280} – absorbance of solution at 280 nm

AAA – ATPase protein family, ATPases Associated with various cellular Activities

ATP - adenosine-5'-triphosphate

BSA – bovine serum albumin

bp – base pair

bZIP – basic-region leucine zipper

CNC – cap ‘n’ collar

CSP – chemical shift perturbation

CP – 20S Core Particle of the 26S proteasome

Ddi – DNA damage-inducible protein

DNA – deoxyribonucleic acid

dNTP – deoxynucleotide triphosphate

E. coli – Escherichia coli

EDTA – ethylenediaminetetraacetic acid

ER - endoplasmic reticulum

ERAD - endoplasmic reticulum-associated degradation

FPLC – fast-performance liquid chromatography

HABA - 2-[4 -hydroxy-benzeneazo]benzoic acid

HDD – Helical domain of Ddi

HIV – human immunodeficiency virus

HU - hydroxyurea

IPTG – isopropyl-β-D-thiogalactopyranosid

K_d – dissociation constant

LB – lysogeny broth (also known as Luria-Bertani broth)

MTOC – microtubule organising centre

NER – nucleotide excision repair

NMR – nuclear magnetic resonance

PAGE – polyacrylamide gel electrophoresis

PCR – polymerase chain reaction

PDB – protein data bank

Pull-down – affinity chromatography interaction assay

RP – 19S Regulatory Particle of the 26S proteasome

RPI – relative peak intensity

RVP – retroviral-like protease

SDS – sodium dodecyl sulphate

Sti1-like – similar to yeast protein Sti1

TEMED – N′,N′,N′,N′- tetramethylethylendiamine

T_m – melting temperature

Tris – tris(hydroxymethyl)aminomethane

UBA – ubiquitin-associated domain

UBD – ubiquitin-binding domain

UBL – ubiquitin-like domain

UBX – ubiquitin-regulatory X

UDP – ubiquitin domain protein

UIM – ubiquitin-interaction motif

ULM – ubiquitin-like modifier

UPS – ubiquitin-proteasome system

XPCb – XPC-binding

v/v – volume ratio

w/v – weight to volume ratio

Yeast cell lines with inactivated gene are marked in italic with gene name in small letters preceded by a delta symbol, i.e. *Δddi*.

NMR spectra: with exception of HSQC, all abbreviations used for the NMR spectra refer to functional groups participating in magnetisation transfer during measurement.

HSQC – heteronuclear single quantum coherence

CA – alpha carbon

CB – beta carbon

HN – amide group of the peptide bond

CO – carbonyl group of the peptide bond

1 Introduction

The focus of this thesis represents the human DNA Damage-Inducible Protein homolog 2 (hDDI2) and the detailed characterization of its interaction with RAD23B protein that will enable to further elucidate the molecular basis of DDI2 cellular function. DDI2 has escaped larger scientific attention for a long time until few years ago when the first discoveries started to uncover its cellular roles. Nevertheless, a lot still remains to be solved.

First insights into hDDI2 function are interesting and make DDI2 worth further investigation. One of those is its dimeric domain located at the C-terminal part which displays a striking similarity to aspartyl protease from the Human Immunodeficiency Virus (HIV) and has all structural properties of an active aspartyl protease, though it was not proved to be proteolytically active *in vitro* at that time[1]. Later on, DDI2 has been found to play an important role in two areas closely connected with cancer: DNA damage repair[2], and regulation of proteotoxic stress response[3]. In DNA damage repair, DDI2 is one of the key proteins that allow the restart of stalled replisomes, a crucial process whose absence can lead to severe DNA damage upon exposure to replication-stalling agents[2]. During the regulation of proteotoxic stress response, DDI2 acts as an activator of cellular response to proteasome inhibition through activation of TCF11/Nrf1 transcription factor upregulating proteasome genes for proteasome resynthesis. It is the DDI2 protease domain that is crucial for this activation by proteolytically cleaving the transcription factor into its active form. The stress response mediated by DDI2 enables certain types of cancer, such as multiple myeloma and mantel cell lymphoma, to resist chemotherapy based on proteasome inhibitors and thus makes DDI2 a novel therapeutic target in blood cancer treatment[3].

The thesis will mainly focus on the DDI2 function in proteotoxic stress response. To put it into a broad context however, more general information about cellular pathways that hDDI2 relates to will be referred to form a necessary background: these are namely the Ubiquitin Proteasome System (UPS), principles of ubiquitin signalling and Endoplasmic Reticulum-Associated Degradation of proteins (ERAD pathway).

Regarding the experimental part, the work was initially started by my former lab colleague Jan Belza. Jan used proteomic mass spectrometry techniques to detect interaction of hDDI2 with human UV excision repair protein RAD23 homolog B (hRAD23B), which he further verified by immunoprecipitation in the cellular context[4]. In this thesis, the interaction sites of DDI2-RAD23B are mapped and further characterized by nuclear

magnetic resonance (NMR) spectroscopy. Since hRAD23B protein similarly to hDDI2 participates both in the ERAD pathway and in DNA damage repair, the detailed biophysical interaction described further might be an important piece of puzzle in reconstructing the overall picture of the molecular mechanisms of hDDI2.

2 Aims of the thesis

The thesis focuses on achieving the following goals:

1) Provide a thorough summary of knowledge concerning the human DDI2 protein incorporated in the wider context
2) Experimentally examine the interaction between hDDI2 and human protein RAD23B:
    a. by pull-down experiments with recombinant DDI2 and RAD23B to map the interaction
    b. by protein nuclear magnetic resonance analysis to characterize DDI2-RAD23B interaction in detail
3) Discuss the obtained experimental results with the literature and suggest possible future directions for the following research

3 Theoretic background

3.1. Principles of ubiquitin signalling

Ubiquitin is a small, 76 amino acid long protein (8.6 kDa), highly evolutionary conserved in eukaryotic organisms and expressed in all types of cells and tissues, as its name suggests. In human, it is encoded by several genes: *RPS27A* and *UBA52* encoding ubiquitin fused to ribosomal proteins and *UBB* and *UBC* genes encoding polyubiquitin. Ubiquitin fulfils a signaling role by being fused to cellular proteins. Ubiquitin is then recognized by ubiquitin-binding factors, altering fate of the modified protein. Most prominent function of ubiquitination is targeting proteins for degradation in the proteasome.

Ubiquitin has an extremely stable fold that is resistant to both chemical and thermal denaturation over the wide pH range, with melting temperature exceeding 90°C at neutral pH[5–7]. Its characteristic fold consists of five-stranded β-sheet, α-helix of 3.5 turns and a short four-residue 3_{10}-helix. The β-sheet adopts a concave shape and the α-helix fits into its groove, tightly bound through hydrophobic interactions, a conformation frequently referred to as "β-grasp fold"[8,9] (see Figure 1).

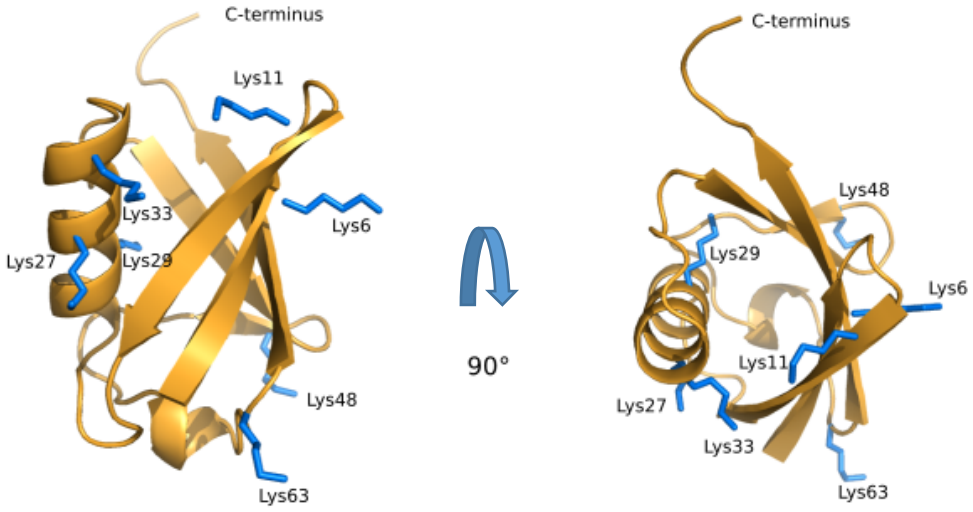

Figure 1: Structure of human ubiquitin (PDB code 1UBQ)[8]. Lysin sidechains are highlighted in blue. The picture was created in PyMol software (Schrodinger, LLC).

There are several features of the ubiquitin structure which are important: six C-terminal residues forming a flexible tail, seven lysine residues and N-terminal methionine (all carrying free amino group), and two hydrophobic surfaces centred on isoleucine 44 (Ile44-patch) and isoleucine 36 (Ile-36 patch), respectively[10].

The hydrophobic patches act as interaction surfaces, responsible for recognition of ubiquitin by other molecules. Through its C-terminus, ubiquitin is fused to other proteins due to an enzymatic action of specialised enzymes (see further), forming the isopeptide bond between its terminal carboxyl group and a free amino group on targeted protein, usually a lysine sidechain. The combination of ubiquitin's ability to form bonds with amino groups of other proteins, combined with the fact that ubiquitin itself carries seven lysine sidechains and a free N-terminus, is crucial for ubiquitin's ability to form various chains.

3.1.1. Polyubiquitin chains

Even though protein ubiquitination starts as a fusion of a single ubiquitin molecule to a lysine sidechain (monoubiquitination), this modification is often expanded through the formation of polyubiquitin chain. During chain formation, ubiquitin itself is modified with other ubiquitin moieties on either of its seven lysine residues or on its N-terminus, while the location of inter-ubiquitin fusion is crucial to the resultant chain topology. Connection through different lysine residues causes different orientation of consecutive ubiquitins, with intermolecular interactions between chain constituents contributing to specific conformations for each type of linkage. This results in different topologies of polyubiquitin chains with different relative positions and solvent exposure of hydrophobic patches, thus presenting a unique interaction surface. This is then recognised by various ubiquitin-binding modules, with specific chain surfaces enabling to distinguish one linkage type from other. As notable example can serve the comparison of Lys48 and Lys63-linked ubiquitin chains: where Lys63-linkage adopts a loose structure with conformational flexibility between ubiquitin molecules, Lys48-linkage forms a compact condensed conformation stabilised by interaction of hydrophobic patches of constituent ubiquitins.

Polyubiquitin chains can be categorised into several groups: single linkage (homotypic) chains, mixed (heterotypic) chains and branched (or forked) chains. Branched chains are special case of mixed chain that appears when single ubiquitin moiety within a chain is ubiquitinated at more than one lysine at once. The best explored types of chains are the homotypic chains, with each linkage having their own set of known functions, as each linkage is utilised for signalling in different context of cellular activity. However, mixed linkages increase the amount of possible ubiquitin signals enormously and evidence of their importance is still growing[10,11].

3.1.2. Regulation of ubiquitination

Regulation of this incredibly variable system of protein modification is achieved due to a large number of specific enzymes providing two functions that counteract each other: ubiquitination and deubiquitination.

Ubiquitination is executed in a process of three consecutive steps called E1-E2-E3 cascade, each step performed by specialised class of enzymes[12]. The first step is called ubiquitin activation and it is carried out by an enzyme called E1 (2 E1 enzymes in humans), which uses energy of ATP for creation of ubiquitin~E1 complex connected via macroergic thioester bond. In the next step, activated ubiquitin is transferred from E1 enzyme to ubiquitin conjugating E2 enzyme, binding to cysteine in its active site. After being charged with ubiquitin, E2 enzyme passes the activated ubiquitin onto E3 enzyme, also known as ubiquitin ligase, which is responsible for transferring the activated ubiquitin molecule onto its final target. While there is typically only one E1 enzyme (two in humans) and several tens of E2 enzymes, number of E3 ligases goes to hundreds. It is the task of a particular E3 enzyme to specifically recognise the target ubiquitination site and bring the specificity of the linkage during polyubiquitin chain formation, thus the need for large number of E3s with variable modes of action[10,12].

As the opposite to E1-E2-E3 cascade act the deubiquitinases (DUBs). DUBs represent the group of cysteine proteases and metalloproteases that are able to recognise and cleave isopeptide bond at the C-terminus of ubiquitin. They are typically divided into two groups based on their specificity: the first group represents the substrate-specific DUBs, cleaving polyubiquitin regardless of the type of linkage. They, however, need to be targeted to their substrate through interaction with other domains or for example by an adaptor protein. Second group are the linkage-specific DUBs that cleave all polyubiquitin of a specific type but are unable to cleave other types of linkages[10].

3.1.3. Ubiquitin binding domains

Each E3 ligase or DUB in order to create and modify specific ubiquitin chain needs to be able to interact with ubiquitin and distinguish between various types of chain. Similarly, turning different ubiquitin modifications into signals of a particular cellular pathway requires a binding of various factors to these signals in a selective manner. This role is often played by Ubiquitin Binding Domains (UBDs).

Over 20 different families of UBDs have been identified so far, possessing several distinct modes of ubiquitin binding. Even though the structures of UBDs do vary, several

groups possessing similar structures can be identified. The largest group consists of various α-helical structures, ranging from single helical motifs to bundles of multiple helices. Next important groups represent zinc finger motives with affinity to ubiquitin and ubiquitin receptors with pleckstrin-homology folds[13].

The first studied were the α-helical UBDs. One of them are the Ubiquitin-Associated Domains (UBA), which were initially detected by bioinformatic analysis of multiple E2, E3 enzymes, DUBs and other proteins related to the ubiquitin system[14]. Typically, UBA is approximately a 45 amino-acid long domain, consisting of three helices bound together by a hydrophobic core. It can be found within various proteins from the E3 ligases and deubiquitinases to several kinases. Even though the known structures of the UBA domains have generally very similar fold, their modes of ubiquitin binding can differ, especially their preference for distinct polyubiquitin chain types[13].

Other very important UBDs are the Ubiquitin-Interacting Motives (UIMs), short sequences of approximately 20 amino acids with canonical sequence identified as Φ-x-x-Ala-x-x-x-Ser-x-x-Ac, where Φ stands for a large hydrophobic residue and Ac stands for acidic residue[15]. The first two UIMs were identified within the ubiquitin binding proteasome subunit S5a[16]. Many others were identified based on sequence homology in various other proteins participating in the ubiquitin proteasome system, lysosomal degradation and vesicular transport[15]. The UIM motif is generally too short to form an independent tertiary structure, but it forms an α-helix that can be incorporated into various protein folds[15]. Apart from canonical UIM many related short helical sequences with affinity to ubiquitin exist, such as MIU (Motif Interacting with Ubiquitin) or DUIM (Double Ubiquitin Interacting Motif), with very similar functions[13] .

The binding affinity of the most UBDs to monoubiquitin is quite low, with dissociation constants (K_d) usually in the range of 10-500 μM[17]. This is understandable in context of ubiquitin concentration within intracellular environment, which is in low micromolar range[18,19]. Furthermore, the low affinities to monoubiquitin protect UBDs from unselective binding to free ubiquitin[17]. High affinity interactions in ubiquitin signalling are achieved through a different principle, which also allows for a high selectivity: the principle of avidity[20]. Avidity stands for the combination of multiple low affinity interactions that result in much stronger interaction by synergy. The combination of several interaction sites is very often seen in ubiquitin recognition, where two or more UBD/UIM domains are frequently present within one protein[13,20]. Similar principle was already discussed above regarding the binding surfaces presented by different polyubiquitin chains.

To name a few human protein examples, tandem UIMs are present in proteasomal ubiquitin receptor subunit S5a and deubiquitinase Ataxin-3, both showing selectivity for the Lys48-linked chains. The DNA damage-repair protein Rap80 and its tandem UIMs selectively binding to the Lys63-linked chains serves as another example[21–24]. It is very common for the UIMs to appear in a tandem, connected by a flexible linker of various length[13]. The length of a linker is usually crucial for the specificity of the chain recognition, because it allows the UIMs to be oriented in a proper way in relation to the position of the binding sites within the polyubiquitin chain. It was shown, that short linker from Ataxin-3 orients UIMs in a way that they fit onto the interaction sites in the Lys48-linked chain and the long linker of Rap80 does the same for the interaction with more extended Lys63-linked chains. The importance of a linker for a chain selectivity was further accented when it was shown that replacing the Rap80 linker with the linker from Ataxin-3 shifts the preference of Rap80 from Lys63-linked chains towards Lys48-linked chains[24].

Alternatively to tandem UBDs, some proteins achieve avidity with a single UBD that contains multiple ubiquitin-binding surfaces. This is the case for the human DNA-repair proteins RAD23A and RAD23B or the yeast Ddi1 protein, whose C-terminal UBA domains possess two ubiquitin binding sites. Individually, each binding site has relatively low affinity for ubiquitin with K_d approximately 400 μM, but through proper spatial orientation they are able to cooperate when binding to the Lys48-linked polyubiquitin, lowering the effective K_d under 10 μM[25,26]. In contrast, UBA domain from the yeast protein Dsk2 and its human homolog Ubiquilin1 possess only one ubiquitin-binding site and thus shows no preference for one type of polyubiquitin[27–29].

3.2. Proteasome

The best described function of protein ubiquitination is protein targeting for degradation. Properly ubiquitinated cytoplasmic proteins are transported to and degraded in the specialised compartmentalised protease complex, the 26S proteasome (proteasome in short). Yeast proteasome will be described here, but human proteasomes are very similar, only the naming differs[30].

3.2.1. Structure of the proteasome

Proteasome is a large protein complex, residing within cytoplasm and nucleus. It comprises of 33 distinct protein subunits that closely cooperate during recognition and

proteolytic cleavage of a substrate. The 26S proteasome can be divided into two main parts: the 20S Core Particle (CP) and the 19S Regulatory Particle (RP), with the CP forming central barrel-like structure, that can be capped on either one or both sides by the RP (Figure 2)[31].

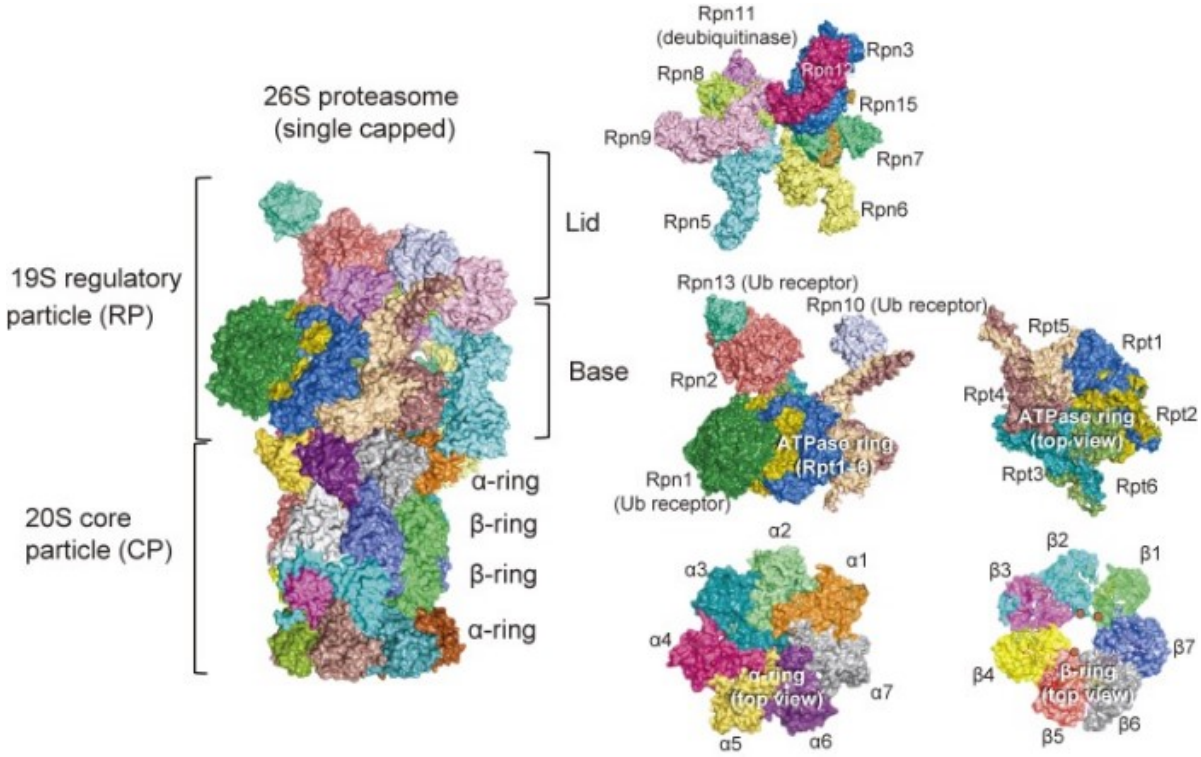

Figure 2: Structure of the yeast proteasome with individual subunits indicated by a different colour coding. Proteolytic active sites on the inside of the β-ring are marked by red dots. Adopted from publication[31].

The CP consists of four rings, each containing seven distinct protein subunits. There are two types of such rings, denoted α and β, that are stacked onto each other in α-β-β-α manner, thus creating a hollow barrel. Seven subunits of α-ring are called α1- α7, and seven subunits of β-ring are called β1- β7. The CP carries the proteolytic centre of the proteasome, created by proteolytic active sites on subunits β1, β2 and β5 in the middle of the barrel[32]. Polypeptides that enter the CP are digested into small fragments ranging from 2 to 10 amino-acid residues[33]. The entrance into the CP is too narrow to allow folded proteins in and is further covered by the RP[32,34].

The RP fulfils the non-proteolytic roles of the proteasome that means the substrate recognition and translocation of the polypeptide chain inside the CP for degradation. It is composed of two subcomplexes: the base and the lid[31].

The base contains hexameric AAA-ATPase complex of six homologous subunits Rpt1-Rpt6, which form a ring that covers the end of the CP barrel. This ATPase is able to

use energy from the ATP for translocation of polypeptide chains from outside the proteasome into the CP through the centre of the AAA-ATPase ring[35]. Other parts of the base subcomplex are four non-ATPase subunits: Rpn1, Rpn10 and Rpn13, that exhibit binding affinity for ubiquitin and act as receptors for ubiquitinated substrates, and Rpn2, a scaffold protein that contacts the lid subcomplex[36–38].

The lid subcomplex comprises of nine non-ATPase subunits: Rpn3, Rpn5-9, Rpn11, Rpn12 and Rpn15. Rpn5 and Rpn6 stabilise interaction of RP with CP by making contacts with α-ring of CP[35]. Important part of the lid subcomplex is Rpn11, which carries a metalloprotease domain and acts as integral proteasome deubiquitinase that removes ubiquitin moieties from polypeptide substrate while it is translocated into the proteasome[39,40].

3.2.2. Function of the proteasome

The CP is quite passive during the process of protein degradation, its purpose is to keep proteolytically active sites concealed within the barrel to prevent unspecific proteolysis. Without the RP, entry into the proteolytic cavity is closed. Once the substrate enters the cavity, it is proteolytically cleaved in exergonic manner, without need for additional energy[31,34,41].

Degradation of proteins in the proteasome is energetically demanding process, requiring input of ATP[41]. It needs at least one third of energy required for its synthesis, without taking additional ATP used for ubiquitination into account. This is caused by action of the RP, which performs complex tasks to ensure selectivity of the degradation process[30,42].

Once the interaction between a substrate and the ATPase ring occurs, cyclic conformation changes of the ring fuelled by the ATP hydrolysis exert a force that moves the polypeptide chain through the central pore of the ring. Nature of this interaction that needs to bind a wide variety of substrates stays elusive, even though the hydrophobic tyrosine residues directed into the lumen of the pore are likely important[30,43,44]. Movement of the RP base also influences the CP and causes widening of the CP entry[34,45,46].

Before ubiquitinated substrate interacts with the ATPase ring, it needs to bind to the ubiquitin receptor outside the entry into the proteasome[41]. The Lys48-linked chains were established historically as the signal for proteasomal degradation, with a minimal tetraubiquitin Lys48-chain[47,48]. However, it was shown that intact proteasomes bind to both Lys48 and Lys63-linked polyubiquitin chains *in vitro* without strong preference, while they

do not bind homotypic Lys11-linked chains[41,49–51]. Also, degradation of substrates modified with the Lys48 or Lys63-chains have comparable rates *in vitro*. In contrast, upon proteasome inhibition, all types of ubiquitin chains accumulate *in vivo*, which might suggest their role in degradation, with exception of the Lys63-linked chains[52]. This might be due to other factors binding to Lys63-chains within context of their specific function, which outcompete the affinity of the proteasome for these chains *in vivo*[53]. The reasons for these discrepancies stay unclear[30].

Recently it was shown, contradicting the previous reports[48], that even multiple short ubiquitin chains or even multiple monoubiquitination can lead to degradation[54,55]. A concept of "ubiquitination threshold" was established, meaning that not a specific type of ubiquitination but rather high enough amount of attached ubiquitin molecules is needed for degradation[11]. This is due to a competition between the interaction with the proteasome and simultaneous deubiquitination occurring within the proteasome proximity[30,56].

There are two steps of a substrate binding: the first is a reversible binding of ubiquitin chains to ubiquitin receptors, the second step is an interaction of a substrate with the ATPase ring, which is irreversible and requires ATP[41]. After the substrate reversibly binds to the proteasome, a relatively long time is needed for the ATPase ring to engage the substrate. In the meantime, ubiquitination of the substrate can decrease to the extent that it dissociates from the ubiquitin receptor and the degradation does not occur[30,56]. Thus, more numerous or longer ubiquitin chains increase the dwell-time of a substrate on the proteasome, increasing the likelihood of degradation[57].

Deubiquitination of a substrate associated with the proteasome is performed by the proteasome-associated deubiquitinases. In yeast there are two of them, Rpn11 that is an integral part of the lid subcomplex, and Ubp6 (USP14 in human), that binds to Rpn1 through its N-terminal ubiquitin-like (UBL) domain (more about UBL domains in section 3.3.2)[39,40,56,58,59]. In mammals, one more proteasome-associated deubiquitinase UCH37 is known, which binds to Rpn13 homolog ADRM1[60,61]. Ubp6/USP14 and UCH37 engage the substrate during the first phase of reversible binding. If it takes too long for the ATPase ring to engage or the ubiquitination of the substrate is low, they can prevent the degradation through substrate dissociation[30,56]. Rpn11 on the other hand doesn't engage ubiquitin chains until the substrate is irreversibly bound in the ATPase pore. It is positioned within the lid subcomplex directly over the ATPase pore, its active site only opens periodically after the ATPase ring activation which fuels cyclic conformational changes in the whole RP[39,40,45]. As bulky ubiquitin chains can clog the proteasome and

prevent substrate from translocation, the role of Rpn11 is to remove the chains that remain on the substrate after translocation starts and to facilitate the substrate entry into the ATPase pore[45]. All three deubiquitinases prevent degradation of ubiquitin molecules and recycle them for further use. Apart from that their functions differ: while inactivation of Rpn11 blocks the degradation, inactivation of Ubp6/USP14 and UCH37 promotes the rate of substrate degradation[30,45]. Further, Ubp6/USP14 and UCH37 have the ability to allosterically promote the ATPase activity of the RP when they engage ubiquitin chains on the substrate. This mechanism allows basal ATPase activity of the proteasome to be low and increase only after the substrate is bound[30,34,46].

Historically, it was believed that disordered parts of a protein are what targets proteins for degradation, due to an observation that the more ordered the proteins are, the longer are their lifetimes[62]. After the discovery of ubiquitin role in degradation, the theory of specific ubiquitin degradation signal was preferred. Recently, both factors are considered to participate in determining the protein lifetime within cells[30,41,63]. The substrate interaction with the ATPase ring pore seems to be the rate-limiting step in proteasomal degradation and the unfolded protein sequence containing exposed hydrophobic sites is needed for the ATPase ring interaction with the substrate[64,65]. More ordered protein sequence needs more ATP to be degraded, and takes longer to engage by the ATPase ring, giving more time to the substrate to be deubiquitinated and dissociated from the proteasome[42]. The role of ubiquitination is to keep the substrate positioned at the proteasome long enough for the ATPase to engage it[30]. Ubiquitin might also help by destabilising the fold of the ubiquitinated protein, promoting the degradation[66]. Often, ubiquitin-dependent AAA-ATPase Cdc48/p97 works as a defoldase in cooperation with the proteasome to facilitate degradation (in bigger detail in chapter 3.4.2.)[67].

3.3. Ubiquitin-like proteins

Although ubiquitin has its characteristic β-grasp fold, it is not the only protein adopting this kind of structure. This structure is widespread within both eukaryotes and prokaryotes. While in prokaryotes its function includes both the interactive and the enzymatic function, in eukaryotes only the function of an interaction module is conserved[9]. The group of the β-grasp fold harbouring proteins is called the Ubiquitin-like (UBL) proteins[68,69], although some authors suggested more descriptive name of the Ubiquitin-fold

(UFD) proteins, due to the fact that even though their fold resembles that of ubiquitin, their sequence similarity to ubiquitin is generally low[70,71].

This group of proteins can be further divided into two subgroups. The first group form the Ubiquitin-like Modifiers (ULMs), which undergo an enzymatic cascade and are conjugated to other proteins. The second group are proteins containing ubiquitin-like fold, but they don't get fused to other proteins and usually contain the β-grasp fold domain together with multiple other domains. This group of proteins is called the Ubiquitin-Domain Proteins (UDPs)[69,71–73].

3.3.1. Ubiquitin-like modifiers

The ULMs are functioning in a very similar way to ubiquitin: they are small proteins adopting the β-grasp fold with extended C-terminus that can be conjugated to amino groups of other proteins. This is done through a similar enzymatic cascade to ubiquitin, with the activating E1 enzyme, conjugating E2 enzyme and with E3 ligase targeting ULM to its final target. Each ULM has its own set of enzymes independent from each other. As the number of targets of the most ULMs is considerably lower than for ubiquitin, the number of E2 and E3 enzymes dedicated for a particular ULM is lower as well in comparison to ubiquitin[69,74].

Typical examples of ULMs are the proteins from the SUMO family, regulating various stress responses, Nedd8 that can be attached to ubiquitin E3 ligases and regulates their activity, or protein Atm8, which is involved in autophagy initiation[74].

3.3.2. Ubiquitin-domain proteins

The second group of the proteins harbouring the β-grasp fold are the Ubiquitin-domain proteins (UDPs). Unlike ULMs, no conjugation to other proteins occurs through UDPs, and the β-grasp fold is found here as an independent integral domain within a context of a multidomain protein. Three main categories of these integral β-grasp fold domains have been established so far: the Ubiquitin-like (UBL) domains, the Ubiquitin-regulatory X (UBX) domains and the Phox and Bem1 (PB1) domains. The UBL domains were recognised first, based on their homology with ubiquitin. The UBX and the PB1 domains were described later and classified separately, because despite adopting the β-grasp fold their sequence homology with ubiquitin is negligible[73,75].

The UBL domains are the most numerous and the best studied of the three groups. They are typically located at the N-terminus of a protein and have high sequence similarity with ubiquitin. They can be often found within various E3 ligases, deubiquitinases and

proteasome-associated factors, proteins involved in proteasomal degradation pathway, further accenting their connection to ubiquitin[75]. The first UBL domain was detected in the yeast protein Rad23. Interestingly, this UBL domain can be replaced with the ubiquitin sequence without affecting the Rad23 function[76]. As more UBL domains were identified, their ability to bind to the ubiquitin-binding subunits of the proteasome emerged as their frequent property, even though not all UBLs are able to do so[73]. Examples of the proteins carrying the proteasome-binding UBLs are the yeast proteasome-associated protein Rad23 and Dsk2 and their human homologues, the human E3 ligase Parkin or the deubiquitinase Usp14. Besides proteins involved in proteasomal-degradation pathway, UBLs can be found in several immune-response-associated kinases[75].

The β-grasp fold often acts as an interaction module. Apart from the well-known association of many UBLs with the regulatory subunit of the proteasome, many UBX domains mediate the interaction with the AAA-ATPase Cdc48/p97, a multifunctional defoldase and segregase[77,78]. The best characterized protein with the UBX domain, p47 from yeast, uses this domain to bind the AAA-ATPase Cdc48/p97 to cooperate with it in membrane fusion events. The human transmembrane protein UBXD8 uses its UBX domain to summon Cdc48/p97 to the cytosolic side of the ER membrane[78–80]. The UBLs often bind to UBA domains within the context of the proteasome-degradation pathway[73,81].

The β-grasp fold domain can also regulate activity of an enzyme they are part of. The human E3 ubiquitin ligase Parkin, a mitophagy regulator, needs to be ubiquitinated in order to be active. It contains UBL domain that inhibits this ubiquitination, thus downregulating Parkin activity[82]. Furthermore, this UBL domain can be phosphorylated, which results in increase of Parkin enzymatic activity[83]. Similar case can be found in the human ubiquitin-specific-protease (USP) family of deubiquitinases, many of which carry a UBL domain[71,84]. The USP14 uses its UBL domain to bind to the proteasome, which leads to allosteric activation of USP14 due to a conformation change[59,85]. USP4 catalytic activity is regulated by competition of the UBL domain with ubiquitin in the active site, which helps to dissociate cleaved ubiquitin and increases substrate turnover[86]. USP7 carries a sequence of five UBLs and an active conformation is adopted after binding of UBL domains outside of the USP7 active site. Regulation of such a setup is then achieved by a cofactor binding to the UBLs, dissociating them from their original position and thus changing the activity of USP7[71,87].

3.4. Endoplasmic reticulum-associated degradation (ERAD)

Endoplasmic reticulum (ER) is an organelle where secretory and transmembrane proteins enter their dedicated path. It is estimated that approximately one third of proteins synthetized by an eukaryotic cell reach their destination through ER[88]. Transport of proteins into the ER is performed cotranslationally, through a protein complex called translocon which recognises the ER-targeting sequence on the N-terminus of a nascent polypeptide and transports it into the ER as it is synthetized[89]. This means that folding of a newly synthetized polypeptide takes place inside the ER lumen[88].

Spatial limitations of the ER together with high number of newly synthetized proteins lead to an environment crowded with partially folded proteins with exposed hydrophobic sites, where proper protein folding is a difficult task. To cope with this problem, ER contains a wide pallet of chaperones to assist with proper maturation of proteins[90]. Furthermore, oxidation of disulphide bonds takes place in the ER, and glycosylation sites within proteins are modified with oligosaccharide residues[88].

Even though cells posses' mechanisms to maximise successful protein folding in the ER, incidence of improper maturation or misfolding of proteins is high[91]. Cells deal with this problem with help of proteasomal degradation of unfit ER-resident proteins. This specialised quality-control mechanism is called the Endoplasmic Reticulum-Associated Degradation (ERAD) pathway[90,92]. ERAD pathway faces a considerable obstacle: constituents of UPS needed for degradation are localised in the cytosol, while the ERAD substrates are located inside the ER lumen or anchored to the ER membrane. The ATPase complex Cdc48 in yeast, called p97 or VCP in mammals, must be thus summoned to the ER membrane to provide mechanical energy for transport into the cytosol[93,94]. Therefore, the ERAD pathway consists of three subsequent steps: 1) recognition of a protein within the ER that needs to be degraded, 2) transport of a substrate protein across the ER membrane into the cytosol, a step commonly called retrotranslocation, 3) degradation of a substrate by the UPS in the cytosol[93,95].

3.4.1. Recognition and retrotranslocation of ERAD substrate

ERAD substrates comprise all proteins within the lumen or membrane of the ER, but their recognition differs based on the location of the misfolded lesion: this can be either in the ER lumen, in the transmembrane region, or in the part of protein oriented on the cytosolic side. In yeast, two transmembrane E3 ligases are known to participate in ERAD: Doa10 takes care of proteins with a misfold in the cytosolic part, while Hrd1 manages

degradation of luminal proteins and proteins with misfold in their transmembrane part[96,97]. In mammals, the situation is more complicated, as over ten E3 ligases have been shown to associate with the ER and have influence on the ERAD pathway[98]. However, homologs of both Doa10 and Hrd1 play central role in the mammalian ERAD as well. Although the majority of the ERAD functioning has been investigated in *S. cerevisiae*, ERAD is a conserved pathway and basic principles are applicable also to mammalian cells[99,100].

The best characterized of the ERAD E3 ligases is Hrd1 (HRD1 in mammals). It has seven transmembrane segments and a C-terminal domain possessing E3 ligase catalytic activity. Its ligase activity has long been established as crucial for retrotranslocation of ERAD substrate into the cytosol, and recently it was also shown to be the main component of the transmembrane channel for substrate retrotranslocation[100–103]. In high concentration (e.g. upon overexpression or in *in vitro* liposomal system), Hrd1 forms oligomers and it can promote substrate retrotranslocation on its own, but it loses specificity for misfolded proteins in such circumstances[103]. Under physiological conditions, it associates into a complex with several other proteins that regulate its action and manage substrate recognition. The most important is transmembrane domain protein Hrd3 (SEL1L in mammals). Its luminal domain interacts with both misfolded proteins, luminal chaperones and glycan moieties of luminal glycoproteins and is crucial for substrate recognition[97,102,104]. What characteristics target an ER protein for degradation is a complex matter: in general, an important role play misfolded hydrophobic lesions and state of glycans on the protein surface. After conjugation of a glycan to a protein in the ER, the glycan is progressively trimmed by the ER enzymes. If the protein doesn't reach maturity in time, specific moieties are revealed on the surface of the glycan, which summons specific Hrd3-interacting chaperones that direct protein towards ERAD[95,100,104,105].

Hrd1 further associates with protein called Der1 (Derlin-1, -2 and -3 in mammals) and with Usa1 protein (HERP in mammals) mediating the association. Der1 protein is a multi-spanning membrane protein, related to rhomboid proteases but without catalytic activity. It seems to participate in retrotranslocation channel formation, it interacts with substrate and promotes its entry into the channel[103,106,107]. Other parts of the complex are the E2 ligase Ubc7 (UBE2G1 and UBE2G2 in mammals), which associates with the complex on the cytosolic site, and Ubx2 (UBXD8 in mammals)[95,96]. Ubx2 is a transmembrane protein, possessing an UBX domain on its cytosolic side. This domain interacts with Cdc48/p97 ATPase and summons it to the cytosolic side of the complex[80]. Hrd1 complex is probably able to present part of the substrate to the cytosolic site, where it

is ubiquitinated by Hrd1 E3 ligase domain. Engagement by Cdc48/p97 then provides mechanical force for extraction of the whole substrate from the ER[100,102,108].

3.4.2. Cdc48/p97

The Cdc48/p97 ATPase is a protein performing a huge variety of tasks within the cellular environment. Name Cdc48 refers to its yeast homolog, while mammalian homologs are referred to as p97 or VCP (Valosin Containing Protein). It resides in the cytoplasm and in the nucleus. It is ubiquitously expressed and often constitutes up to 1% of the cytosolic proteome[109]. Mutations in Cdc48/p97 relate to a variety of health disorders and its full inactivation is lethal, showing its importance for cellular functioning[77,110,111]. The Cdc48/p97 uses the ATP hydrolysis to exert mechanical force on protein substrates[112]. This ability of Cdc48/p97 is utilised in numerous types of cell activity: clearance of protein aggregates, unfolding of proteins, membrane fusion events, cytoskeletal organisation or cell division. It is crucial for chromatin organisation, transcription, replication and some DNA damage-repair pathways, because it can extract proteins bound to the DNA and pass them for degradation. Transport of ERAD substrates across the ER membrane is thus one of its many roles[77]. Universality of the Cdc48/p97 functioning is achieved in a modular fashion: Cdc48/p97 acts as a motor unit that can bind a variety of adaptor and regulatory proteins, which in turn direct it to its substrate. Other important aspect for Cdc48/p97 substrate recognition is a ubiquitination of a target protein, which is recognized by the Cdc48/p97-associated adaptors that contain UBDs[77,109,113].

The Cdc48/p97 forms a hexameric barrel structure with a central pore. Each monomer consists of the N-domain located at the N-terminus, then D1 and D2 domains, both possessing an ATPase activity, and the unstructured C-terminal tail. The barrel structure is formed by two stacked rings, the first one formed by the D1 domains and the second one by the D2 domains, with N-domains and C-terminal tails extending in the opposite directions from the central barrel. The C-terminal tails and the N-domains act as binding sites for cofactors that cooperate with Cdc48/p97[114–116].

One of the most important cofactors of Cdc48/p97 are Ufd1 and Npl4 proteins (in yeast, mammalian versions are called UFD1 and NPLOC4) which form a heterodimer (UN complex) that binds to the N-domain of Cdc48/p97 and participates in many processes performed by Cdc48/p97, including ERAD[117–119]. Npl4 carries an UBX domain which binds to N-domain of Cdc48/p97, and both members of the UN complex carry ubiquitin-binding sites. The UN complex serves as a ubiquitin recognition module, binds ubiquitinated substrates of Cdc48/p97 and directs them towards the ATPase

domains[94,108,120,121]. The UN complex has a strong preference for the Lys48-linked ubiquitin chains. This might have an important influence on overall *in vivo* Lys48-chain preference of the proteasome which acts downstream of the Cdc48/p97[122].

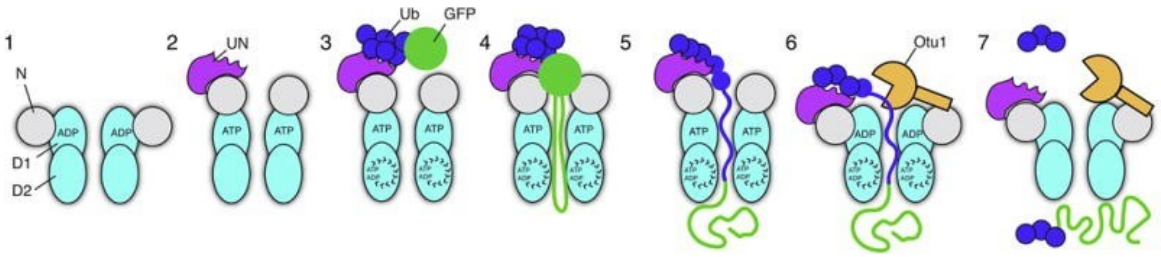

Figure 3: Schematic representation of the stages of a substrate processing by the Cdc48 ATPase. Individual domains of Cdc48 and other cofactors are depicted. GFP is a model substrate. Adopted from publication[108].

The detailed mechanism of function was described for the yeast homolog Cdc48 using an *in vitro* assay of unfolding the ubiquitinated GFP construct and crosslinking experiments. It was shown that the substrate is translocated through the central pore of the Cdc48 in several consecutive steps (see Figure 3)[108]. The N-domains can adopt two conformations, either "up-conformation" or "down-conformation" (with N-domains co-planar with D1 ATPase ring), depending on the conformational changes of D1 ring driven by ATP binding[115,116]. The UN complex binds to the N-domain in both states and without a presence of the substrate it inhibits ATPase activity of Cdc48. Upon substrate binding to the UN, the D1 ring and the N-domains stabilize in the "up-conformation" and the ATPase activity of the D2 ring is enhanced. ATP hydrolysis of the D2 ring then drives the substrate through both the D1 and the D2 pore. During translocation, Lys48-linked ubiquitin chains remain attached to the UN, until the translocation is completed. Next step is ATP hydrolysis by the D1 ring, which brings N-domains to the down-conformation. This exposes ubiquitin chain bound to the UN to the activity of deubiquitinases, namely Otu1 deubiquitinase (YOD1 in mammals)[108]. Otu1/YOD1 resides in the proximity of the Cdc48 complex, bound to the Cdc48 N-domain through its UBX domain[123,124]. When ubiquitin chain bound to the UN is exposed to its action, Otu1 cleaves it which leads to release of the substrate. It was also shown, that short ubiquitin chains of four to ten ubiquitins can be transported through the pore as well and remain conjugated to the substrate[108]. They are probably able to refold after translocation and can be recognized by the Cdc48 downstream factors, even though this needs further experimental verification[108,125].

3.4.3. Degradation

After the substrate is released by the Cdc48/p97 into the cytosol, proteasomal degradation occurs. How exactly the substrate is transferred between the Cdc48/p97 and the proteasome remains unclear. It is likely that degradation follows translocation closely, to prevent presence of unfolded proteins into the cytoplasm as this can lead to aggregation. Solubility of the ERAD substrates is enhanced by BAG6 chaperone complex, which binds to ERAD substrates, prevents their aggregation and improves efficiency of their degradation[126,127]. Ubiquitin chains left on the substrate from retrotranslocation are mostly long enough to trigger proteasomal degradation, but can be further extended by ubiquitin ligase Ufd2 (E4a and E4b in mammals) which is associated with C-terminal tail of Cdc48/p97[30,108,128–130]. Other important enzyme acting on the ERAD substrates is glycanase (Png1 in yeast, NGLY1 in mammals), which interacts with Cdc48/p97 and performs deglycosylation of glycoprotein substrates before they are degraded in the proteasome[131,132].

Yeast proteins Rad23 and Dsk2 were shown in genomic screen to have a positive influence on degradation of ERAD substrates in yeast. Loss-of-function mutation of both Rad23 and Dsk2 proteins led to a delayed degradation of a model ERAD substrate and its accumulation in the cytoplasm where it formed aggregates[133]. This is consistent with established roles of these proteins and their mammalian homologues (RAD23A and B for Rad23, Ubiquilin-1, -2, -3 and -4 for Dsk2) in recognition of proteasomal substrates (more in chapter 3.5.). Double knock-out of Rad23 and Dsk2 leads to a significant decrease in the proteasome bound ubiquitinated substrates[122]. Furthermore, both Dsk2 and Rad23 interact with various factors participating in ERAD. Both Dsk2 and Rad23 interact with Ufd2 ubiquitin ligase, which is bound to Cdc48[81,130,134]. Rad23 and its mammalian homolog RAD23B interact with glycanase Png1/NGLY1[135–137]. Next, mammalian homologs RAD23A and B interact with Ataxin-3, mammalian deubiquitinase that is implicated in regulation of ERAD[138–141]. As Png1/NGLY1, Ufd2 and Ataxin-3 all interact with Cdc48/p97, they probably summon Rad23 and Dsk2 to the proximity of the Cdc48/p97 for substrate capture[130,136,137,140,142].

3.5. UBL-UBA proteins: transient proteasomal receptors

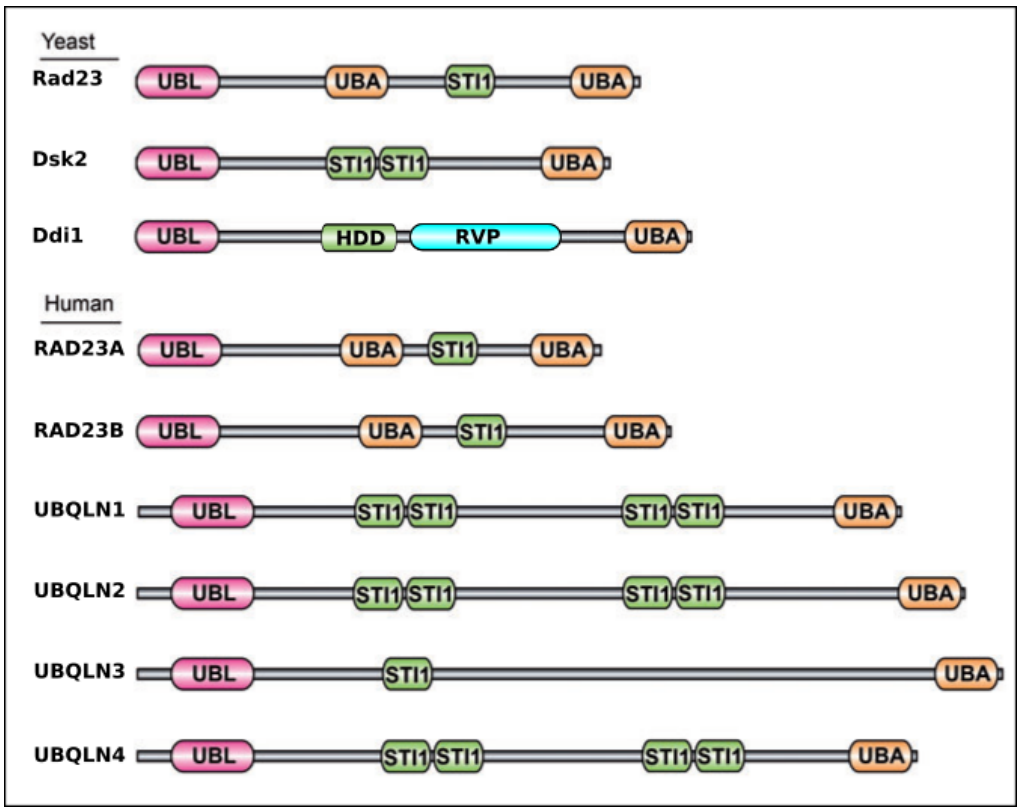

Figure 4: Schematic domain organisations of important UBL-UBA proteins: yeast Rad23, Dsk2 and Ddi1, and two human RAD23 proteins with four human Ubiquilin (UBQLN) proteins. Adopted from publication[143] and edited.

The UBL-UBA protein family is formed by proteins carrying UBL domain on the N-terminal end and one or more UBA domains. Typical members of this family are the yeast proteins Dsk2 and Rad23 (mentioned already in chapter 3.4.3) with their orthologs (including human RAD23B protein) and Ddi1, the yeast ortholog of the human DDI1 and DDI2 proteins.

Other domain often featured within the UBL-UBA proteins is a Sti1-like domain, which name was derived from similarity of these domains to the yeast Sti1 (Stress-Inducible 1) protein[144,145]. These domains usually act as a protein-protein interaction module in context of UBA-UBL proteins. Domain organisations of important UBL-UBA proteins are depicted in Figure 4.

3.5.1. Yeast Rad23

Yeast Rad23 is the first discovered and the best characterized within the UBL-UBA protein family. Its UBL domain is highly similar to ubiquitin and can even be replaced by ubiquitin without affecting Rad23 functions[76]. It interacts with the proteasome via Rpn1

ubiquitin receptor[146]. It features two UBA domains: the first one lays in the middle of the protein (UBA1), the second one is located at the very C-terminus of the protein (UBA2)[14,147]. Both UBA domains of Rad23 were shown to bind Lys48 and Lys63-linked chains, with selectivity towards Lys63-chains being disputed[24,29]. The fourth domain is a Sti1-like domain, located in between the two UBA domains[148]. Rad23 was reported to dimerise through its C-terminal half but without UBA2 domain involved[149].

Rad23 lacks any enzymatic activity, but its many interaction-mediating domains allow it to function as a scaffold and adaptor protein. Rad23 has two main functions: it participates in the process of proteasomal protein degradation and in DNA damage repair[148].

Rad23 was proposed to act as an alternative transiently-associated proteasome receptor, due to the fact that it is able to interact with both the proteasome and the polyubiquitin chains. Overexpression of Rad23 was shown to reverse the growth defects of the cells carrying Δ*rpn10* Δr*ad23* double deletion, and functional UBL domain and both UBA domains are important in the process. Also an increase in polyubiquitinated substrate binding to the proteasome was observed in the presence of Rad23[150]. Further, as discussed in chapter 3.4.3., Rad23 is an essential component of the ERAD pathway: it interacts with Ufd2 through its UBL domain and with Png1 through its Sti1-like domain, binds to polyubiquitinated substrates downstream of Cdc48 and promotes their degradation[122,130,133–136].

The Rad23-bound polyubiquitin chains have been shown to be protected both from elongation and deubiquitination by modifying enzymes *in vitro*. It was also shown to stabilise ubiquitinated substrates, which counters the hypothesis that it simply acts as a receptor transferring proteins into the proteasome[148,150–153]. Rad23 was found to promote proteasomal proteolysis *in vitro* only if present in defined amount[154]. Exact mechanism on how Rad23 promotes degradation of proteasomal substrates thus remains elusive, but it clearly plays a key role in managing the protein degradation[148,154].

Depletion of Rad23 increases the cell sensitivity to the UV light exposure[155]. Rad23 participates in nucleotide excision repair (NER) pathway through interacting with Rad4, DNA-damage sensor that binds to photolesions[156]. Interaction with Rad23 stabilises Rad4 and together they form a complex that initiates assembly of a repair complex on the damaged site[157,158]. Interaction with Rad4 is mediated by the Sti1-like domain and its character is analogous to the interaction of Rad23 with Png1[135,159]. Also the interaction of

Rad23 with ubiquitinated proteins is elevated upon DNA damage[150]. Deletion of Rad23 UBL domain increases sensitivity to UV exposure[160].

3.5.2. Human RAD23A and RAD23B

In humans, the *RAD23* gene was duplicated, resulting in two protein homologs: a 363 amino acid long RAD23A and a 409 amino acid long RAD23B[161], both sharing the domain organisation of the yeast Rad23 (see Figure 4 on page 30). The two proteins were found to be expressed across a wide array of tissues in both human and mouse. The highest levels of mRNA for both proteins were found in testis, with RAD23B being expressed as a shorter protein variant lacking the UBL domain through alternative splicing[162–164]. *RAD23B* gene knock-out in mice was described to cause high embryonal lethality and severe developmental defects in surviving mice, while *RAD23A* gene knock-out caused no significant effect on mice development. However, double knock-out of both *RAD23* genes increased embryonal lethality to 100%, suggesting that *RAD23A* can partly compensate for the loss of *RAD23B*[165,166]. This is in agreement with the fact that RAD23B protein is usually present in markedly higher amount than RAD23A protein in mammalian cells[162,167].

Both human homologs retained the role in DNA damage response: they bind the XPC protein (homolog of the yeast Rad4) and stimulate NER activity. The interaction is mediated by the Sti1-like domain, which is referred to as XPCb (XPC-binding) domain for the human homologs[168–170].

Even though the function of RAD23 proteins in proteasomal degradation was not explored in such a detail as for the yeast homolog, important similarities exist. They retained both the ability to bind the Lys-48 linked ubiquitin chains and the ability to bind proteasomal ubiquitin receptors using the UBL domain[162,171]. Both have been shown to interact with the S5a proteasomal subunit (homolog of yeast Rpn10), RAD23B was further described to interact with PSMD2 (homolog of the yeast Rpn1)[171,172]. The RAD23A UBL domain was successfully used as affinity chromatography matrix for purification of the human proteasomes, the RAD23B UBL domain as an affinity chromatography matrix allowed purification of both the proteasomes and the p97/VCP ATPase complexes from human cells[173,174]. Further, the human RAD23 proteins interact with Ataxin-3 deubiquitinase (via their UBL domain) and NGLY1 glycanase (via their XPCb domain), enzymes participating in ERAD and interacting with p97/VCP ATPase[131,137,138,141]. Both human RAD23 proteins were reported to form intramolecular interaction between the UBL

and UBA domains when not bound to the proteasome or the polyubiquitin, with RAD23A specifically described to not form a dimer.[175,176].

3.5.3. Yeast Dsk2 and human Ubiquilin proteins

Dsk2 is another yeast protein containing the UBL domain on its N-terminus and the UBA domain on its C-terminus (see Figure 4 on page 30). The UBL domain of Dsk2 interacts with Rpn1 and Rpn10 subunit of the proteasome, and its UBA domain interacts with both Lys48 and Lys63-linked chains without specific preference[177–180]. Binding of Dsk2 to the proteasome is enhanced when Rpn10 is depleted, because the population of Rpn10 that resides freely in the cytoplasm restricts access of Dsk2 to the proteasome[179]. In the central region of the Dsk2, two Sti1-like domains are located, that were indicated in the binding to the Hsp70-like chaperone family [143,181].

Genetic studies in yeast revealed a role of Dsk2 in ERAD (mentioned in chapter 3.4.3) and in duplication of microtubule organising centre (MTOC), a processes in which function of Rad23 was implicated as well[133,182]. Similarly to Rad23, Dsk2 protects bound polyubiquitin chains from deubiquitination, thus stabilising them, with overexpression of Dsk2 causing polyubiquitin accumulation and cytotoxicity[178,179]. Dsk2 was described to form a homodimer via the UBA domains and to form a heterodimer with Rad23, with the UBA of each protein binding to UBL domain of the other[183,184].

In mammals, the original *DSK2* gene has been multiplied: a human genome carries four genes encoding "Dsk2" proteins called Ubiquilin1-4 (UBQLN1-4). UBQLN1 and UBQLN2 proteins are often marked alternatively as PLIC-1 and PLIC-2 proteins (Protein Linking Integrin-associated protein with Cytoskeleton). While UBQLN1 is expressed in all tissues and UBQLN2 and 4 in most tissues, UBQLN3 is expressed only in testis and is the most evolutionary diverged[185,186]. All of them contain the UBL and the UBA domain and the Sti1-like domains in the middle: UBQLN3 carries one, while other UBQLN proteins carry four Sti1-like domains (shown in Figure 4 on page 30)[143,185].

UBQLN1, UBQLN2 and UBQLN4 were described to participate in degradation of proteasomal substrates[187,188]. The UBA domain of UBQLN1 binds Lys48 and Lys63-linked ubiquitin chains and even monomeric ubiquitin[27]. The UBQLN1/2 UBL domain binds to ARDM1 proteasomal receptor through its pleckstrin-like domain, S5a proteasomal receptor and Ataxin-3 through their UIM motifs[172,189,190]. The UBQLN1 is also responsible for summoning p97/VCP complex to ERAD substrates[191]. Together with UBQLN2 and UBQLN4 it was shown to bind HRD1 cofactor protein HERP and promote ERAD[192].

UBQLN2 was described to interact with RAD23A protein, with the UBA of each protein binding to UBL domain of the other[193].

However, in contrast to human RAD23 proteins which seem to function specifically within proteasomal degradation, UBQLN functions exceed into autophagy pathways[143,185]. UBQLN1, UBQLN2 and UBQLN4 were shown to prevent negative effects of nutrient depletion by promoting autophagy[194]. UBQLN4 helps summon UBQLN1 to autophagic complexes[195]. UBQLN1 and UBQLN2 associate with protein aggregates and aggresomes, clusters of aggregated proteins that cells gather in proximity of MTOC to contain their toxic effects. They promote aggresome maturation and target proteins to lysosomal degradation via autophagy[190,194,196]. UBQLN1 interacts via its UBL domain with UIM motif of autophagy factor EPS15 and depletion of UBQLN1 reduces aggresome formation. Deletion of its UBL domain causes negative dominant phenotype, where formation of aggresomes is blocked and association of EPS15 with dispersed protein aggregates is prevented[190]. In line with this fact, Dsk2 promotes a vacuole-mediated clearance of protein aggregates in yeast by inclusion body formation[197].

This wide functioning of the UBQLN proteins contributes to the recent view of the individual cellular proteolytic pathways being closely tied to each other. Upon impairment of proteasomal function and aggregate creation, proteotoxic stress can be ameliorated by redirecting towards autophagic pathways[143,185]. On the other hand, the proteasomes participate in aggresome clearance, a process which UBQLN2 was shown to mediate together with the heat-shock chaperones[198].

3.5.4. Yeast Ddi1

The yeast protein Ddi1 (DNA damage-inducible protein 1) is the third member of the UBL-UBA family in yeast and carries the UBL domain on its N-terminus and the UBA domain on its C-terminus (see Figure 4 on page 30)[199]. However, Ddi1 is atypical among the UBL-UBA family.

Firstly, Ddi1 and its homologs carry a conserved retroviral protease (RVP) domain homologous to the proteases of retroviruses including HIV[199]. The RVP structures show that the RVP domain adopts a typical dimeric aspartate protease fold with Asp220 from both subunits forming a catalytic site[200,201]. Ddi1 exists as a dimer due to the RVP domain[201]. While no *in vitro* proteolytic activity of Ddi1 was reported so far, Ddi1 from *Leishmania major* was reported to be active at acidic pH. However, the biological data

from yeast indicate a high importance of Asp220 for Ddi1 function as also discussed further[202,203].

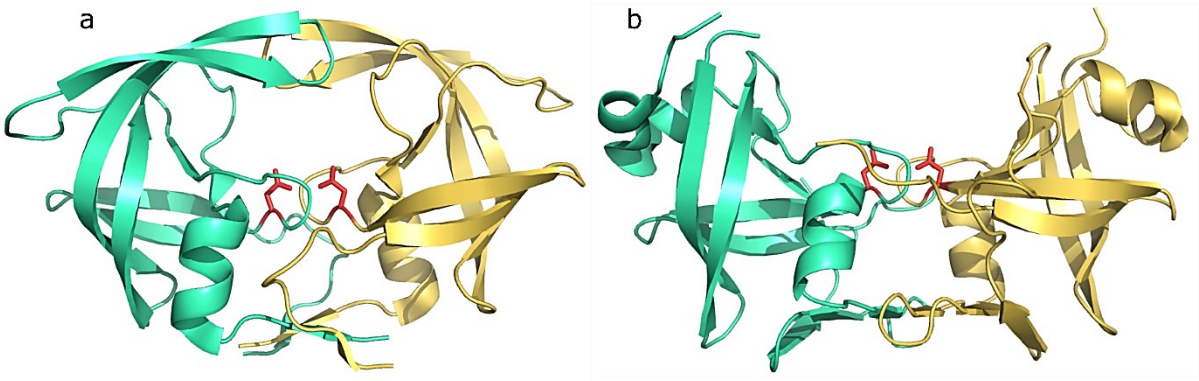

Figure 5: Comparison of the structure of the HIV protease and the RVP domain of the yeast Ddi1. The chains of each dimer are coloured differently. Two aspartates in the catalytic centre in red. **a)** Structure of the HIV protease (PDB code 4LL3)[204] **b)** Structure of the yeast Ddi1 RVP domain (PDB code 2I1A)[200]. Picture was created using PyMol software (Schrodinger, LLC).

Based on its sequence similarity with Sti1-like domains, the helical domain was identified within Ddi1 adjacent to the N-terminus of the RVP domain. Its structure was however found to differ from the Sti1-like domains of other UBL-UBA proteins and it was named the Helical domain of Ddi1 (HDD), due to its conservation among Ddi1 homologs. Structure of the yeast Ddi1 HDD domain revealed that it's composed of two independent helical bundles, the N-terminal one consisting of four and the C-terminal one consisting of two helices[201]. Structure of the Ddi1 HDD domain is shown in Figure 6 on the next page.

The N-terminal UBL of Ddi1 adopts the β-grasp fold (see Figure 6 on the next page) but shares only 14% sequence identity with ubiquitin[201,205]. It doesn't bind the proteasome subunit Rpn10 and its interaction with Rpn1 is much weaker than for Rad23 or Dsk2[206–208]. On the other hand, the Ddi1 UBL domain is able to bind ubiquitin, unexpected behaviour among UBL domains[201,205]. The K_d of its interaction with monoubiquitin was determined as 45 μM, which is comparable to UBA domains binding affinities to monoubiquitin (4-500 μM). In comparison, isolated UBA domain of Ddi1 was reported to bind monoubiquitin with the K_d of 150 μM. The atypical behaviour of Ddi1 UBL domain questions its classification as UBL domain[205] .

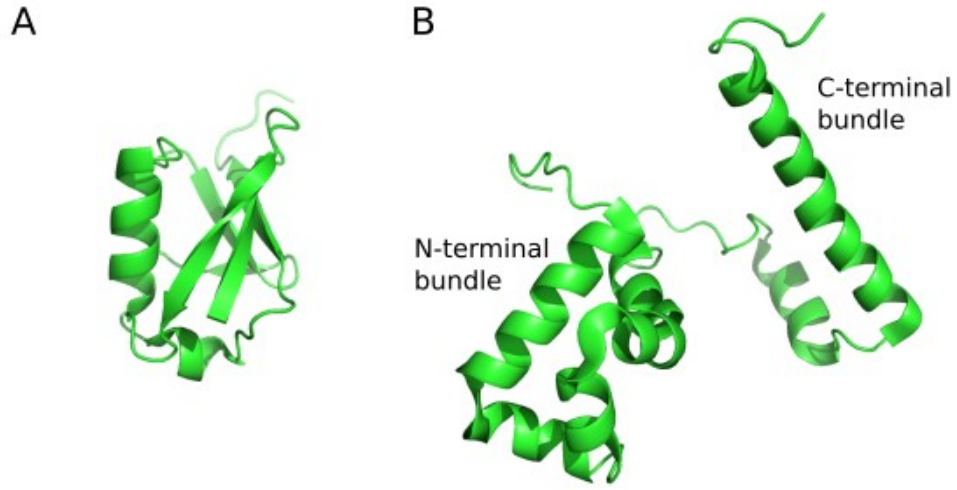

Figure 6: **A)** Structure of the yeast Ddi1 UBL domain (PDB code 2N7E)[201]. **B)** Structure of yeast Ddi1 HDD domain (PDB code 5KES)[201]. Two distinct helical bundles are indicated. The picture was created using PyMol software (Schrodinger, LLC).

With two domains able to bind ubiquitin, the full-length Ddi1 was reported to bind to monoubiquitin with the K_d of 320 μM with 1:2 stoichiometry, and to Lys48-linked diubiquitin with the K_d of 77 μM and 1:1 stoichiometry. Further, ΔUBL Ddi1 binds monoubiquitin with the K_d of 43 μM (1:1 stoichiometry) and ΔUBA Ddi1 binds monoubiquitin with the K_d of 310 μM (1:1 stoichiometry). This suggest that the UBL and UBA domains cooperate for selectivity towards ubiquitin chains instead of monoubiquitin[201].

DDI1 gene shares bidirectional promoter with *MAG1*, glycosidase involved in base-excision repair, and its transcription is induced upon DNA damage[209]. It was found that Ddi1 has overlapping functions with Wss1 protein, a metalloprotease which is involved in DNA damage-repair by removing DNA-protein crosslinks. Deletion of both *WSS1* and *DDI1* genes causes severe sensitivity to hydroxyurea (HU), which acts as a replication stalling agent. Overexpression of Ddi1 is able to alleviate the HU sensitivity, but if the catalytic Asp220 is mutated or the C-terminal helix bundle of Ddi1 HDD domain is deleted, this ability is lost. Truncated variant of Ddi1 containing only C-terminal HDD bundle and RVP domain is still able to protect the yeast cells from HU effect. Proteolytic activity of Ddi1 and the C-terminal bundle of HDD domain are thus the key components for Ddi1 function in DNA repair[203]. The data are also supported by a recent preprint by Serbyn and colleagues identifying Ddi1 as a protease that contributes to DNA-protein crosslinks removal[210].

It was found that Ddi1 is necessary for degradation of Ho endonuclease[211]. This nuclear protein is phosphorylated upon activation of certain DNA-damage response pathways and exported into cytoplasm, where it is ubiquitinated and degraded in proteasome[212]. Ddi1 interacts via its UBL domain with Ufo1, protein which recruits Ho endonuclease to ubiquitination complex[213]. It was found that Ho endonuclease accumulates in *Δddi1* cells, but not in *Δrad23* or *Δdsk2* cells[211].

Ddi1 together with Rad23 was shown to supress temperature and HU sensitivity of the cells carrying mutations in Pds1, a mitotic checkpoint control protein. For both Rad23 and Ddi1, the C-terminal UBA domains are important to fulfil this role, for Ddi1 the catalytic Asp220 is necessary as well[214,215]. Ddi1 was reported to form heterodimer with Rad23, with the UBA of each protein binding to UBL domain of the other[149].

Ddi1 interacts with the Cdc48 ATPase. Together, they have been described to transport the membrane hydrolase Cps1 into a vacuole for degradation. Ddi1 binds to Cps1 after it is ubiquitinated. Cdc48 and Ddi1 were also shown to disassemble insoluble Cps1 oligomers. Rad23 and especially Dsk2 have overlapping roles with Ddi1 in this process[216].

Ddi1 was found to be involved in cellular trafficking and was independently named Vsm1 (v-Snare master 1) in this context. It is a negative regulator of exocytosis, it interacts with both exocytic Snc1 and endocytic Snc2 v-Snare proteins, and with t-Snare protein Sso1[217,218]. The Sso1 binding is mediated by the linker region between RVP and UBA domain of Ddi1[215]. *Δddi1* cells display elevated levels of protein secretion. Both UBL domain and catalytic Asp220 are important for negative exocytosis regulation[217,219].

3.6. Human DDI2 and its function

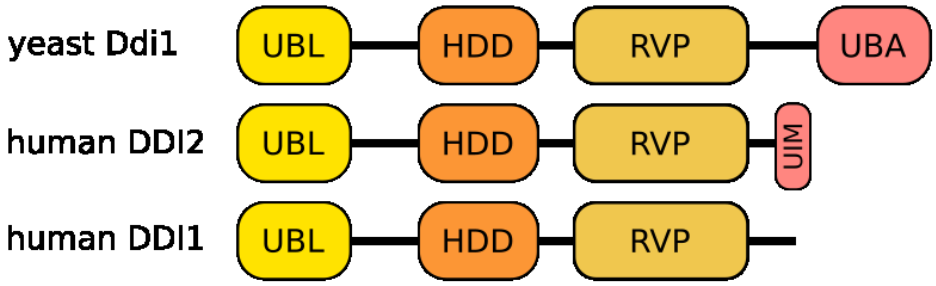

Figure 7: Schematic depiction of domain organisations of the yeast Ddi1 and human DDI2 and DDI1 proteins. Human homologs lost the UBA domain during evolution. DDI2 carries UIM motif on its C-terminus, which is not present in DDI1.

Human homologs of Ddi1, DDI1 and DDI2, carry the UBL domain on their N-terminus, the HDD domain and the RVP domain similarly to the yeast Ddi1, however they lost the UBA domain during evolution. Instead, the *RSC1A1* gene was inserted into the locus, resulting in the UBA domain becoming a part of the RSC1A1, protein trafficking regulator. In mammals, the original *DDI2* gene was further duplicated through a retrotransposition event. This resulted in one original gene (*DDI2* in humans) and one new gene consisting of only one exon (*DDI1* in humans), encoding two proteins with 70% sequence identity and 81% similarity[1]. For comparison of domain organisation of the yeast Ddi1 and human DDI1 and DDI2, see Figure 7 (page 37).

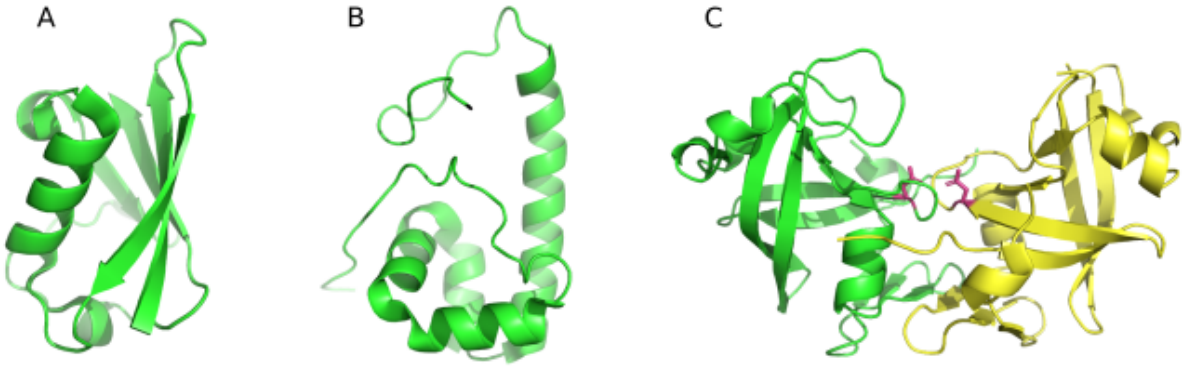

Figure 8: Structure of individual domains of the human DDI2 protein. **A)** Structure of DDI2 UBL domain (PDB code 2N7D)[1]. **B)** Structure of DDI2 HDD domain (PDB code 5K57)[1] **C)** Structure of DDI2 RVP domain (PDB code 4RGH)[1]. Two chains forming the dimer are coloured differently. The catalytic aspartates are depicted in red. The picture was created using PyMol software (Schrodinger, LLC).

DDI2 is the better studied protein among the two human homologs. Structures of its UBL, HDD and RVP domains have been already solved and are presented in Figure 8. As a potential replacement for the missing UBA domain, a single UIM motif was discovered through bioinformatic analysis at the C-terminus of DDI2. However, when the capability of DDI2 to bind ubiquitin was examined, DDI2 was shown to be unable to bind Lys48 and Lys63-linked diubiquitin chains in affinity chromatography assay. The interaction of the DDI2 UIM motif with ubiquitin was determined by NMR with the K_d of 2.2-3 mM, specific, yet quite weak and most likely non-physiological[1]. The interaction of human DDI2 UBL with ubiquitin is in the range of 0.42-1.1 mM, which is significantly weaker compared to the yeast Ddi1-UBQ binding ($K_d \sim 45\mu M$), but already in the range of values typical for the UPS system[1]. Furthermore, taking into account that DDI2 is a dimer, it is possible that DDI2 binds with higher affinity the longer ubiquitin chains through avidity effect, but this possibility hasn't been addressed experimentally so far.

Proteomic techniques revealed that DDI2 co-purifies with a range of proteins forming the 19S proteasomal subunit, most prominently PSMD2 (homolog of the yeast Rpn1), the proteasome ubiquitin receptor that interacts with UBL domains[2]. This has not been further explored and the complete data from the experiment are not present in the publication. Thus, it is not clear if it is a direct interaction, or whether this could be mediated by another binding partner.

3.6.1. DDI1 and DDI2 in replication stress

Other set of proteins co-purified with DDI2 were various replication factors, including MCM helicase, DNA polymerase delta and others. This led to further discovery that both DDI1 and DDI2 have mutually redundant roles in response to the replication stress. Depletion of DDI1 and DDI2 caused sensitivity of cells to replication stalling agents such as HU. This was found to be due to inability of DDI1/2 depleted cells to remove RTF2, replication stalling factor, from the replication fork and restart replication. It was suggested that DDI1/2 targets RTF2 for proteasomal degradation[2].

3.6.2. DDI2 activates TCF11/NRF1 transcription factor

Human DDI2 protein was identified as essential for activation of TCF11/NRF1 transcription factor. Modification of TCF11/NRF1 into a transcriptional activator requires its proteolytic cleavage, but this step is blocked in DDI2 depleted cells, offering strong evidence for the importance of DDI2 proteolytic activity[3].

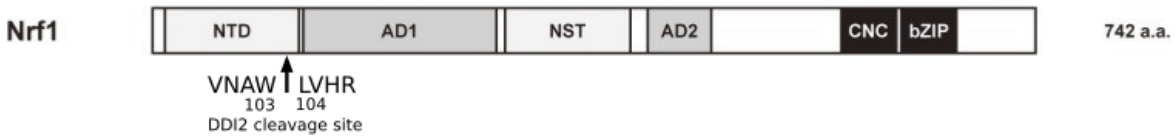

Figure 9: Schematic domain organisation of TCF11/NRF1 protein. DDI2 cleavage site between Trp103 and Leu104 is marked. NTD – N-terminal domain with transmembrane helix; AD - acidic domain; NST – Asn/Ser/Thr-rich domain; CNC - cap 'n' collar domain; bZIP - basic-region leucine zipper domain. Adopted from publication[31] and edited.

TCF11/NRF1 was demonstrated to manage proteasome synthesis[220–222]. It belongs to the cap 'n' collar/basic-region leucine zipper (CNC-bZIP) family of transcription factors together with NRF2 and NRF3. While NFR2 was shown to regulate response to oxidative stress, TCF11/NRF1 manages cellular response to proteotoxic stress. CNC-bZIP factors bind via its bZIP domain to antioxidant response element (ARE) sequences present in promoter regions of genes crucial to stress response and activate their transcription[31]. ARE

sequence 5‘-(A/G)TGACTCAGC-3‘ was identified as a binding site of TCF11/NRF1 and is among others present within the promoters of all 33 known proteasome subunits[223]. The bZIP domain of TCF11/NRF1 is located on its C-terminus, flanked by the CNC domain. On its N-terminus, TCF11/NRF1 carries a transmembrane helix. Two acidic domains (AD) and Asn/Ser/Thr-rich (NST) domain are located in the middle region[31] (see Figure 9, page 39).

TCF11/NRF1 activation pathway is set up as feedback mechanism monitoring the level of proteasomal activity[31,224]. This mechanism is schematically shown in Figure 10. TCF11/NRF1 is translocated into the ER lumen upon synthesis and anchored to the ER membrane through its N-terminal transmembrane helix. In the ER lumen it is glycosylated on the NST domain. From the ER lumen, it is constitutively targeted to the ERAD pathway. Its recognised by the HRD1 complex, translocated into the cytoplasm by the action of p97/VCP ATPase, deglycosylated by NGLY1 and degraded in the proteasome[31,224]. Inactivation of NGLY1 results in TCF11/NRF1 failing to translocate into the nucleus[225].

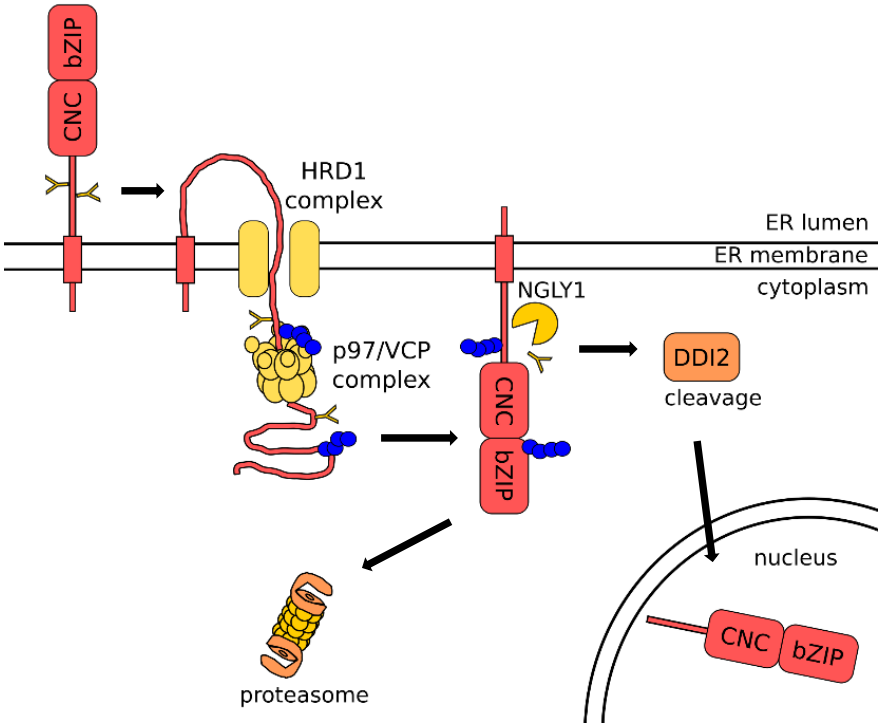

Figure 10: Schematic representation of TCF11/NRF1 signalling pathway. TCF11/NRF1 is retrotranslocated via HRD1/p97 pathway and deglycosylated by NGLY1. Then it is either degraded in the proteasome or cleaved by DDI2 and translocated to the nucleus. Degradation is much more prevalent, but when the proteasome activity decrease, more TCF11/NRF1 escapes degradation and gets activated. Pink - TCF11/NRF1; blue - ubiquitination; yellow forks - glycosylation.

When proteasomal activity is impaired, either due to saturation of proteasomes by other substrates or by action of proteasome inhibitor drugs, TCF11/NRF1 quickly accumulates in cells. Undegraded TCF11/NRF1 dwells on the cytoplasmic side of the ER membrane, where it is proteolytically processed by DDI2. Deglycosylated TCF11/NRF1 with trimmed off N-terminal transmembrane domain is then transported into the nucleus where it activates proteasome synthesis[224]. While TCF11/NRF1 presence is almost undetectable under normal conditions in cells due to constant degradation, some TCF11/NRF1 molecules probably escape degradation even when proteasome is not impaired, as basal proteasome expression is managed by TCF11/NRF1 as well[3,220,221].

TCF11/NRF1 is cleaved between position Trp103 and Leu104, with surrounding sequence conforming to the cleavage motif of retroviral aspartyl proteases[224,226]. Depletion of DDI2 decreases the basal proteasome expression and blocks the cleavage of TCF11/NRF1 in cells treated with proteasome inhibitors. Expression of DDI2 protein restores the cleavage, but not expression of DDI2 with mutated catalytic Asp252 in the RVP domain. Expression of DDI2 with deletion of UBL domain restored the cleavage only partially[3].

NRF3, related transcription factor regulating cell proliferation, is cleaved by DDI2 as well upon proteasome inhibition. It is similar in structure to TCF11/NRF1 and the cleavage site is conserved in sequence[227]. Its role in response to proteasome inhibition is unclear.

3.6.3. Mechanism of TCF11/NRF1 cleavage by DDI2

The exact mechanism of TCF11/NRF1 cleavage by DDI2 is not known, as *in vitro* cleavage of TCF11/NRF1 by DDI2 could not be reconstituted with recombinant proteins and other factors are apparently necessary for the cleavage to occur[3]. *In vivo*, TCF11/NRF1 is most likely both ubiquitinated and defolded after p97/VCP translocation[108], which may have influence on DDI2 ability to cleave it. RAD23B is known to interact with ERAD substrates after retrotranslocation, it interacts with several p97/VCP-associated proteins including NGLY1 (as described in chapter 3.4.3, page 29), and was found to interact with DDI2 in another thesis[4]. RAD23B is thus very likely to be involved in the TCF11/NRF1 pathway and might coordinate the processing of TCF11/NRF1 after it is translocated to cytoplasm by p97/VCP.

4 Methods

4.1. Material and instruments

4.1.1. Chemicals

- **Bio-Rad, Hercules (USA)**
  agarose
- **Biosynth AG, Staad (Switzerland)**
  Isopropyl-β-D-1-thiogalactopyranoside (IPTG)
- **Biotika, Slovenská Ľupča (SR)**
  ampicillin
- **Biotium, Hayward (USA)**
  GelRed
- **Cambridge Isotope Laboratories, Andover (USA)**
  Ammonium chloride (^{15}N-labelled), D-glukose (^{13}C-labelled)
- **Lach-Ner, Neratovice (CR)** acetic acid, ethanol, formaldehyde, hydrochloric acid, isopropanol, potassium acetate, potassium dihydrogen phosphate, sodium dihydrogen phosphate
- **Merck, Billerica (USA)**
  Deuterium oxide
- **New England BioLabs, Ipswitch (USA)**
  Phusion® DNA polymerase, 5x Phusion HF buffer, NheI, EcoRI, NdeI, KpnI, SacI, 10x NEBuffer 1.1, 10x NEBuffer 2.1, 10x NEBuffer 3.1, 10x EcoRI buffer, 10x CutSmart buffer, deoxynucleotide mix (dNTPs), Antarctic phosphatase, Antarctic phosphatase reaction buffer,
- **Penta, Prague (CR)**
  methanol, sodium hydroxide, magnesium sulphate heptahydrate, sodium sulphate, boric acid, phosphoric acid, ammonium sulphate
- **Roche Diagnostics GmbH, Basel (Switzerland)**
  inhibitor cocktail Complete Mini (EDTA-free)
- **Serva, Heidelberg (Germany)**
  Bromophenol blue, Coomassie Brilliant Blue G-250

- **Sigma-Aldrich, Buchs (Switzerland)**

  ethylenediaminetetraacetic acid (EDTA), sodium dodecylsulphate (SDS), LB medium, LB agar, 2-mercaptoethanol, glycerol, HEPES, imidazole, saccharose, sodium azide, N,N'-methylenbisakrylamide, calcium chloride, thiamine, D-biotin, BME Vitamin 100x solution, ferrous chloride tetrahydrate, manganese(II) chloride tetrahydrate, cobalt(II) chloride hexahydrate, zinc chloride, copper(II) chloride dihydrate, calcium chloride dihydrate, sodium molybdate dihydrate, sodium chloride, desthiobiotin, bovine serum albumin (BSA)
- **Thermo Scientific (USA)**

  5% (v/v) Blocker Casein in PBS solution
- **USB, Cleveland (USA)**

  acrylamide, glycine, TRIS

4.1.2. Instruments

- GeneAmp PCR System 2400 thermocycler instrument (Perkin-Elmer)
- Incubator IPP 400 (Memmert GmbH)
- Incubator Innova® 44/44 R (New Brunswick Scientific)
- Horizontal electrophoresis apparatus (Gibco)
- Vertical electrophoresis apparatus (Bio-Rad)
- Western blotting apparatus (Bio-Rad)
- Near-Infrared Fluorescence Imaging System Odyssey CLx (LI-COR)
- Camera Quantum ST4 (Vilber Lourmat)
- Homogeniser EmulsiFlex-C3 (Avestin)
- Centrifuge Sorvall Evolution RC (Thermo Scientific)
- Spectrophotometer UNICAM UV 500 (Thermo Scientific)
- Spectrophotometer NanoDrop™ One (Thermo Scientific)
- FPLC system ÄKTAExplorer (Amersham Pharmacia Biotech)
- FPLC column HiLoad Superdex 75pg 16/60 (Amersham Pharmacia Biotech-GE Healthcare Bio-Sciences)
- FPLC column HiLoad Superdex 200pg 16/60 (Amersham Pharmacia Biotech-GE Healthcare Bio-Sciences)
- Thermoblock Thermocell MB102 (BIOER Technology)
- Spectrometer Bruker Avance 850 Hz (Bruker BioSpin GmbH)

4.1.3. Primers

The primers used in the thesis were supplied by Sigma Aldrich and are listed in Table 1. The sequences complementary to the PCR template are written in bold, overhang sequences are written in italic. For the complementary section of each primer, the melting temperature was predicted using Exiqon Tm Prediction on-line tool (https://www.exiqon.com/ls/Pages/ExiqonTMPredictionTool.aspx) and the temperatures are listed in Table 1.

Table 1: Primer names, sequences and melting temperatures

Primer name	Primer sequence	T_m
F_NStrep_Ddi2	*AATCCAATGCGGCTAGC***ATGCTGCTCACCG**	56 °C
R_NStrep_Ddi2	*CGACGGAGCTCCTACTA***TGGCTTCTGACGC**	56 °C
F_CStrep_Ddi2	*AAGAAGGAGATATACAT***ATGCTGCTCACCG**	56 °C
R_CStrep_Ddi2	*TACAGGTTCTCGGTACC***TGGCTTCTGACGC**	56 °C
F_NStrep_RVP	*AATCCAATGCGGCTAGC***CAGAACATTGAGGAAAAC**	54 °C
R_NStrep_Ddi2ΔUIM	*CGACGGAGCTCCTACTA***CTCTGGTAGCTCTCC**	55 °C
R_CStrep_UBL	*TACAGGTTCTCGGTACC***GTCTGCATTCTCCTTCTG**	58 °C
F_CStrep_Ddi2ΔUBL	*AAGAAGGAGATATACAT***ATGCCTCGACCTCCA**	52 °C
F_Nstrep_HDD	*AATCCAATGCGGCTAGC***ATGCAGCAGTCCCACTCA**	58 °C
R_Nstrep_HDD	*CGACGGAGCTCCTACTA***TTGCCTTATATCTTCTTCTA**	52 °C
R1_Cstrep_ΔHDD	**CAATGTTCTGTGTTCCTGGTGG**	64 °C
F2_Cstrep_ΔHDD	**CAGGAACACAGAACATTGAGGAAAAC**	65 °C
F_RAD23_FL	*ATCAAGCTAGC***ATGCAGGTCACCC**	54 °C
R_RAD23_FL	*AGTAAGAATTCCTATCA***ATCTTCATCAAAGTTC**	55 °C
R_RAD23_UBL	*AGTAAGAATTCCTATCA***TGTGGACACTGC**	51 °C
F_RAD23_UBA1	*ATCAAGCTAGC***GTGACGGGTCAG**	52 °C
R_RAD23_UBA1	*AGTAAGAATTCCTATCA***TCTATCTCCAGGG**	47 °C
F_RAD23_UBA2	*ATCAAGCTAGC***CAAGTAACACCTCAGG**	54 °C
F_RAD23_ΔUBL	*ATCAAGCTAGC***CCAGCACCAGCTACAACTC**	64 °C
R_RAD23_ΔUBA2	*AGTAAGAATTCCTATCA***AATGTAGTTCATATGACC**	51 °C
T7_F	**TAATACGACTCACTATAGGG**	55 °C
T7_R	**GCTAGTTATTGCTCAGCGG**	62 °C

4.1.4. Vectors

- p905 vector (Grantz Saskova laboratory, FUS)
- Strep-p905 vector (Grantz Saskova laboratory, FUS)
- pTRE-Tight vector carrying N-Flag-RAD23B gene (Grantz Saskova laboratory, FUS)
- pTRE-Tight vector carrying N-HA-Ddi2 gene (Grantz Saskova laboratory, FUS)

4.1.5. Other material

- **IBA Lifesciences (Germany)**

  Strep-Tactin® Sepharose resin

- **QIAGEN (USA)**

  NiNTA Agarose resin, GelPilot 1kb Ladder, GelPilot 100bp Plus Ladder, QIAquick Gel Extraction Kit

- **ZYMO Research (USA)**

  DNA Clean & Concentrator kit, Zyppy Plasmid Miniprep kit

- **Spectrum Laboratories (USA)**

  Dialysis membrane Spectrapore 6-8kDa MWCO, Dialysis membrane Spectrapore 3.5kDa MWCO

- **Novagen (USA)**

  E. coli strain BL21(DE3)RIL, *E. coli* strain TOP10

- **Millipore (USA)**

  15ml Amicon Ultra centrifugal filters 3000 MWCO, 15ml Amicon Ultra centrifugal filters 10000 MWCO, 2ml Amicon Ultra centrifugal filters 3000 MWCO, 2ml Amicon Ultra centrifugal filters 10000 MWCO

4.1.6. Software

- UGENE[228]
- Sparky (UCSF) with NMRFAM extension[229] (https://nmrfam.wisc.edu/nmrfam-sparky-distribution)
- TITAN[230] (www.nmr-titan.com)
- Exiqon TM Prediction Tool (QIAGEN) (https://www.exiqon.com/ls/Pages/ExiqonTMPredictionTool.aspx)
- Expasy Protparam[231] (https://web.expasy.org/protparam)
- The PyMOL Molecular Graphics system 2.0 (Schrodinger, LLC)

4.2. Cloning

4.2.1. Polymerase chain reaction

The polymerase chain reaction (PCR) was used to amplify specific DNA sequences using as a DNA template either the pTRE-Tight plasmid encoding DDI2 protein or the pTRE-Tight plasmid encoding RAD23B protein. The primers (see Table 1 in Materials

chapter 4.1.3., page 44) were designed to produce specific extensions at the ends of the amplified DNA to allow the ligation into the expression vector. The composition of the PCR reaction mixture is listed in Table 2. PCR programme used for the amplification reaction is listed in Table 3. PCR was performed in GeneAmp PCR System 2400 thermocycler instrument (Perkin-Elmer).

Table 2: Composition of PCR reaction mixture

Component	Volume (µl)
DNA template (10 ng/µl)	1
5x HF buffer	6
Primer F (10µM)	1
Primer R (10µM)	1
dNTPs (10mM)	1
Deionised water	19.8
Phusion polymerase	0.2

Table 3: PCR program

Step	Temperature	Duration	Cycles
1	98 °C	30 s	-
2	98 °C	10 s	-
3	50 °C	20 s	-
4	72 °C	45 s	30x to Step 2
5	72 °C	5 min	-

The primers were designed manually using UGENE software[228] and their melting temperatures of the sections complementary to the template DNA were calculated using online Exiqon TM Prediction Tool (QIAGEN). Individual PCR products, their expected size and primer combination for their amplification are listed in Table 4.

Table 4: PCR products, primer combinations and expected sizes

Construct	Forward primer	Reverse primer	Expected size
C-Strep DDI2 FL	F_Cstrep_Ddi2	R_Cstrep_Ddi2	1231 bp
N-Strep DDI2 FL	F_NStrep_Ddi2	R_NStrep_Ddi2	1231 bp
$DDI2^{82-399}$ ΔUBL	F_Cstrep_Ddi2ΔUBL	R_Cstrep_Ddi2	991 bp
$DDI2^{1-81}$ UBL	F_Cstrep_Ddi2	R_Cstrep_Ddi2UBL	277 bp
$DDI2^{\Delta116-212}$ ΔHDD1	F_Cstrep_Ddi2	R_Cstrep_Ddi2ΔHDD	372 bp
$DDI2^{\Delta116-212}$ ΔHDD2	F2_Cstrep_Ddi2ΔHDD	R_Cstrep_Ddi2	586 bp
$DDI2^{116-212}$ HDD	F_NStrep_Ddi2HDD	R_NStrep_Ddi2HDD	328 bp
$DDI2^{212-399}$ RVP-UIM	F_NStrep_Ddi2RVP	R_NStrep_Ddi2	595 bp
$DDI2^{1-360}$ ΔUIM	F_NStrep_Ddi2	R_NStrep_Ddi2ΔUIM	1114 bp

DDI2$^{1-212}$ UBL-HDD	F_NStrep_Ddi2	R_NStrep_Ddi2HDD	670 bp
RAD23B FL	F_Rad23b_FL	R_Rad23b_FL	1255 bp
RAD23B$^{1-82}$ UBL	F_Rad23b_FL	R_Rad23b_UBL	274 bp
RAD23B$^{185-233}$ UBA1	F_Rad23b_UBA1	R_Rad23b_UBA1	175 bp
RAD23B$^{360-409}$ UBA2	F_Rad23b_UBA2	R_Rad23b_UBA2	175 bp
RAD23B$^{83-409}$ ΔUBL	F_Rad23b_ΔUBL	R_Rad23b_FL	1009 bp
RAD23B$^{1-360}$ ΔUBA2	F_Rad23b_FL	R_Rad23b_ΔUBA2	1108 bp

4.2.2. Restriction cleavage of DNA

Cleavage of the DNA was performed using commercially supplied restriction enzymes (NEB). For cloning of the RAD23B constructs, both insert sequences and the p905 vector were cleaved using NheI and EcoRI endonucleases. For cloning of the DDI2 constructs with N-terminal tag, combination of NheI and SacI restriction endonucleases was used to cleave the Strep-p905 vector. For cloning of the DDI2 constructs with C-terminal tag, combination of NdeI and KpnI restriction endonucleases was used to cleave the Strep-p905 vector. Composition of the used reaction mixture is listed in Table 5. Commercial reaction buffers NEBuffer (New England Biolabs) were used for individual restriction endonucleases according to manufacturer's instructions. The cleavage was done in 37°C (incubator IPP 400) for 5 h and was performed sequentially: after the first cleavage, the DNA was purified using the DNA Clean & Concentrator kit (Zymo Research) and added into next cleavage mixture. The cleavage was analyzed using agarose electrophoresis, which was used for final purification of the DNA as well (chapter 4.2.3. on the following page). Prior to agarose electrophoresis purification, the cleaved vectors (not the inserts) were dephosphorylated: 6 µl of Antarctic phosphatase buffer (NEB), 3 µl of deionised water and 1 µl of Antarctic phosphatase were added to the restriction cleavage reaction and incubated in 37 °C (incubator IPP 400) for 30 min.

Table 5: Restriction cleavage reaction mixture

NEBuffer	5 µl
Restriction endonuclease	1 µl
DNA solution	40 µl
Deionised H_2O	9 µl
Total volume	***50 µl***

4.2.3. Agarose electrophoresis

- *1% agarose gel: 0.5 g agarose, 50 ml TAE buffer, 1 µl GelRed 10 000x (Biotium)*
- *6x sample buffer: 40% (w/v) saccharose, 0.1% (w/v) bromphenol blue*
- *TAE buffer: 40 mM Tris, 20 mM acetic acid, 1 mM EDTA, pH 8.0*

Agarose gel solution was prepared in the horizontal electrophoresis apparatus (Gibco) according to manufacturer instructions. A DNA sample was mixed 5:1 with the 6x sample buffer and applied onto the gel. Electrophoresis was run at 120 V. Separation of the DNA was monitored by temporarily removing the gel from the apparatus and visualising it under UV light using camera Quantum ST4 (Vilber Lourmat). Depending on presumed size of the DNA, molecular markers GelPilot 100bp Plus Ladder (100-1500 bp range) or GelPilot 1kb Ladder (1000-10 000 bp range) from QIAGEN were run along the samples and used for size determination. If required, the separated DNA band was purified by cutting the area out of the gel and extracting the DNA from the gel using QIAquick Gel Extraction Kit (QIAGENE) according to manufacturer instructions.

4.2.4. Ligation – T4 ligase

50 ng of the p905 vector (cleaved and purified according to chapter 4.2.2., page 47) were used for the ligation reaction and the insert DNA coding for the respective RAD23B variant (cleaved and purified according to chapter 4.2.2., page 47) was added to reach 3:1 molar ratio of insert to vector. 2 µl of the T4 ligase buffer (NEB) were added and then the volume of the mixture was adjusted to 19 µl with deionised water. Finally, 1 µl of T4 ligase (NEB) was added and the reaction was incubated overnight at laboratory temperature. 5 µl of the reaction were then used to transform 40 µl of TOP10 *E. coli* competent cells according to chapter 4.2.6 (page 49).

4.2.5. Ligation – Gibson assembly

- *Gibson reaction mix: 6.7% PEG-8000, 133.3 mM Tris-HCl pH 7.5, 13.3 mM $MgCl_2$, 13.3 mM DTT, 0.27 mM each of the 4 dNTPs, 1.3 mM NAD, 12 µU/µl T5 exonuclease (NEB), 33.3 µU/µl Phusion polymerase (NEB), 5.7 µU/µl Taq ligase (NEB)*

50 ng of the Strep-p905 vector (cleaved and purified according to chapter 4.2.2., page 47) were used for the ligation reaction, the insert DNA coding for the respective DDI2 variant was added to reach 3:1 molar ratio of insert to vector. For the C-Strep DDI2 ΔHDD construct, which was assembled from two DNA fragments, each fragment was

added to reach 2:1 molar ratio to the vector. The DNA mixture was added to 15 µl of the Gibson reaction mixture, reaction volume was adjusted to 20 µl and the reaction was incubated in 50°C for 1 h in GeneAmp PCR System 2400 thermocycler instrument (Perkin-Elmer). 5 µl of the reaction were then used to transform 40 µl of TOP10 *E. coli* competent cells according to chapter 4.2.6.

4.2.6. Transformation of *E. coli*

- *LB medium: 20 g/l LB broth powder (Sigma Aldrich)*

40 µl of the competent *E. coli* culture was mixed with the respective DNA solution (either 5 µl of ligation reaction mix or 50 ng of purified plasmid) and incubated on ice for 15 min. Next, the culture was incubated in 42 °C for 90 s and then immediately put on ice for 5 min. 1 ml of the sterile LB medium was added to the culture and the culture was incubated for 1 h in 37 °C. 200 µl of the culture was then spread on the sterile LB agar containing ampicillin to select the transformed clones. The plasmids were purified from the selected clones using Zyppy Plasmid Miniprep Kit (ZYMO Research) according to manufacturer instructions.

4.2.7. Sequencing

For sequencing, 80-100 ng of the purified plasmid in 5 µl in deionised water were mixed with 5 µM water solution of sequencing primer. All constructs were sequenced with T7_F primer, constructs with inserts longer than 500 bp were additionally analysed using T7_R primer in separate sample (primer sequences in Table 1, page 45). The sequencing mixtures were sent to Eurofins-GATC company to be analysed by Sanger sequencing. Sequences provided by the company were aligned and compared with the designed sequences in UGENE software[228] to ensure the desired result was obtained from the cloning procedure.

4.3. Protein analysis methods

4.3.1. SDS-PAGE

- *Stacking gel: 5% (w/v) acrylamide solution (acrylamide and N,N'-bisacrylamide in a ratio 35.7:1), 250 mM Tris-HCl pH 6.8, 0.1% (w/v) SDS, 0.1% (w/v) ammonium persulfate (APS), 0.02% (v/v) TEMED*

- *Resolving gel: acrylamide solution (acrylamide and N,N'-bisacrylamide in a ratio 35.7:1) with variable acrylamide content depending on molecular weight of analysed protein, 250 mM Tris-HCl pH 8.8, 0.1% (w/v) SDS, 0.1% (w/v) ammonium persulfate (APS), 0.02% (v/v) TEMED*
- *Running buffer: 25 mM Tris pH 8.8, 250 mM glycine, 0.1% (w/v) SDS*
- *6x sample buffer: 360 mM Tris, pH 6.8, 30% (v/v) glycerol, 10% (w/v) SDS, 4% (v/v) 2-mercaptoethanol, 0.01% (w/v) bromophenol blue*
- *Fixing solution: 40% (v/v) methanol, 10% (v/v) acetic acid*
- *Staining solution: 1.2% (w/v) Coomassie G-250, 10% (v/v) phosphoric acid, 10% (w/v) ammonium sulphate, 20% (v/v) methanol, prepared according to chapter 4.3.2.*

Gels for Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) were prepared and performed in vertical electrophoresis apparatus (Mini-PROTEAN® System, BioRad) according to the manufacturer instructions. Resolving gels of different percentages of the acrylamide were used based on the molecular weight of the analysed samples. Protein sample was mixed 5:1 with 6x sample buffer, incubated for 10 min at room temperature and applied onto the gel. Electrophoresis was run at 150 V and the whole apparatus was cooled with ice. Molecular weight standard named All Blue Marker (BioRad) was run alongside the samples to allow live monitoring of the separation.

After electrophoresis, the gel was removed from the instrument and incubated in the fixing solution for at least 30 min. Next, the gel was transferred into the staining solution and left shaking for at least 1 h. After the staining, the gel was thoroughly washed with distilled water. Then it was left shaking for at least 2 h, preferably overnight, to decrease the background staining. The gel was scanned with LI-COR Odyssey CLx Near-Infrared Fluorescence Imaging System using Coomassie G-250 fluorescence at 800 nm.

4.3.2. Colloidal Coomassie G-250 solution preparation

1) 118 ml of 85% phosphoric acid was added to 100 ml of deionised water on magnetic stirrer and thoroughly mixed
2) 100 g of ammonium sulphate was added to the solution and mixed until complete dissolution
3) 1.2 g of Coomassie G-250 was added to the solution and mixed until complete dissolution
4) Deionised water was added upon stirring until volume of 800 ml was reached

5) 200 ml of anhydrous methanol was added upon stirring, reaching total volume of 1000 ml

Exact sequence of ingredients mixing is necessary for reaching maximal sensitivity of the staining. Prepared solution was filtered through a paper filter, transferred into a brown bottle and stored in 4 °C.

4.3.3. Western blotting

- *Blotting buffer: 12.5 mM Tris-glycine pH 8.3, 10% (v/v) methanol*
- *PBST buffer: 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 137 mM NaCl, 2.7 mM KCl, pH 7.4, 0.05% (v/v) Tween 20*
- *Blocking buffer: PBST buffer, 5% (v/v) Blocker Casein in PBS solution (Thermo Scientific)*

Samples for western blot analysis were first separated in polyacrylamide gel according to chapter 4.3.1. (page 49). The gel was then transferred into Mini Trans-Blot® Cell (Bio-Rad) apparatus. The apparatus was filled with blotting buffer and electroblotting was performed at 100 V for 1 hour. Proteins were transferred to nitrocellulose membrane (BioRad).

The membrane was rinsed with PBST buffer and subsequently incubated in blocking buffer. After blocking, the membrane was transferred into the blocking buffer containing primary antibody or iBody for at least 4 h. Anti-polyHis 800 iBody (IOCB, dilution 1:5000) was used to detect His-tagged proteins and StrepMAB-Classic-HRP (IBA, dilution 1:10 000) was used to visualize Strep-tagged proteins. After incubation, the membrane was rinsed 3x for 5 min in PBST buffer. As anti-polyHis 800 iBody contains the fluorophore molecules, the detection was possible immediately. The membrane with bound StrepMAB-Classic-HRP was incubated with fluorophore-labelled secondary antibody IRDye® 800CW Goat anti-Mouse IgG (LI-COR, dilution 1:25 000) in the blocking buffer for at least 1 h. The membrane was then rinsed 3x for 5 min in PBST buffer. The visualization was done using Odyssey CLx Near-Infrared Fluorescence Imaging System (LI-COR) using fluorescence at 800 nm and simultaneous detection of All Blue Marker fluorescence at 700 nm.

4.3.4. Protein concentration measurement

The protein concentration was measured spectrophotometrically using Thermo Scientific™ NanoDrop™ One spectrophotometer, at 280 nm (A_{280}). A clean buffer was used to zero the instrument, then absorbance of a protein solution was recorded.

Theoretical absorbance of a solution containing protein of interest in concentration of 1 mg/ml, and the molecular weight of the protein were calculated based on the protein sequence using the online Expasy Protparam tool (QIAGEN). Predicted properties and measured A_{280} were then used to calculate mass concentration and molar concentration of the protein solution using the following equations:

$$c_m = \frac{A_{280}(measured)}{A_{280}(1mg/ml)}\ mg/ml$$

$$c = \frac{c_m}{M} * 10^6\ \mu mol/l$$

c_m…… *mass concentration [mg/ml]*

c…… *molar concentration [µmol/l]*

A_{280}*(measured)*..... *Measured absorbance of solution at 280 nm*

A_{280}*(1mg/ml)*..... *Predicted absorbance of solution for protein concentration 1mg/ml*

M…… *Molecular weight of measured protein [g/mol]*

4.4. Recombinant protein preparation

4.4.1. Bacterial protein expression

- *<u>LB medium:</u> 20 g/l LB broth powder (Sigma Aldrich), 100 µg/ml ampicillin*
- *<u>M9 minimal medium</u>: 6.8 g/l $Na_2HPO_4 \cdot 12H_2O$, 3g/l KH_2PO_4 , 11.2 mg/l Na_2SO_4, 50 mg/l $CaCl_2 \cdot 2H_20$, 5 mg/l $FeCl_2 \cdot 4H_2O$, 1.2 mg/l $MnCl_2 \cdot 4H_2O$, 0.8 mg/l $CoCl_2 \cdot 6H_2O$, 0.34 mg/l $ZnCl_2$, 0.3 mg/l $CuCl_2 \cdot 2H_2O$, 20 µg/l H_3BO_3, 0.61 mg/l $Na_2MoO_4 \cdot 2H_2O$, 5 mg/l EDTA, 0.25 g/l $MgSO_4 \cdot 7H_20$, 1 mg/l d-Biotin, 1 mg/l thiamine, 0.81 g/l NH_4Cl, 4 g/l glucose, 100 µg/ml ampicillin*
- *<u>Buffer A:</u> 50 mM Tris pH 8.0, 50 mM NaCl, 1 pill of Roche Diagnostics GbmH Complete Mini inhibitor cocktail per 20 ml of buffer*

All protein variants were expressed using BL21(DE3)RIL *E. coli* strain . Cells were transformed with plasmid encoding gene of interest according to the chapter 4.2.6. (page 49) Isotopically labelled proteins for the NMR analysis were expressed in the M9 medium, where the ^{15}N isotope was supplied in the form of isotopically enriched ammonium chloride and ^{13}C isotope in the form of isotopically enriched glucose. The unlabelled proteins were expressed in the LB medium. Apart from the media composition,

expression and purification procedures for ^{15}N-labelled, $^{15}N/^{13}C$-labelled and unlabelled proteins were the same.

Inoculum was prepared by growing overnight bacterial culture in 5 ml of LB medium. Erlenmeyer flask containing 0.5 l of growth medium was prepared, autoclaved, after cooling supplied with 100 µg/ml ampicillin and inoculated with 5 ml of inoculum. The culture was incubated in Innova® 44/44 R incubator in 37 °C and 220 RPM.

Optical density at 595nm (OD_{595}) was checked during the bacterial growth using UNICAM UV 500 spectrophotometer. When the OD_{595} reached approximately 0.7, 100 µl sample was taken for SDS-PAGE analysis and IPTG was added to final concentration of 0.75 mM to induce the protein expression. Then, the temperature was decreased to 18 °C and culture was left shaking overnight.

Before the harvest, a 100 µl sample was taken for SDS-PAGE analysis. The bacterial culture was centrifuged (6000 g, 20 min, Sorvall Evolution RC centrifuge), the supernatant was discarded, and the pellet was resuspended in 40 ml of buffer A. The suspension was homogenised in Potter-Elvehjem homogeniser to get rid of solid particles. The resuspended culture was then lysed by three passages through EmulsiFlex-C3 instrument. The lysate was then centrifuged (30 000 g, 30 min, Sorvall Evolution RC centrifuge), the pellet was discarded, and the soluble fraction was used for further purification.

4.4.2. Affinity chromatography: Strep-tagged protein purification

- *Buffer W: 50 mM Tris pH 8.0, 150 mM NaCl*
- *Buffer E: 50 mM Tris pH 8.0, 150 mM NaCl, 10 mM desthiobiotin*
- *Buffer R: 50 mM Tris pH 8.0, 150 mM NaCl, 1 mM HABA*
- *PBS buffer: 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 137 mM NaCl, 2.7 mM KCl, pH 7.4*

40 ml of bacterial lysate (prepared according to chapter 4.4.1., page 52) containing recombinant Strep-tagged protein was mixed in a falcon tube with 2 ml of IBA Strep-Tactin® Sepharose resin in buffer W and was left to incubate on a rotary mixer in 4 °C for 3 hours. After incubation, mixture of lysate and resin was then transferred into 5 ml gravity flow column and flow through fraction was collected. Resin collected in the column was washed two times with 5 ml of buffer W. The column output was plugged, and 3 ml of buffer E were added into the column. Resin was incubated for 1 hour with

buffer E to ensure complete elution. Next, 3 ml of buffer E containing eluted protein were collected. 5 ml of buffer R were added to the column for resin regeneration, changing the resin colour to red. The column was then washed with buffer W until red colour disappeared, signalling completed regeneration. Resin was then transferred into a falcon tube, mixed with the flow through fraction and the whole process was repeated, yielding 6 ml of buffer E containing Strep-tagged recombinant protein in total. Eluted fraction was transferred into a Spectrapore dialysis membrane tube and dialysed overnight against 3 l of PBS buffer. After dialysis, protein solution was centrifuged (16000 g, 20 min, Eppendorf 5415 R centrifuge) to remove precipitate and protein concentration was measured according to chapter 4.3.4. (page 51).

4.4.3. Affinity chromatography: His-tagged protein purification

- *Buffer W: 50 mM Tris pH 8.0, 300 mM NaCl, 20 mM imidazole*
- *Buffer E: 50 mM Tris pH 8.0, 300 mM NaCl, 250 mM imidazole*
- *PBS buffer: 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 137 mM NaCl, 2.7 mM KCl, pH 7.4*

40 ml of bacterial lysate (prepared according to chapter 4.4.1., page 52) containing recombinant His-tagged protein was mixed in a falcon tube with 2 ml of QIAGEN NiNTA Agarose resin in buffer W and was left to incubate on a rotary mixer in 4 °C for 3 hours. After incubation, mixture of lysate and resin was then transferred into 5 ml gravity flow column and flow through fraction was collected. Resin collected in the column was washed two times with 5 ml of buffer W. Resin was incubated for 1 hour with 3 ml of buffer E to ensure complete elution. Next, 3 ml of buffer E containing eluted protein were collected. Resin was washed with 5 ml of buffer W and then transferred into a falcon tube, mixed with the flow through fraction and the whole process was repeated, yielding 6 ml of buffer E containing His-tagged recombinant protein in total. Eluted fraction was transferred into Spectrapore dialysis membrane tube and dialysed overnight against 3 l of PBS buffer. After dialysis, protein solution was centrifuged (16000 g, 20 min, Eppendorf 5415 R centrifuge) to remove precipitate and protein concentration was measured according to chapter 4.3.4. (page 51).

4.4.4. Gel filtration chromatography

- *PBS buffer: 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 137 mM NaCl, 2.7 mM KCl, pH 7.4, filtered, degassed*

Protein solution was concentrated with Amicon Ultra centrifugal filters (Millipore) to reduce its volume to 2 ml. Concentrated solution was centrifuged (16000 g, 30 min, Eppendorf 5415 R centrifuge) to remove precipitate, and concentration was measured according to chapter 4.6. Solution was injected into the FPLC system (ÄKTA Explorer, Amersham Pharmacia Biotech) using a 2 ml loading tube. System was connected to either SuperdexTM 75pg 16/60 or 200pg 16/60 FPLC columns, depending on the molecular mass of a purified protein (75pg column for proteins up to 30 kDa, 200pg for proteins of higher molecular mass). PBS buffer was used as a mobile phase with flow set to 0.5 ml/min and purification process was monitored using A_{280} sensor. Relevant fractions were analysed using SDS-PAGE (chapter 4.3.1., page 49), fractions containing protein of interest were combined and concentration was measured according to chapter 4.3.4. (page 51).

4.5. Protein-protein interaction pull-down assay

The affinity pull-down assay allows testing of interaction between two proteins. First protein is bound to the resin through an affinity tag (bait-protein). The second protein (prey-protein) is then incubated with the resin, the resin is subsequently washed with a buffer and finally the elution buffer is applied to release the bait-protein from the resin. Interaction is confirmed when the prey-protein is detected in the eluted solution.

4.5.1. Affinity pull-down using Strep-Tactin resin

- *PBS buffer: 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 137 mM NaCl, 2.7 mM KCl, pH 7.4*
- *PBS-BSA buffer: PBS buffer, 2 mg/ml BSA*
- *Buffer E: PBS buffer, 10 mM desthiobiotin*

25 µl of IBA Strep-Tactin® Sepharose resin were added into a 2 ml gravity flow column and subsequently washed three times with 1 ml of PBS buffer. Next, 1 ml of bacterial lysate containing bait-protein prepared according to chapter 4.4.1. (page 53) was added into the column and incubated with the resin for 2 h at laboratory temperature. Resin was washed twice with 200 µl of PBS buffer. Next, 200 µl of 40 µM His-tagged prey-protein in PBS-BSA buffer were added into the column and incubated with the resin

for 2 hours at laboratory temperature. After the incubation, the resin was washed twice with 100 µl of PBS buffer. Finally, 50 µl of buffer E were added for 30 min to elute proteins from the resin. The eluate was analysed by SDS-PAGE (chapter 4.3.1., page 49) and western blotting (chapter 4.3.3., page 51).

4.5.2. Affinity pull-down using NiNTA resin

- *PBS buffer: 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 137 mM NaCl, 2.7 mM KCl, pH 7.4*
- *PBS-BSA buffer: PBS buffer, 2 mg/ml BSA*
- *Buffer E: PBS buffer, 250 mM imidazole*

25 µl of NiNTA (QIAGEN) beads were added into a 2 ml gravity flow column and subsequently washed three times with 1 ml of PBS buffer. Next, 1 ml of bacterial lysate bait-protein prepared according to chapter 4.4.1 (page 53) was added into the column and incubated with the resin for 2 hours at laboratory temperature. Resin was washed twice with 200 µl of PBS buffer, 200 µl of 40 µM Strep-tagged prey-protein in PBS-BSA buffer were added to the column and incubated with the resin for 2 hours at laboratory temperature. After the incubation, the resin was washed twice with 100 µl of PBS buffer. Finally, 50 µl of buffer E were added for 30 min to elute proteins from the resin. The eluate was analysed by SDS-PAGE (chapter 4.3.1., page 49) and western blotting (chapter 4.3.3., page 51).

4.6. Nuclear magnetic resonance experiments

- *PBS buffer: 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 137 mM NaCl, 2.7 mM KCl, pH 7.4, 10% deuterium oxide (v/v)*

All measurements were performed on 850 MHz Bruker Avance spectrometer (Bruker BioSpin GmbH) equipped with a triple resonance (15N, 13C, 1H) cryoprobe at IOCB Prague. PBS buffer with 10% deuterium oxide content was used for all measurements. Spectra visualisation, adjustments and resonance assignments were done manually in Sparky-NMRFAM software (UCSF). Data acquisition and processing were done under the guidance from Ing. Václav Veverka, Ph.D. and Pavel Srb, Ph.D. (IOCB Prague).

4.6.1. The DDI2$^{1-212}$ UBL-HDD sequence resonance assignment

To be able to interpret the $^{15}N/^{1}H$ HSQC spectrum, the peaks in the spectrum need to be assigned to individual residues of the analysed protein. For the sequence resonance assignment, 350 µl of protein sample with 220 µM $^{15}N/^{13}C$ labelled DDI2$^{1-212}$ UBL-HDD protein was used to collect the spectra. The $^{15}N/^{1}H$ HSQC 2D spectrum and the $^{15}N/^{13}C/^{1}H$ HNCO, HN(CA)CO, HNCACB and CACB(CO)NH 3D spectra were collected. During assignment, previously solved ^{15}N/1H HSQC spectra and peak lists of the DDI2 UBL domain and DDI2 HDD domain[1] solved in our laboratory were used for comparison with the newly assigned peaks of the DDI2$^{1-212}$ UBL-HDD to aid with the assignment process. Summary of assignment approach using complementary spectra is shown in Figure 11.

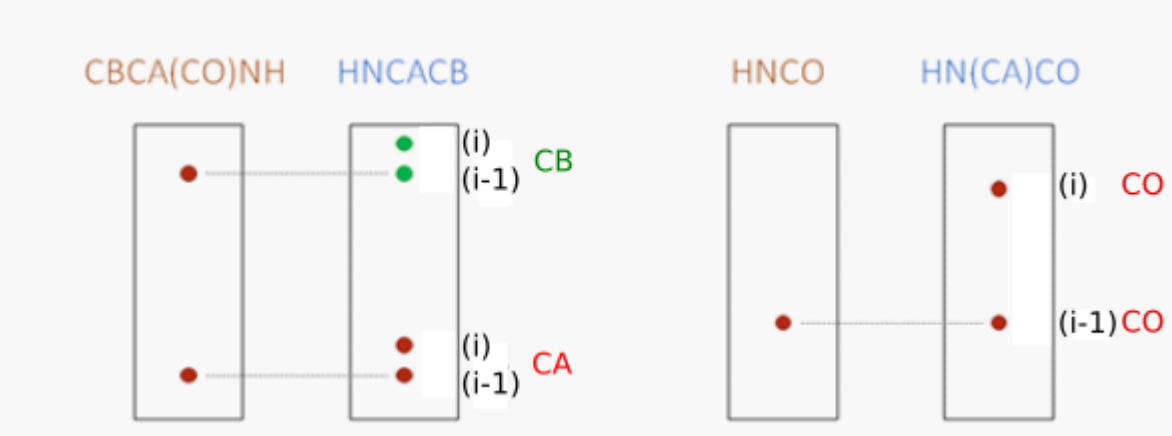

Figure 11: Schematic representation of assignment strategy used in this thesis. For residue in position (i) in the sequence, the peaks of specific chemical shifts in the ^{13}C axis are detected for the residues in position (i) and (i-1). The peaks are related to $^{15}N/^{1}H$ HSQC 2D spectrum through their ^{1}H and ^{15}N chemical shifts. This setup allows for successive assignment of the $^{15}N/^{1}H$ HSQC peaks along the sequence.

4.6.2. The DDI2$^{1-212}$ UBL-HDD and RAD23B FL/RAD23B$^{1-82}$ UBL interaction characterisation

For detailed characterisation of the interaction of the DDI2$^{1-212}$ UBL-HDD with the RAD23B FL protein and the RAD23B$^{1-82}$ UBL, six samples containing 25 µM ^{15}N-labelled DDI2$^{1-212}$ UBL-HDD protein were used, containing varying concentrations of unlabelled binding partner set to 0-fold, 0.5-fold (12.5 µM), 1-fold (25 µM), 2-fold (50 µM), 4-fold (100 µM) and 8-fold (200 µM) molar excess of either RAD23B FL or RAD23B$^{1-82}$ UBL. $^{15}N/^{1}H$ HSQC spectra were recorded for each sample at 25 °C.

Evaluation of titrations was done in Sparky-NMRFAM software. Spectra recorded with growing concentration of the binding partner were overlaid and the changes in peak positions manually tracked. Perturbation analysis tool of Sparky-NMRFAM software was used for evaluation. For each peak that was tracked reliably, the relative peak intensity (RPI) was calculated as ratio of the peak intensity with defined amount of the binding partner to the peak intensity without binding partner present. Chemical shit perturbation (CSP) was calculated as distance between peak position in 1:0 and 1:8 molar ratio spectra using formula:

$$CSP = \sqrt{(\omega^{H}_{1:0} - \omega^{H}_{1:8})^2 + (0.2 * \omega^{N}_{1:0} - 0.2 * \omega^{N}_{1:8})^2}$$

CSP ..... Chemical shift perturbation
$\omega^{H}_{1:0}$..... Peak chemical shift in ppm on ^{1}H axis without the binding partner
$\omega^{H}_{1:8}$..... Peak chemical shift in ppm on ^{1}H axis with 8-fold molar excess of the binding partner
$\omega^{N}_{1:0}$..... Peak chemical shift in ppm on ^{15}N axis without the binding partner
$\omega^{N}_{1:8}$..... Peak chemical shift in ppm on ^{15}N axis with 8-fold molar excess of the binding partner

Average CSP and standard deviation of the CSP were calculated for the set of evaluated peaks. Peak CSPs higher than one standard deviation than the average were classified as significant.

TITAN, 2D line-shape fitting software, was applied to the acquired sets of titration spectra for dissociation constant (K_d) determination. TITAN software uses both shifts in peak positions and changes of peak intensities for K_d calculation. The Phe17, Val21, Asp22, Ala23, Leu33, Leu36, Glu44, Glu26 peaks for the RAD23B$^{1-82}$ UBL titration and the Phe(-6), Val71, Asp133, Phe198, Asn145, Leu137, Arg155, Phe193, Arg9 peaks for the RAD23B FL titration were used for fitting. The K_d fitting and calculations were done by Pavel Srb, Ph.D.

5 Results

5.1. Designed protein constructs

To investigate the interaction of human DDI2 and RAD23B proteins on the molecular level, we designed various DNA constructs encoding N- or C-terminally Strep-tagged full-length and domain-deleted variants of DDI2 protein and full-length and domain-deleted variants of RAD23B with N-terminal His-tag. Individual constructs of DDI2 were designed according to previous publication from our group[1], constructs encoding variants of RAD23B protein were designed according to UniProt database[232]. Molecular weights and A_{280} of protein variants predicted with Expasy Protparam (https://web.expasy.org/protparam) are listed in Table 6 (page 61). Schematic representations of the design of individual protein variants are depicted in Figure 12.

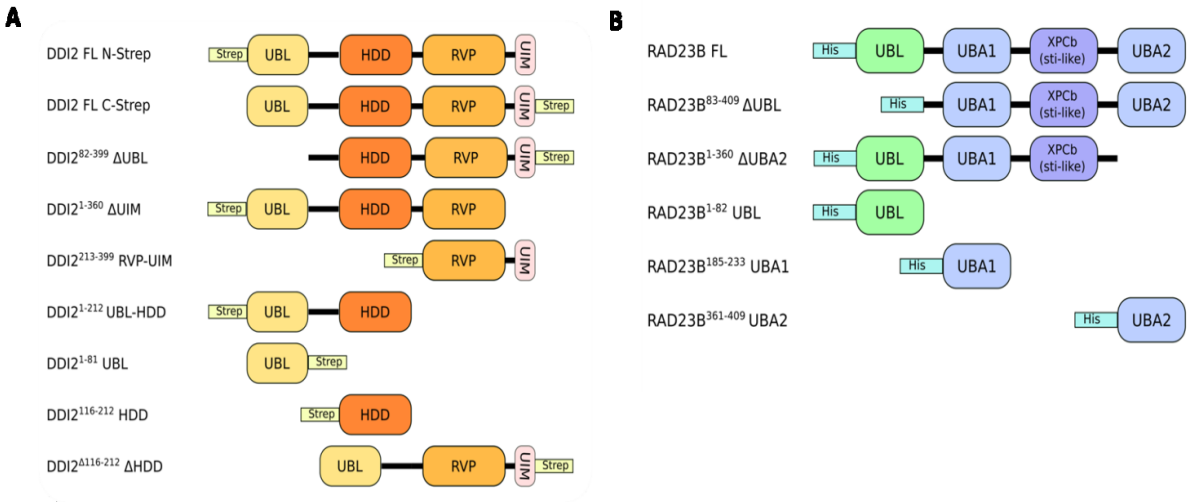

Figure 12: Schematic representation of individual protein constructs representing full-length and truncated variants of DDI2 and RAD23B used in this study with indicated positions of recombinant tags (Strep-tag – yellow rectangle, His-tag – blue rectangle). **A)** Protein constructs of the human DDI2 protein. UBL – ubiquitin-like domain, HDD – helical domain of Ddi1-like proteins, RVP – retroviral protease-like domain, UIM – ubiquitin-like motif. **B)** Protein constructs of the human RAD23B protein. UBL – ubiquitin-like domain, UBA – ubiquitin-associated domain, XPCb – XPC-binding domain (or Sti1-like domain).

Table 6: Molecular weights and A_{280} of the designed protein variants

Protein variant	MW [Da]	A_{280} for c = 1 mg/ml
C-Strep DDI2 FL	46946.15	0.425
N-Strep DDI2 FL	47077.35	0.424
$DDI2^{82-399}$ ΔUBL	37938.04	0.408
$DDI2^{1-81}$ UBL	11580.90	0.990
$DDI2^{\Delta116-212}$ ΔHDD	35934.85	0.555
$DDI2^{116-212}$ HDD	13715.28	0.510
$DDI2^{1-360}$ ΔUIM	42780.65	0.431
$DDI2^{212-399}$ RVP-UIM	23246.56	0.665
$DDI2^{1-212}$ UBL-HDD	26403.58	0.434
RAD23B FL	46067.25	0.259
$RAD23B^{1-82}$ UBL	12123.83	0.369
$RAD23B^{185-233}$ UBA1	8420.32	0.708
$RAD23B^{361-409}$ UBA2	8480.32	0.351
$RAD23B^{83-409}$ ΔUBL	36857.50	0.243
$RAD23B^{1-360}$ ΔUBA2	40501.00	0.258

5.2. Cloning

All DNA constructs were cloned into the expression vectors using protocols described in chapter 4.2. (page 45): the RAD23B constructs were cloned into the p905 vector, the DDI2 constructs were cloned into the Strep-p905 vector. Electrophoretic analysis of the PCR amplification of the respective sequences is shown in Figure 13 on the following page. Representative picture of the electrophoretic analysis of the Strep-p905 vector is shown in Figure 14 on the following page. The amplified sequences were ligated into the vectors, the ligation products were transformed into the TOP10 cells and the amplified plasmids were extracted from the selected clones. The cloning process was verified by Sanger sequencing (GATC, Biotech, Germany).

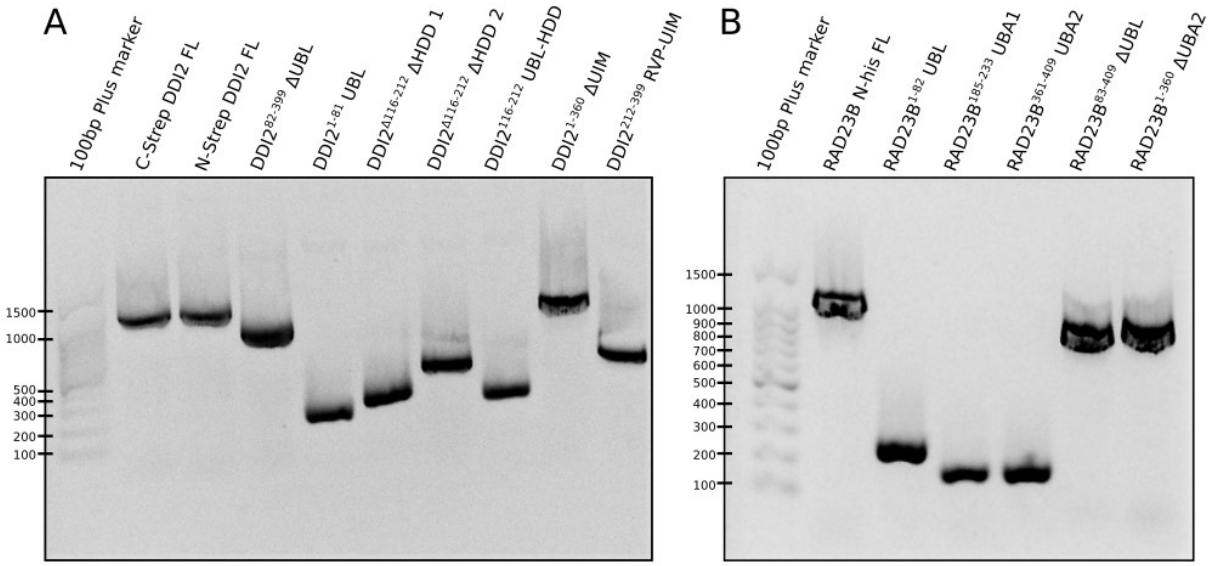

Figure 13: Electrophoretic analysis of the PCR amplification of the respective DNA sequences. 2.5 µl of the PCR reactions and 2 µl of molecular marker were loaded into the agarose gel. Position of the samples is indicated above the lines. For the expected sizes of the PCR products, please refer to the Table 4 (page 46-47). **A)** PCR amplification of the DDI2 constructs. **B)** PCR amplification of the RAD23B constructs.

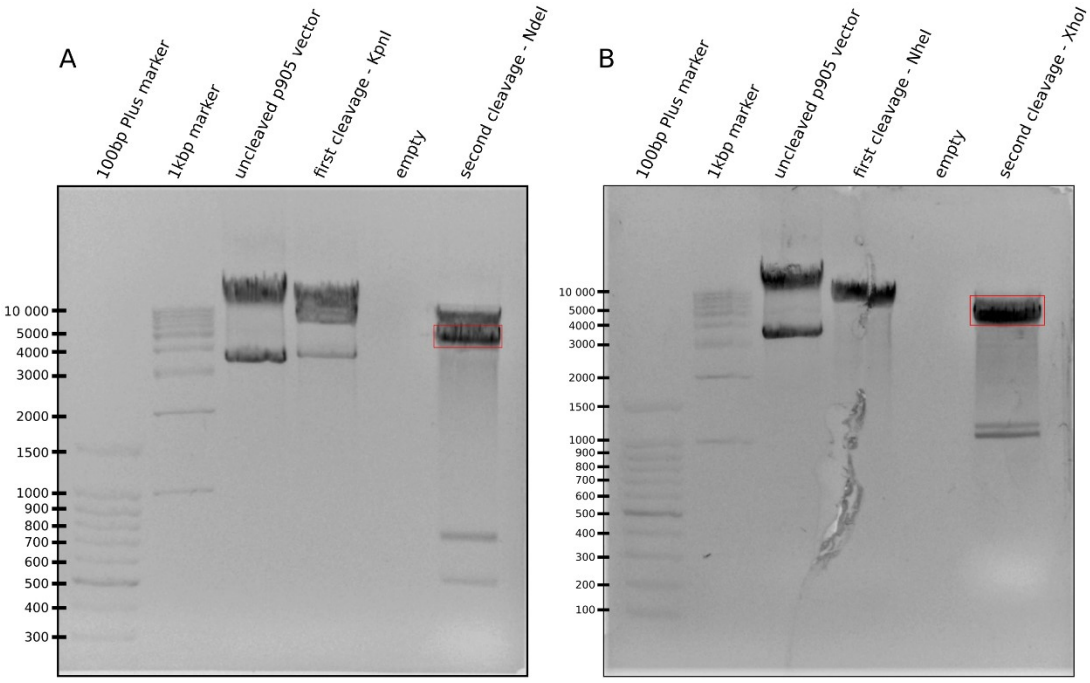

Figure 14: Electrophoretic analysis of the Strep-p905 vector restriction cleavage. **A)** The cleavage of the Strep-p905 vector for the cloning with C-terminal Strep-tag. **B)** The cleavage of the Strep-p905 vector for the cloning with N-terminal Strep-tag. 2 µl of molecular marker, 5 µl of the restriction cleavage mixture and 100 ng of uncleaved vector were loaded into the agarose gel. Positions of the samples are indicated above the lines. Positions of the bands that were cut out and further purified are indicated by the red squares.

5.3. Protein expression and purification

5.3.1. Affinity purification

All proteins were recombinantly expressed in *E. coli* and the lysates were prepared as described in chapter 4.4.1. (page 52). Affinity chromatography was used for the purification of the protein variants as follows: RAD23B FL and RAD23B$^{1-82}$ UBL were purified using NiNTA agarose beads as described in chapter 4.4.3. (page 54), DDI2 FL C-Strep, DDI2 FL N-Strep and DDI2$^{1-212}$ UBL-HDD were purified using Strep-Tactin® Sepharose resin as described in chapter 4.4.2. (page 53). The purification procedure was analyzed by SDS-PAGE according to chapter 4.3.1. (page 49) as follows: the samples representing the lysates, the lysates after incubation with affinity resin ("Flow through"), the samples collected after the first and the second washing steps ("Wash1" and "Wash2") and the elution fraction ("Elution"). The resulting SDS-PAGE gels are shown for RAD23B FL, DDI2 FL C-Strep and DDI2 FL N-Strep in Figure 15, for RAD23B$^{1-82}$ UBL and DDI2$^{1-212}$ UBL-HDD in Figure 16 on the following page.

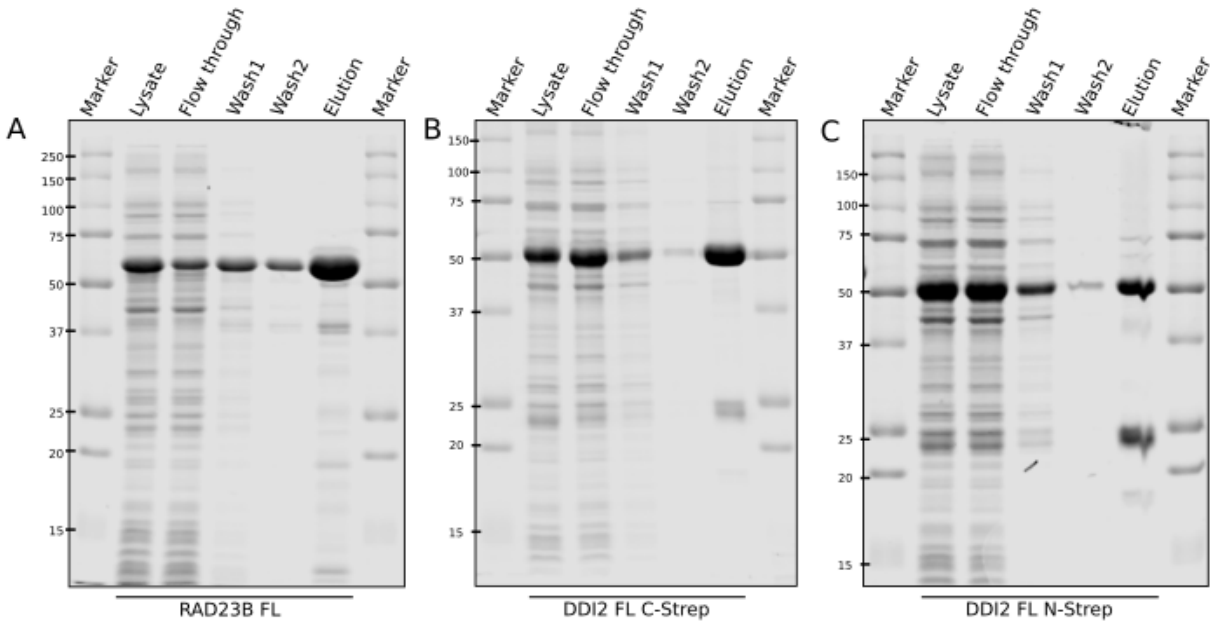

Figure 15: Affinity purification of RAD23B FL (**A**), DDI2 FL C-Strep (**B**) and DDI2 FL N-Strep (**C**) analyzed by SDS-PAGE. Molecular weight marker (1.5 µl) and 2 µl of each sample were loaded onto the 12% polyacrylamide gel. Sample positions are indicated.

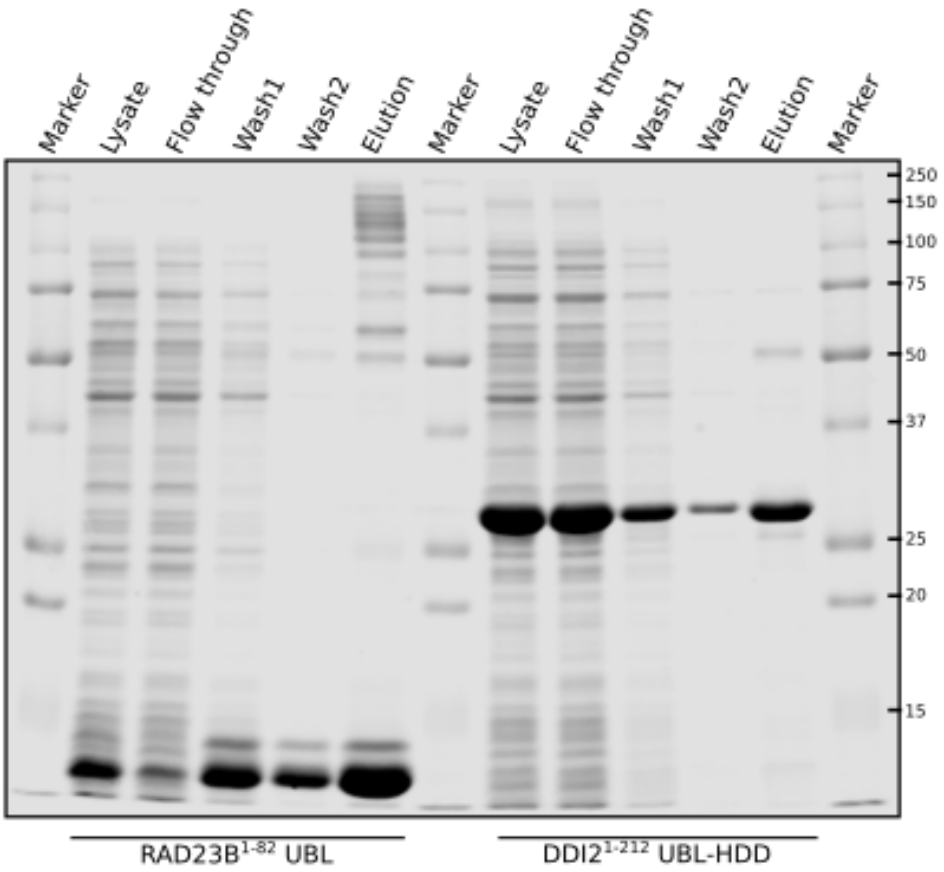

Figure 16: Affinity purification of RAD23B$^{1-82}$ UBL and DDI2$^{1-212}$ UBL-HDD protein variants analysed by SDS-PAGE. Molecular weight marker (1.5 µl) and 2 µl of each sample were loaded onto the 16% polyacrylamide gel. Sample positions are indicated.

5.3.2. Gel filtration chromatography purification

RAD23B FL, RAD23B$^{1-82}$ UBL and DDI2$^{1-212}$ UBL-HDD protein variants were further purified by gel filtration chromatography according to chapter 4.4.4. (page 55). Protein concentrations of the samples loaded into the column were as follows: RAD23B FL ~12 mg/ml, RAD23B$^{1-82}$ UBL ~9 mg/ml, DDI2$^{1-212}$ UBL-HDD ~3.5 mg/ml. Chromatogram corresponding to DDI2$^{1-212}$ UBL-HDD protein purification is shown in Figure 17 (page 64), chromatogram corresponding to RAD23B$^{1-82}$ UBL is depicted in Figure 18 (page 64) and chromatogram corresponding to RAD23B FL is shown in Figure 19 (page 65). Fractions corresponding to the peaks detected at A_{280} nm were further analyzed by SDS-PAGE (chapter 4.3.1., page 49) as follows: fractions A4-B10 and B3-C4 of DDI2$^{1-212}$ UBL-HDD purification, fractions C1-D11 of RAD23B$^{1-82}$ UBL purification, and fractions B11-C12 of RAD23B FL purification. Resulting SDS-PAGE gels are shown in Figure 20 (page 65).

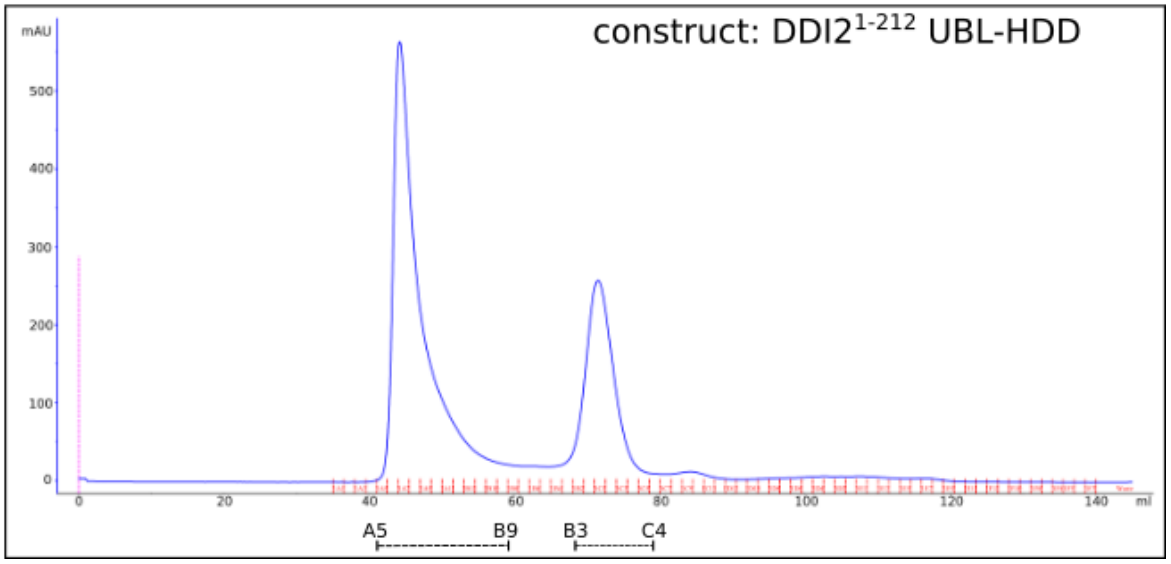

Figure 17: Chromatogram of gel filtration chromatography of DDI2$^{1-212}$ UBL-HDD protein variant. Pink dashed line marks sample injection, blue curve shows A_{280} on the column output in relation to volume of the eluted mobile phase. Red lines mark collected fractions. Black line marks the fractions further analysed by SDS-PAGE.

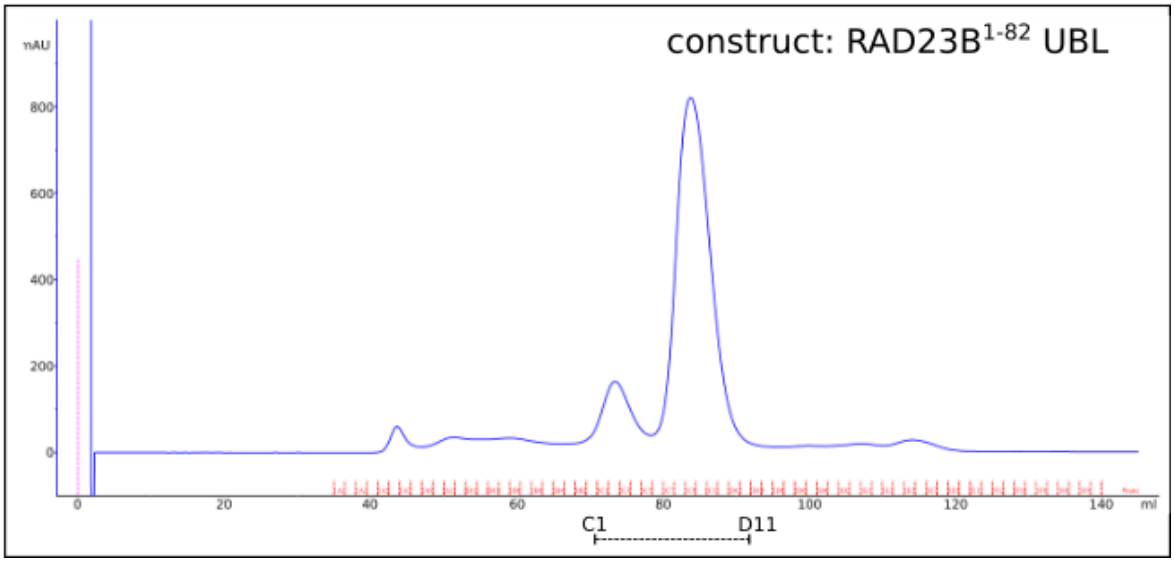

Figure 18: Chromatogram of gel filtration chromatography of RAD23B$^{1-82}$ UBL protein variant. Pink dashed line marks sample injection, blue curve shows A_{280} on the column output in relation to volume of the eluted mobile phase. Red lines mark collected fractions. Black line marks the fractions further analysed by SDS-PAGE.

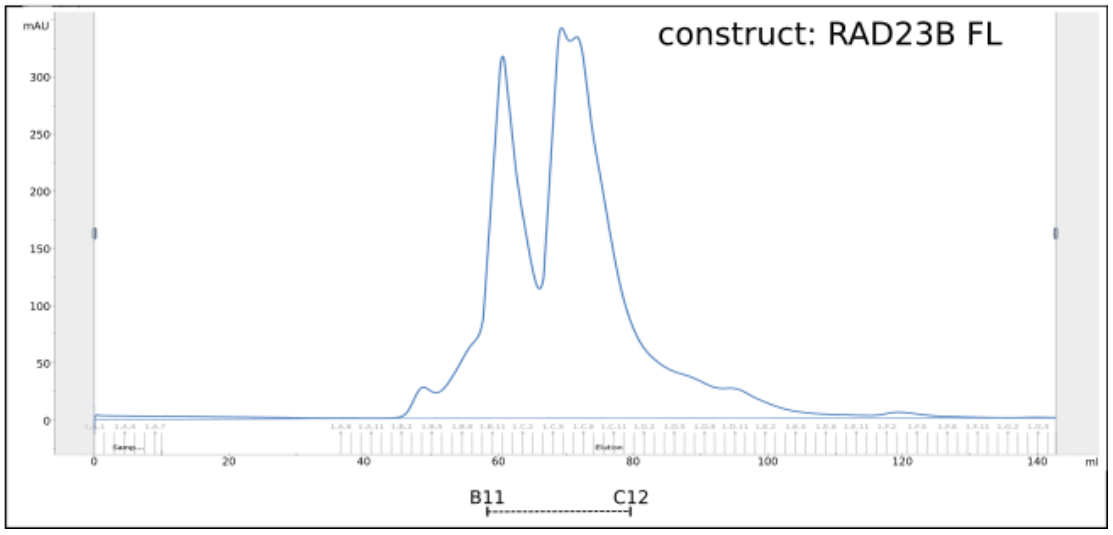

Figure 19: Chromatogram of gel filtration chromatography of RAD23B FL protein variant. Blue curve shows A_{280} on the column output in relation to volume of the eluted mobile phase. Grey lines mark collected fractions. Black line marks the fractions further analysed by SDS-PAGE.

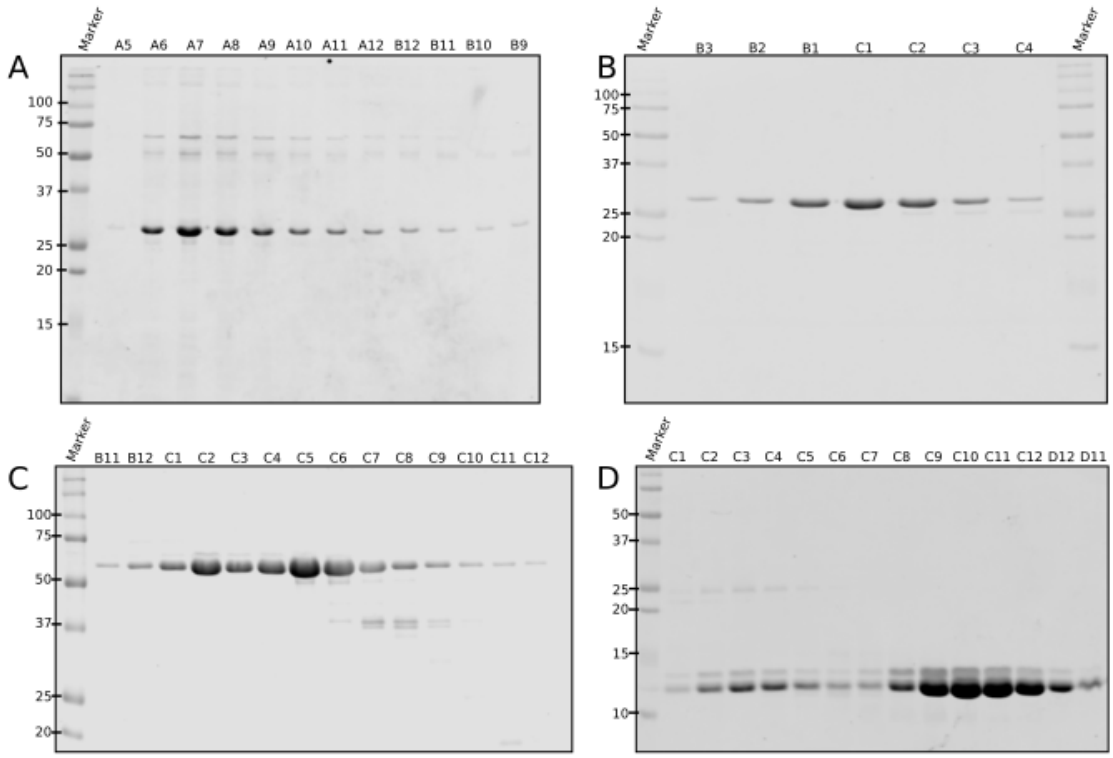

Figure 20: SDS-PAGE analysis of the gel filtration chromatography. 2 µl of the samples prepared from individual indicated fractions and 1.5 µl of molecular weight marker were loaded into the gel. **A)** Analysis of the A4-A12 and B12-B10 fractions of the DDI2$^{1-212}$ UBL-HDD purification using 16% polyacrylamide gel. **B)** Analysis of the B3-B1 and C1-C4 fractions of the DDI2$^{1-212}$ UBL-HDD purification using 16% polyacrylamide gel. **C)** Analysis of the B11-B12 and C1-C12 fractions of the RAD23B FL purification using 12% polyacrylamide gel. **D)** Analysis of the C1-C12 and D12-B11 fractions of the RAD23B$^{1-82}$ UBL purification using 16% polyacrylamide gel.

5.4. Protein-protein interaction pull-down assay

To map the interaction of the human proteins DDI2 and RAD23B, the affinity-based pull-down assays were performed. The bacterial lysates containing overexpressed respective proteins were prepared according to chapter 4.4.1. The DDI2 FL N-Strep and the RAD23B FL proteins were purified by affinity chromatography (chapter 4.4.2., page 53 and 4.4.3., page 54), protein concentrations were measured as described in chapter 4.3.4 (page 51) and adjusted to 40 µM, the BSA was added to final concentration of 2 mg/ml to minimize non-specific binding. The pull-down interaction assays were performed as described in chapter 4.5.1. (page 55) to map the interaction of the DDI2 variants with RAD23B FL. Similarly, the interaction of the DDI2 with specific region of the RAD23B was mapped using various RAD23B protein variants and DDI2 FL N-Strep and the pull-down assays were performed as described in chapter 4.5.2. (page 56). Eluted samples were analysed by SDS-PAGE (chapter 4.3.1., page 49) and western blotting (chapter 4.3.3., page 51). Results of DDI2 mapping are shown in Figure 21, results of RAD23B mapping are shown in Figure 22 (page 67).

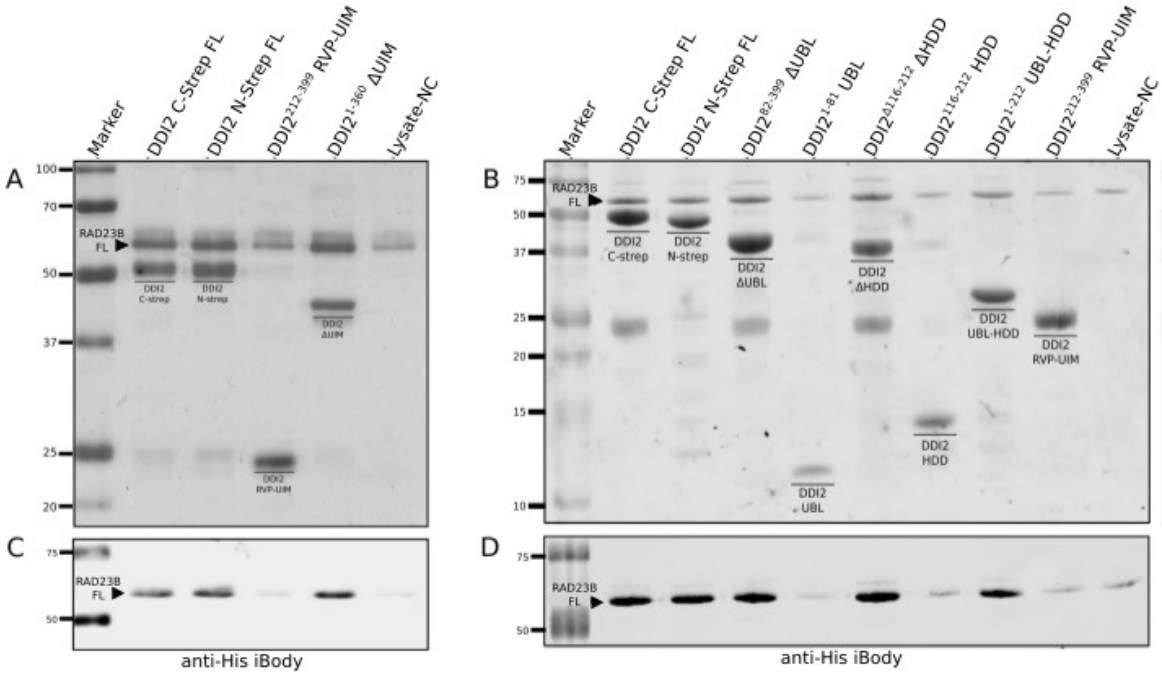

Figure 21: DDI2-RAD23B interaction mapping performed by the pull-down assays using Streptactin resin with the DDI2 protein variants as a bait and RAD23B as a prey. 1.5 µl of molecular weigth marker and 2 µl of the protein samples were loaded into the gel. Bait proteins for each sample are noted at the top and the relevant bait protein bands are marked specifically in the gel. Figures A+C (12% polyacrylamide gel) and B+D (16% polyacrylamide gel) are related and share line markings, as two identical gels were prepared, one was stained by Coomassie and the second was blotted and visualized by anti-His iBody.

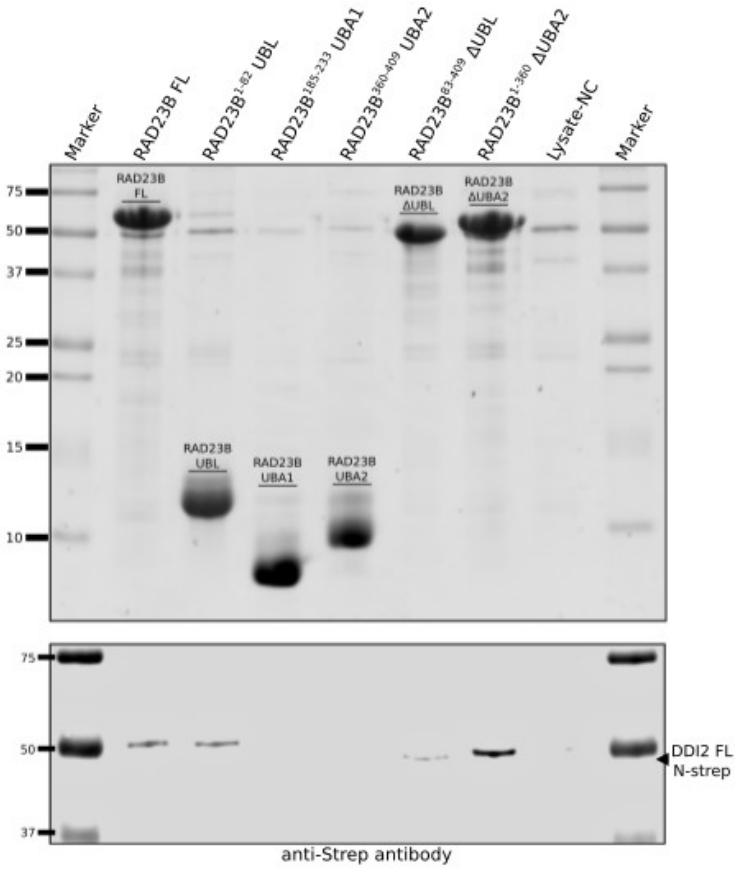

Figure 22: DDI2-RAD23B interaction mapping performed by the pull-down assays using NiNTA resin with RAD23B protein variants as a bait and DDI2 as a prey. 1.5 µl of molecular weigth marker and 2 µl of the protein samples were loaded into the gel. Bait proteins for each sample are noted at the top and the relevant bait protein bands are marked specifically in the gel. 16% polyacrylamide gel was used, blots were visualized using anti-strep antibody. Representative picture of two independent experiments.

5.5. NMR experiments

5.5.1. Sequence assignment

The $^{15}N/^{1}H$ HSQC spectrum of the $^{13}C/^{15}N$ labelled $DDI2^{1-212}$ UBL-HDD protein variant was measured and compared with the $^{15}N/^{1}H$ HSQC spectra of the DDI2 UBL domain and the HDD domain, previously solved by NMR in our laboratory[1]. Overlay of the three $^{15}N/^{1}H$ HSQC spectra is shown in Figure 23 (page 68). The $^{15}N/^{13}C/^{1}H$ HNCO, HN(CA)CO, HNCACB and CACB(CO)NH spectra of the $DDI2^{1-212}$ UBL-HDD were collected and used for assignment of the $^{15}N/^{1}H$ HSQC spectral peaks to individual amino acid residues according to chapter 4.6.1. (page 57). The sequence of the $DDI2^{1-212}$ UBL-HDD was partially assigned to the spectral peaks and the result was used for the following NMR titrations evaluation.

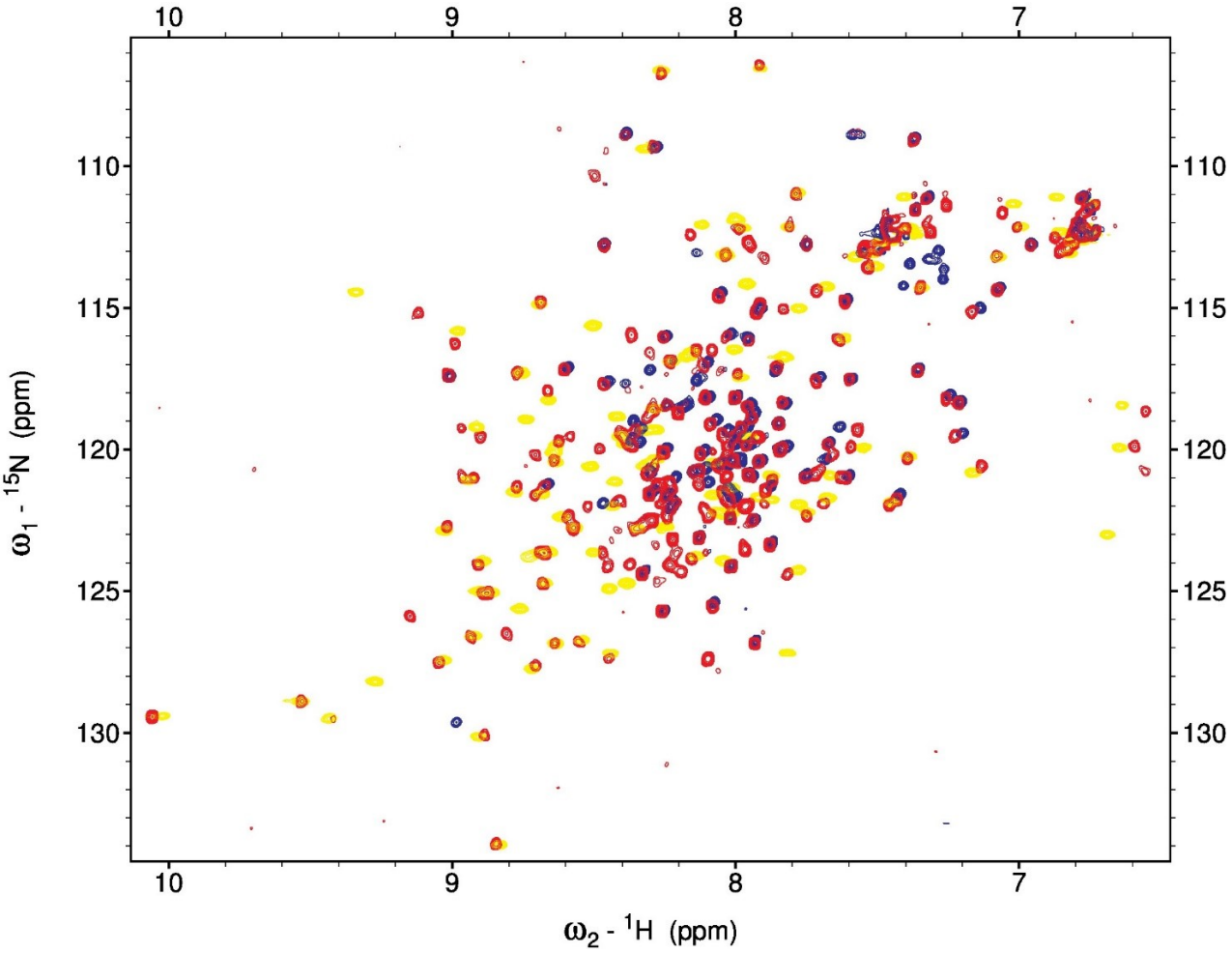

Figure 23: Obtained ^{15}N/^{1}H HSQC spectrum of the DDI2$^{1-212}$ UBL-HDD protein variant overlaid with the ^{15}N/^{1}H HSQC spectra of the DDI2 UBL domain (PDB code 2N7D) and the HDD domain (PDB code 5K57) published previously[1]. Red – spectrum of the DDI2$^{1-212}$ UBL-HDD protein; yellow – spectrum of the DDI2 UBL domain; blue – spectrum of the DDI2 HDD domain.

5.5.2. The DDI2$^{1-212}$ UBL-HDD titration by the RAD23B$^{1-82}$ UBL domain

The ^{15}N/^{1}H HSQC spectra of the DDI2$^{1-212}$ UBL-HDD supplemented with the RAD23B$^{1-82}$ UBL protein variant were recorded. Molar ratios 1:0, 1:0.5, 1:1, 1:2, 1:4 and 1:8 of the DDI2$^{1-212}$ UBL-HDD to RAD23B$^{1-82}$ UBL were used for the measurement. The titration was evaluated according to chapter 4.6.2. (page 57). Overlay of the spectra from the RAD23B$^{1-82}$ UBL titration are shown in Figure 24 (page 69). The shifts of the peak positions for 8-fold molar excess of RAD23B$^{1-82}$ UBL and the changes of relative peak intensity (RPI) for the RAD23B$^{1-82}$ UBL titration are summarised in Figure 25 (page 70). The K_d obtained for the DDI2$^{1-212}$ UBL-HDD interaction with the RAD23B$^{1-82}$ UBL is 234 ± 10 μM.

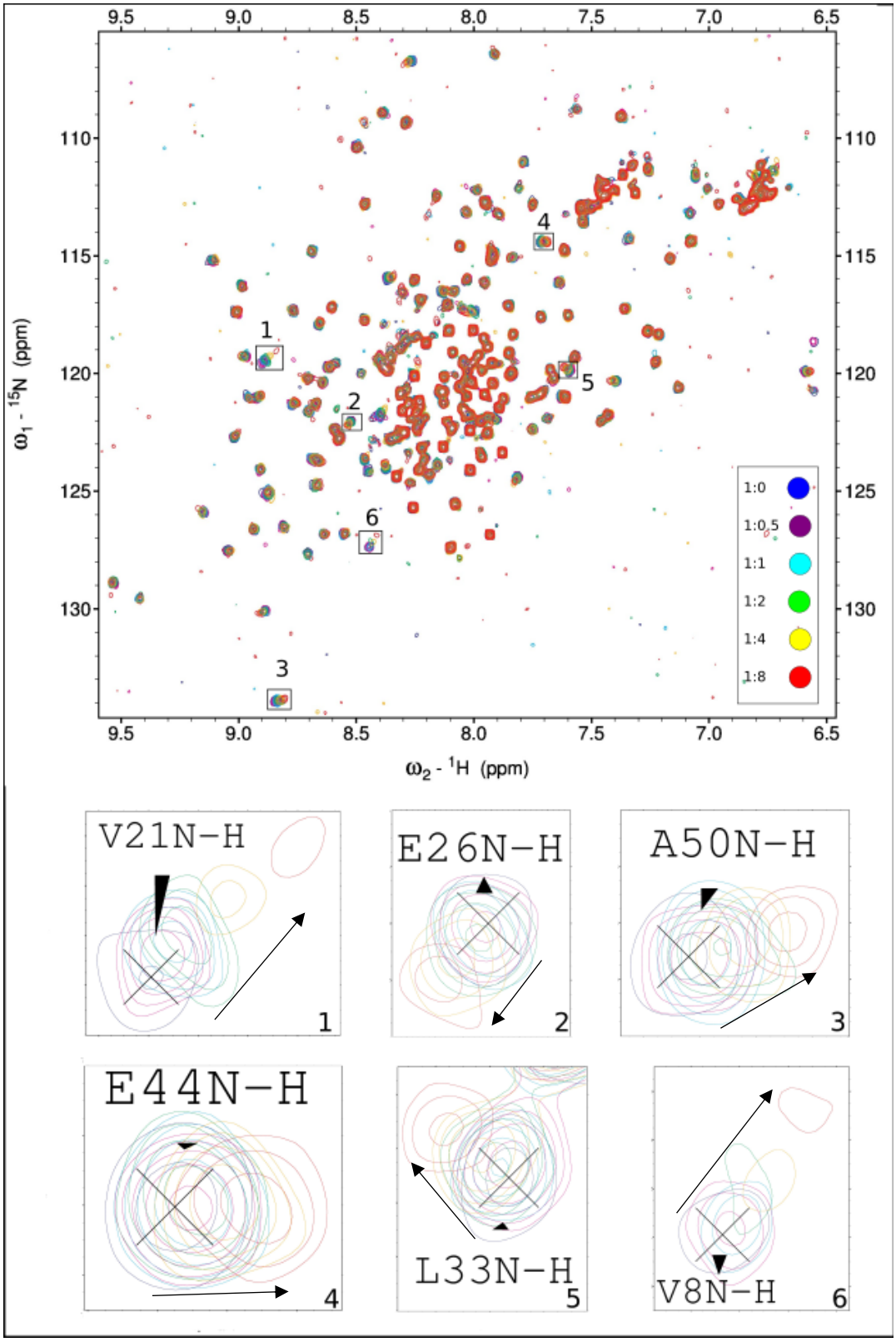

Figure 24: The spectra from the titration of 25 µM DDI2$^{1-212}$ UBL-HDD with RAD23B$^{1-82}$ UBL. The DDI2$^{1-212}$ UBL-HDD spectra measured with 1:0, 1:0.5, 1:1, 1:2-fold, 1:4 and 1:8 molar ratios (DDI2$^{1-212}$ UBL-HDD : RAD23B$^{1-82}$ UBL) were overlayed and the perturbations observed. Selected peak perturbations are shown in detail, with arrow marking peak movement direction and with position of the peaks indicated in the spectrum.

Figure 25: Evaluation of 25 µM DDI2$^{1-212}$ UBL-HDD titration with RAD23B$^{1-82}$ UBL for individual residues. Residues unnasigned or not included in the analysis are marked with red cross. Amino acid residues include N-terminal Strep-tag (residues with positions less than 1). **A)** Evaluation of the chemical shift perturbations (CSP) for 1:8 molar ratio spectrum. Yellow – CSPs at least one standard deviation higher than the average, Orange - CSPs at least two standard deviations higher than the average. Gray – disappearing peaks. **B)** Evaluation of relative peak intensity (RPI) changes in relation to RAD23B$^{1-82}$ UBL concentration. Blue - 1:0.5 molar ratio spectrum. Yellow - 1:2 molar ratio spectrum. Red - 1:8 molar ratio spectrum.

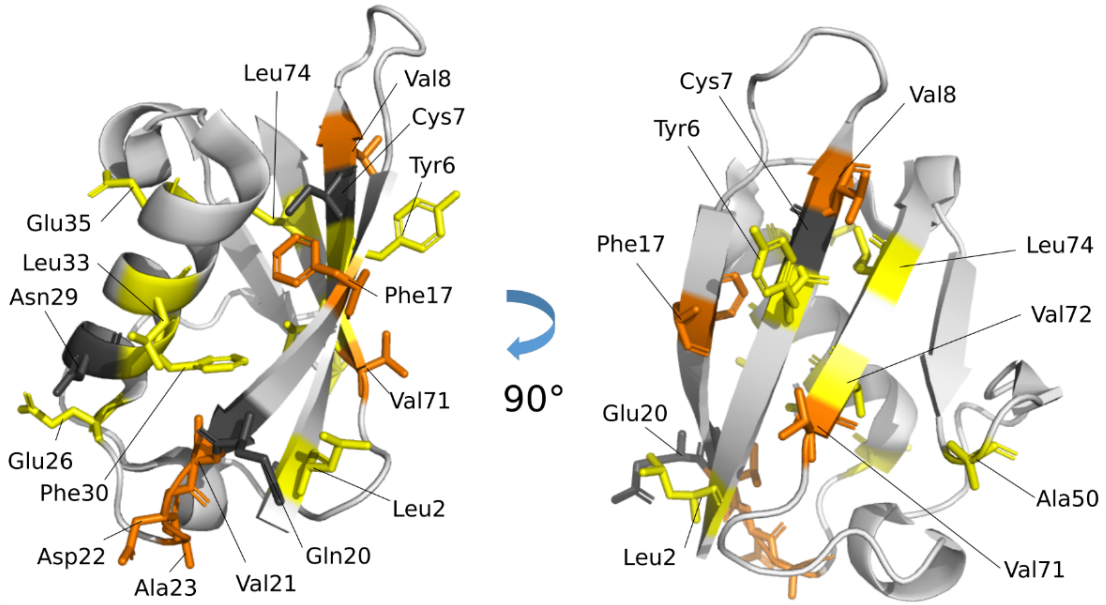

Figure 26: Structure of the UBL domain of the human DDI2 (PDB 2N7D)[1]. Residues with observed significant CSPs for the RAD23B$^{1-82}$ UBL titration are colored, with color scheme corresponding to the graph in Figure 25A (page 70): Yellow – CSPs of at least one standard deviation higher than the average, Orange - CSPs of at least two standard deviations higher than the average. Dark gray – residues with disappearing peaks. The picture was created using PyMol software (Schrodinger, LLC).

Next, the interaction of the DDI2$^{1-212}$ UBL-HDD with RAD23B$^{1-82}$ UBL was mapped. All significant CSPs were detected within the UBL domain of DDI2. Residues with the most significant shifts are Val8, Phe17, and Val71 located within the β-sheet of the UBL domain and three subsequent residues Val21, Asp22 and Ala23 that form a flexible region adjoining the β-sheet. Peaks of Cys7, Gln20 and Asn29 residues disappeared in the courese of titration. Further perturbations of lesser extent were observed in the area of the β-sheet and the opposing α-helix. Positions of the residues with significant CSPs in the DDI2 UBL upon RAD23B UBL binding are shown in Figure 26.

5.5.3. The DDI2$^{1-212}$ UBL-HDD titration by the RAD23B FL

The ^{15}N/^{1}H HSQC spectra of the DDI2$^{1-212}$ UBL-HDD supplemented with RAD23B FL protein variant were recorded using molar ratios 1:0, 1:0.5, 1:1, 1:2, 1:4 and 1:8 of the DDI2$^{1-212}$ UBL-HDD to RAD23B FL. The titration was evaluated according to chapter 4.6.2. (page 58). Overlay of the spectra from the RAD23B FL titration are shown in Figure 27 (page 73). The shifts of the peak positions for the 8-fold molar excess of the RAD23B FL and the changes of relative peak intensity (RPI) for RAD23B FL titration are summarised in Figure 28 (page 74). The K_d obtained for the DDI2$^{1-212}$ UBL-HDD interaction with the RAD23B FL is 23 ± 3 μM.

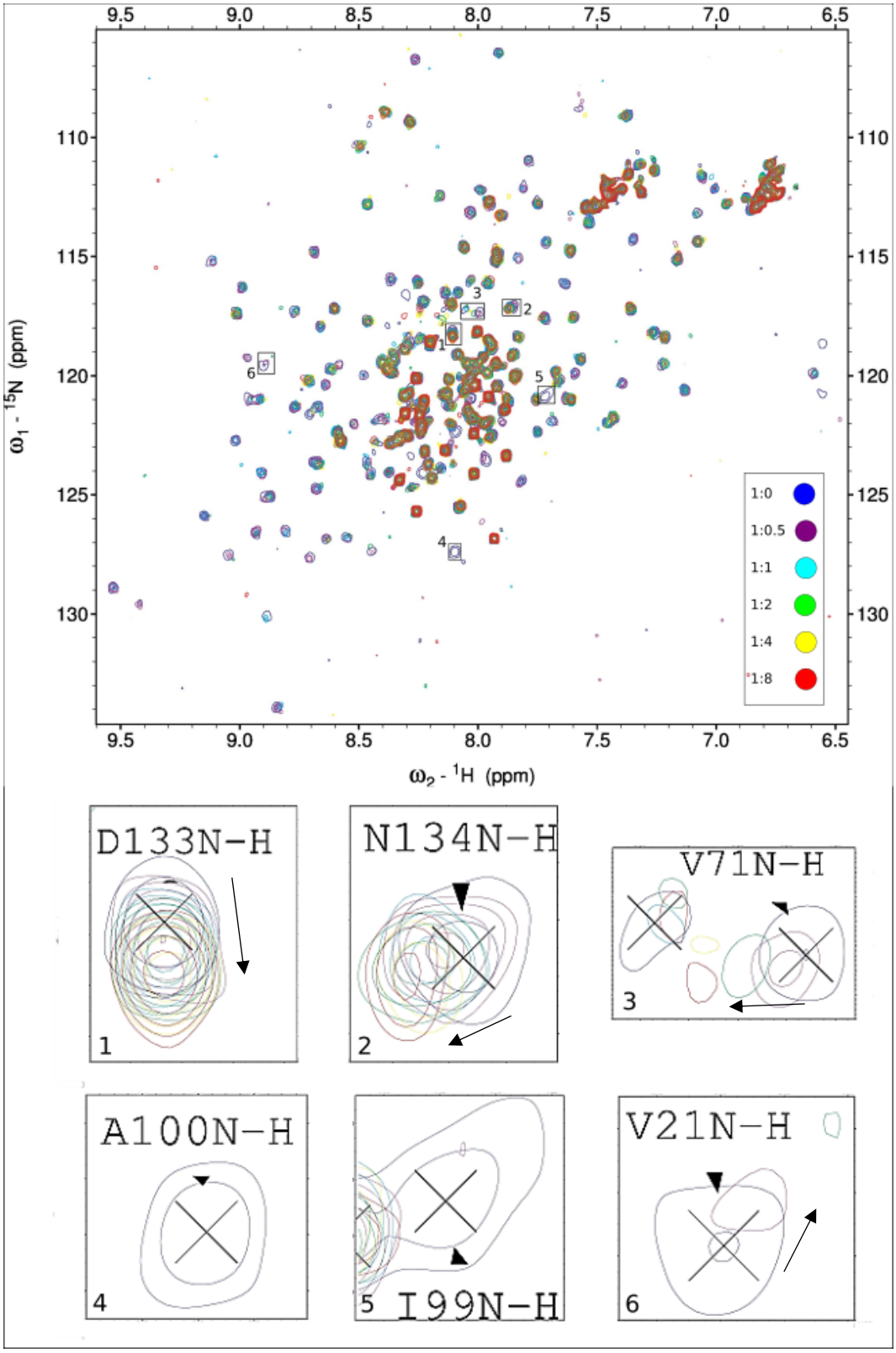

Figure 27: The spectra of the titration of 25 µM DDI2$^{1-212}$ UBL-HDD with RAD23B FL. The DDI2$^{1-212}$ UBL-HDD spectra were measured with 1:0, 1:0.5, 1:1, 1:2-fold, 1:4 and 1:8 molar ratios (DDI2$^{1-212}$ UBL-HDD : RAD23B FL). Selected peak perturbations are shown in detail, with arrow marking peak movement direction and with position of the peaks indicated in the spectrum.

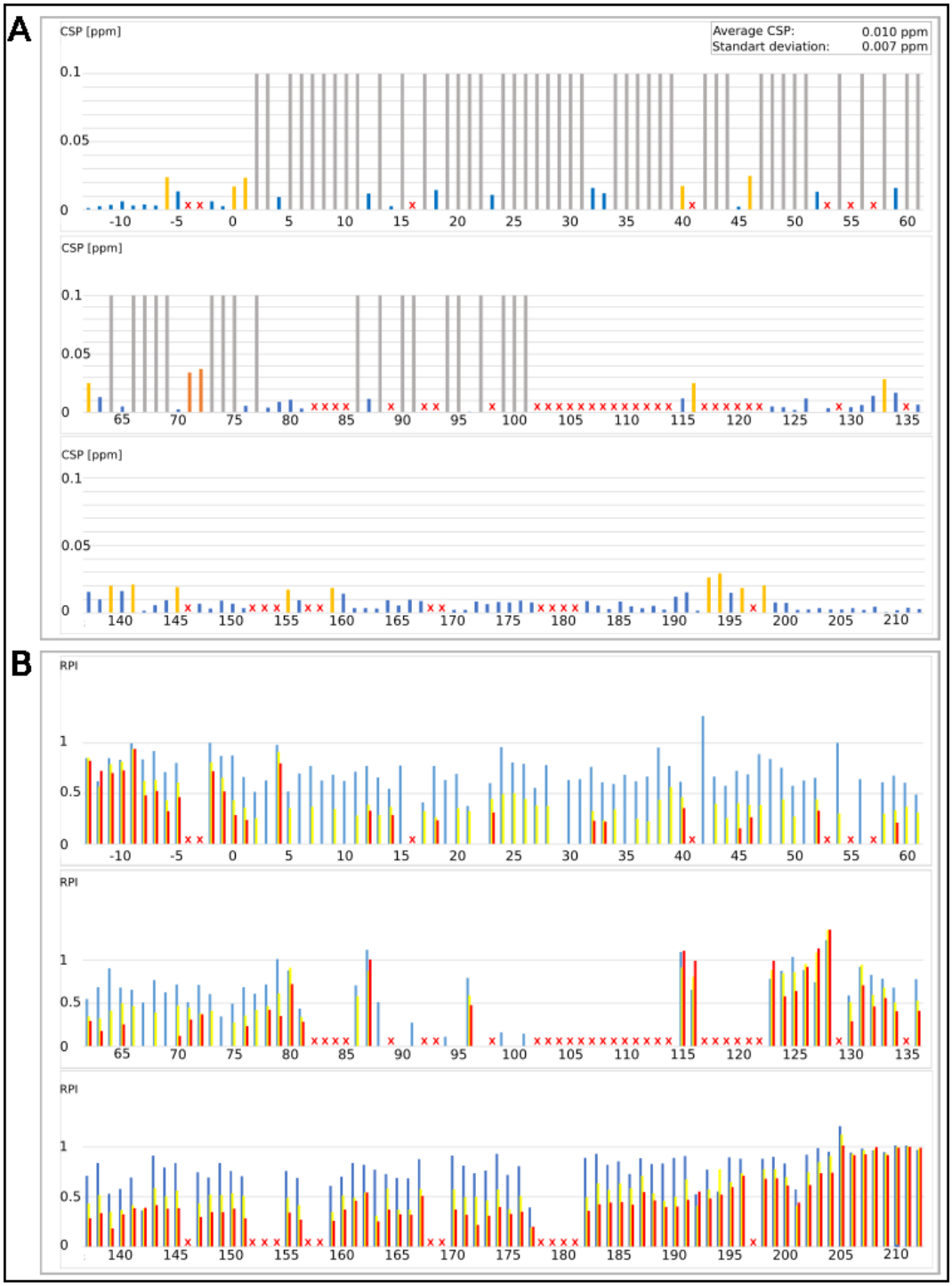

Figure 28: Evaluation of 25 µM DDI2$^{1-212}$ UBL-HDD titration with RAD23B FL for individual residues. Residues unnasigned or not included in the analysis are marked with red cross. Amino acid residues include N-terminal Strep-tag (residues with positions less than 1). **A)** Evaluation of chemical shift perturbations (CSP) for 1:8 molar ratio spectrum. Yellow – CSPs at least one standard deviation higher than the average, Orange - CSPs at least two standard deviations higher than the average. Gray – disappearing peaks. **B)** Evaluation of relative peak intensity (RPI) changes in relation to RAD23B$^{1-82}$ UBL concentration. Blue - 1:0.5 molar ratio spectrum. Yellow - 1:2 molar ratio spectrum. Red - 1:8 molar ratio spectrum.

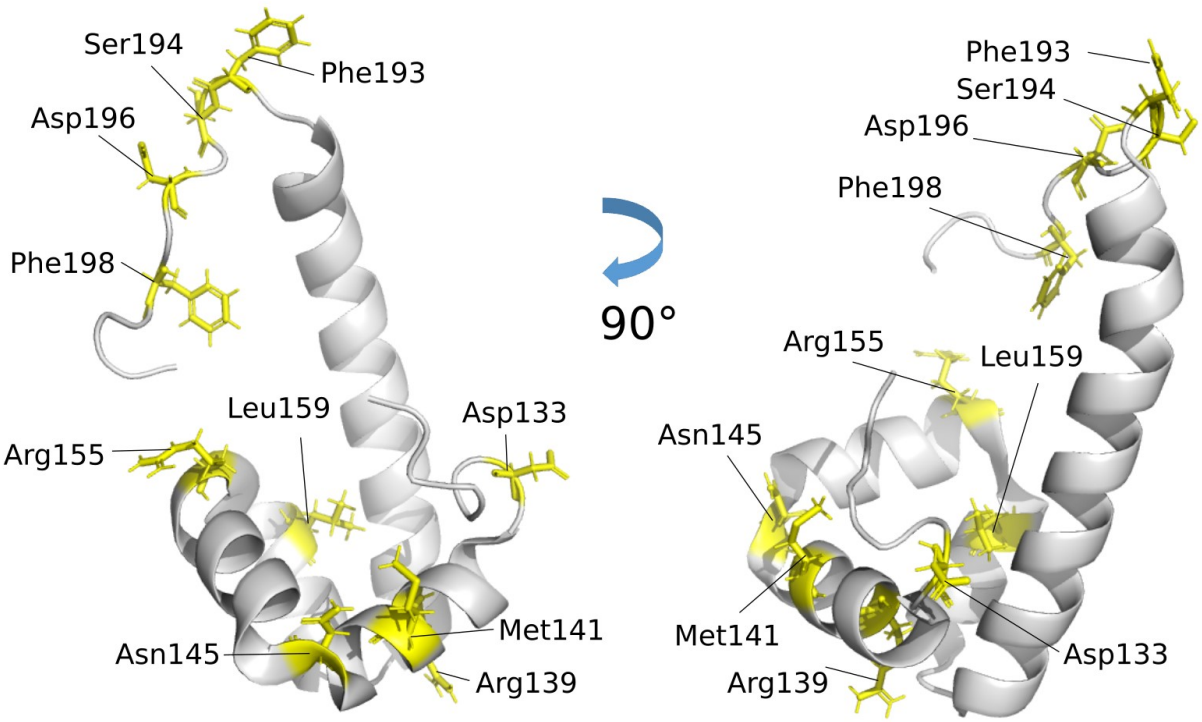

Figure 29: Structure of the HDD domain of the human DDI2 (PDB 5K57)[1] from two different angles. Residues with observed significant CSPs for the RAD23B FL titration are colored, with color scheme corresponding to the graph in Figure 28A (page 73): Yellow – CSPs of at least one standard deviation higher than the average. No disappeared signals or signals with CSP higher than two standard deviations over the average were observed within the HDD domain. The picture was created using PyMol software (Schrodinger, LLC).

Significant relative peak intensity (RPI) changes and disappearing peaks were observed for the signals assigned to the DDI2 UBL domain and to the linker region between UBL and HDD. The CSPs were not mapped on the structure of the DDI2 UBL as for the previous titration due to disappearing peaks. Several significant shifts corresponding to Asp133, Arg139, Met141, Asn145, Arg155, Leu159, Phe193, Ser194, Asp196 and Phe198 residues were detected also within the HDD domain. Positions of residues with significant CSPs in the DDI2 HDD structure (PDB 5K57)[1] are shown in Figure 29.

6 Discussion

6.1. Protein purification

The DDI2 protein variants including DDI2 FL C-Strep, DDI2 FL N-Strep, DDI2$^{1-212}$ UBL-HDD and the RAD23B protein variants RAD23B FL and RAD23B$^{1-82}$ UBL were purified using affinity chromatography. For their use as prey-proteins in the pull-down experiments, the purity from this single purification step was considered sufficient. For the NMR titration experiments, gel filtration was employed as the second purification step. Therefore, the DDI2$^{1-212}$ UBL-HDD, RAD23B FL and RAD23B$^{1-82}$ UBL protein variants were purified this way.

The DDI2$^{1-212}$ UBL-HDD protein was purified using Superdex 75pg 16/60 column. This protein of approx. 26 kDa was expected to be eluted within 70-80 ml of the mobile phase. The corresponding peak was detected, yet a second peak was eluted at approximately 42 ml of the mobile phase, maching the dead volume of the used chromatography system. SDS-PAGE analysis of the purification revealed, that both peaks were formed by the DDI2$^{1-212}$ UBL-HDD. The first peak consisted of DDI2$^{1-212}$ UBL-HDD along with impurities of higher molecular weight, and was not used further as the protein was likely denaturated due to the low elution volume. The second peak eluted at approximately 72 ml was found to be of high purity and was used for NMR experiments.

The RAD23B$^{1-82}$ UBL protein was purified using Superdex 75pg 16/60 column. This protein of approx. 12 kDa was expected to be eluted within 80-90 ml of the mobile phase. The correspondig peak was detected at elution volume of approximately 86 ml. It was preceded by a smaller peak eluted at approximately 74 ml of the mobile phase, likely corresponding to a protein of approximately 25 kDa. SDS-PAGE analysis of the purification procedure revealed, that both peaks corresponded to the RAD23B$^{1-82}$ UBL protein. This indicated, that RAD23B$^{1-82}$ UBL protein partially formes a dimer. While no reference for this behavior was found in the literature for the UBL domain of the human RAD23B protein, the UBL domain of RAD23A protein was described to form weak dimers mediated by hydrophobic residues Leu10, Ile49 and Met75[233]. Considering the high sequence identity of the two homologous domains with all the three residues mediating the dimerisation conserved, it is very likely that the two domains are behaving in the same manner. Detected dimer is probably a result of high concentration (9 mg/ml, approx. 750 µM) of RAD23B$^{1-82}$ UBL loaded into the column.

The RAD23B FL protein was purified using Superdex 200pg 16/60 column. This protein of approx. 46 kDa was expected to be eluted within 70-80 ml of the mobile phase. Unexpectedly, three major peaks were detected during gel chromatography purification, located at 60 ml, 69 ml and 73ml. SDS-PAGE analysis revealed, that all three peaks consisted predominantly of the RAD23B FL protein. This might suggest that RAD23B FL forms tetramers (60 ml peak), dimers (69 ml peak) and monomers (73 ml peak). No previous reports regarding oligomerisation behavior of the full-length human RAD23B protein were found in the literature, only the yeast Rad23 ortholog is known to form dimers[149]. RAD23A was explicitly reported to not form multimers upon investigation by NMR. Unfortunately, the authors are not stating clearly at which concentration it was measured[175]. Even so, RAD23B behavior can be different. An older article reported strikingly different behavior of RAD23A and RAD23B on native electrophoresis, with RAD23A migrating at 70kDa, while RAD23B was migrating at 140 kDa[234]. Apparent size of 140 kDa would correlate well with the peak eluted at 60 ml volume. Apart from posible multimerisation, our data might relate to a conformational behavior of RAD23A/B, as the proteins are not globular and consist of small domains connected by flexible linkers[175]. To explore this behaviour, further gel chromatography experiments can be performed, monitoring the behaviour of the three major peaks in relation to RAD23B FL loading concentration, and using truncated RAD23B variants to probe the effect of RAD23B domains on possible oligomerisation. For purposes of this thesis, the purity of fractions B11-C6 was found sufficient and they were used further for NMR titration experiments.

6.2. Protein-protein interaction pull-down assay

Pull-down experiments with truncated protein variants were used to map the interaction of RAD23B and DDI2. In yeast, the interaction of Rad23 and Ddi1 proteins was described to be mediated by the C-terminal UBA domain of Ddi1 binding to the Rad23 UBL domain and by the internal UBA1 domain of Rad23 binding to the Ddi1 UBL domain[149]. While domain organization of Rad23 is conserved in its mammalian orthologs, the UBA domain of Ddi1 was lost in vertebrates[1]. Potential candidates for replacing the role of the missing UBA domain were considered the putative UIM motif at the C-terminus of the human DDI2, possibly binding to the UBL domain of RAD23B, or the UBL domain of DDI2 that could interact with one of the RAD23B UBA domains.

Regarding RAD23B truncation, the constructs encoding the full-length protein and the UBL, UBA1 and UBA2 domains were cloned to explore whether they can bind the

DDI2 protein. To complement these variants, deletions of N-terminal UBL and C-terminal UBA2 domains were prepared, as internal domain deletions were more complicated to prepare using classical restriction cloning.

For DDI2 DNA constructs preparation, a more flexible method of Gibson assembly was used to clone the variants with deletions. The DDI2 FL N-Strep, DDI2$^{212-399}$ RVP-UIM and DDI2$^{1-360}$ ΔUIM with N-terminal Strep-tag were cloned to explore the role of the C-terminal UIM sequence in the interaction. The DDI2 FL C-Strep, DDI2$^{1-81}$ UBL and DDI2$^{82-399}$ ΔUBL with C-terminal Strep-tag were cloned to explore the role of the N-terminal UBL domain in the interaction. The following constructs DDI2 $^{\Delta 116-212}$ ΔHDD, DDI2 HDD$^{116-212}$ and DDI2 UBL-HDD$^{1-212}$ were cloned after the roles of the UIM and UBL domains were examined.

Several independent pull-down experiments were performed, from which only representative results are shown in the Results section. UIM motif was found to be uninvolved in the interaction, with DDI2 FL N-Strep and DDI2$^{1-360}$ ΔUIM variants binding to the RAD23B FL, but DDI2$^{212-399}$ RVP-UIM not binding (Figure 21 A+C, page 66). Similarly, DDI2$^{1-81}$ UBL was not binding the RAD23B FL, with DDI2 FL C-Strep and DDI2$^{82-399}$ ΔUBL both interacting, without any effect of the UBL deletion observed. Further examination revealed no interaction for the HDD domain and unaffected interaction for the DDI2$^{\Delta 116-212}$ ΔHDD variant. The minimal variant consistently being able to bind RAD23B FL was the DDI2$^{1-212}$ UBL-HDD (Figure 21 B+D, page 66).

Regarding RAD23B, UBL domain was found to be able to bind to the DDI2 protein. On the other hand, RAD23B$^{83-409}$ ΔUBL was still able to bind the DDI2 protein, which indicated the presence of another interaction site within the RAD23B protein. However, I was unable to detect any interaction for RAD23B$^{185-233}$ UBA1 or RAD23B$^{361-409}$ UBA2 variants, while RAD23B$^{1-360}$ ΔUBA2 was binding to the DDI2 protein (Figure 22, page 67).

Thus, the information about the interaction site on RAD23B remains incomplete. The UBL domain of RAD23B participates in the interaction but it is not the only interaction site. For the UBA1 and UBA2 domains, it is uncertain whether they do not interact, or whether they were structurally impaired when expressed in the designed truncated form. With ΔUBA1 variant missing, the role of UBA1 is unclear. While ΔUBA2 variant was still binding to the DDI2 protein, suggesting no involvement of UBA2 domain, the interaction could have been mediated by the present UBL domain of RAD23B. No constructs for elucidation of the role of the XPC-binding domain were cloned. To

overcome these gaps, variants with double deletions of ΔUBLΔUBA1, ΔUBLΔUBA2 and ΔUBLΔXPCb should be included in the tested set of variants in the future.

The minimal DDI2 protein variant binding RAD23B is $DDI2^{1-212}$ UBL-HDD, while the UBL or the HDD domains alone did not preserve the binding. This suggested some kind of cooperation between the UBL and the HDD domains during RAD23B binding or, less likely, a novel binding motif located in the linker region.

Two available options in this stage were either designing and cloning new DDI2 variants to explore the role of the previously omitted $DDI2^{82-116}$ linker region, or proceeding to NMR analysis of the whole $DDI2^{1-212}$ UBL-HDD variant. After consulting Dr. Veverka and Dr. Srb on the plausibility and potential drawbacks of NMR analysis of the 26 kDa $DDI2^{1-212}$ UBL-HDD, it was decided to proceed to NMR.

6.3. NMR experiments

6.3.1. Sequence assignment of the $DDI2^{1-212}$ UBL-HDD protein variant

Sequence assignment of the $DDI2^{1-212}$ UBL-HDD variant was complicated by several factors, including low quality of collected 3D spectra. While HNCO and CBCA(CO)NH experiments provided well resolved signal peaks of a sufficient intensity, the HNCACB spectra and the HN(CA)CO spectra provided low quality signals with high levels of noise. The HN(CA)CO spectra were mostly signals from the (i) position, rendering whole complementary pair of HN(CA)CO + HNCO spectra unusable, as only signals for (i-1) position were visible. The HNCACB spectrum was of low quality but usable, as signals from (i-1) position were noisy or missing, but signals from (i) position were mostly present. As signals of the (i-1) position were present in complementary CBCA(CO)NH spectra, the obtained information was sufficient for following the signals along the sequence (see Figure 11, page 57 for the assignment strategy description). This difference in intensities is understandable, as the magnetisation has to be transferred over longer distances for aquiring (i-1) position signals in HNCACB spectrum and (i) position in HN(CA)CO spectrum, resulting in lower intensity of these signals. In combination with the size of the protein construct which further contributed by overall lowering the intensities of peaks by signal broadening, used time of measurement was insufficient for these peaks to surpass the noise level.

Next factors complicating the signal assignment were also resulting from the large size of the used protein variant, namely sequence repetitions, frequency degeneracies and spectral crowding. Spectral crowding stems from high number of signals confined within

measured spectral range, resulting in peak overlaps that distort peaks or cause low intensity peaks to be concealed by more intensive ones. The sequence repetitions don't allow assigned sequence of signals to be uniquely attributed to a single spot within the sequence, and increase the likelihood of frequency degeneracy. Frequency degeneracy is a situation, where more than one amino acid presents peaks of particular frequencies, making such residues hard to be distinguished from one another in the spectra[235].

These factors contributed to the fact that the complete sequence of the DDI2 UBL-HDD$^{1-212}$ was not assigned to signals in the $^{15}N/^{1}H$ HSQC. However, through comparison with previously solved spectra of the DDI2 UBL domain and the DDI2 HDD domain[1], the most of the UBL domain and HDD domain sequence was assigned. Further, with decreased number of unassigned signals, part of the linker region connecting the two domains was assigned.

6.3.2. The DDI2$^{1-212}$ UBL-HDD titration by the RAD23B UBL domain

The interaction of the DDI2$^{1-212}$ UBL-HDD with the RAD23B$^{1-82}$ UBL was succesfully mapped, suggesting that this interaction is mediated by the UBL domain of the DDI2, as a large majority of the observed CSPs was located within the DDI2 UBL (Figure 25A, page 70). Almost no CSPs were observed in the linker region or the HDD domain. Furthemore, a consistent decrease in RPIs was observed for the UBL domain residues (Figure 25B, page 70). Observed shifts included the hydrophobic residues on the surface of the β-sheet, the residues directed into the hydrophobic core of the domain and the residues of the α-helix opposing the β-sheet (Figure 26, page 71).

The shifts in the hydrophobic core probably arrise from a conformational change of the domain induced by the interaction. Ubiquitin and other UBL domains interact often through the hydrophobic patch located within the β-sheet. Thus a model can be suggested where the intermolecular contacts with the RAD23B UBL domain are made by the hydrophobic residues exposed on the β-sheet of the DDI2 UBL. This results in a conformational change of the β-sheet, influencing the α-helix through the residues of the hydrophobic core that hold the β-sheet and the α-helix together. However, this hypothesis would need further verification.

The K_d of the DDI2$^{1-212}$ UBL-HDD with RAD23B$^{1-82}$ UBL interaction was determined to 234 ± 10 μM. This value is however likely affected by a significant error, as the maximum concentration of 200 μM RAD23B$^{1-82}$ UBL was used for the titration.

Higher concentration of RAD23B$^{1-82}$ UBL would thus need to be used to get closer to a saturated state for accurate K_d calculation.

Previously published interaction of the human DDI2 UBL domain with human ubiquitin shows several similarities to the interaction with the human RAD23 UBL domain[1]. The K_d determined for the interaction of DDI2 UBL with ubiquitin is 0.42-1.1 mM. Significant shifts of the β-sheet residues, in the adjoining flexible region and in the α-helix were detected. This suggests a very similar mode of interaction employing the β-sheet accompanied by a conformational change. Residues Cys7, Val8, Phe17, Val21, Ala23, Phe25, Phe30 and Leu74 were found to be perturbed in both interactions, with other residues perturbed in only one of the interactions, as can be seen in Figure 30. This shows that while different residues might be involved in the interaction, the interaction surfaces spatially overlap, omitting almost certainly a simultaneous binding of both ubiquitin and RAD23 UBL to the DDI2 UBL. If this has some biological consequences, it remains to be determined.

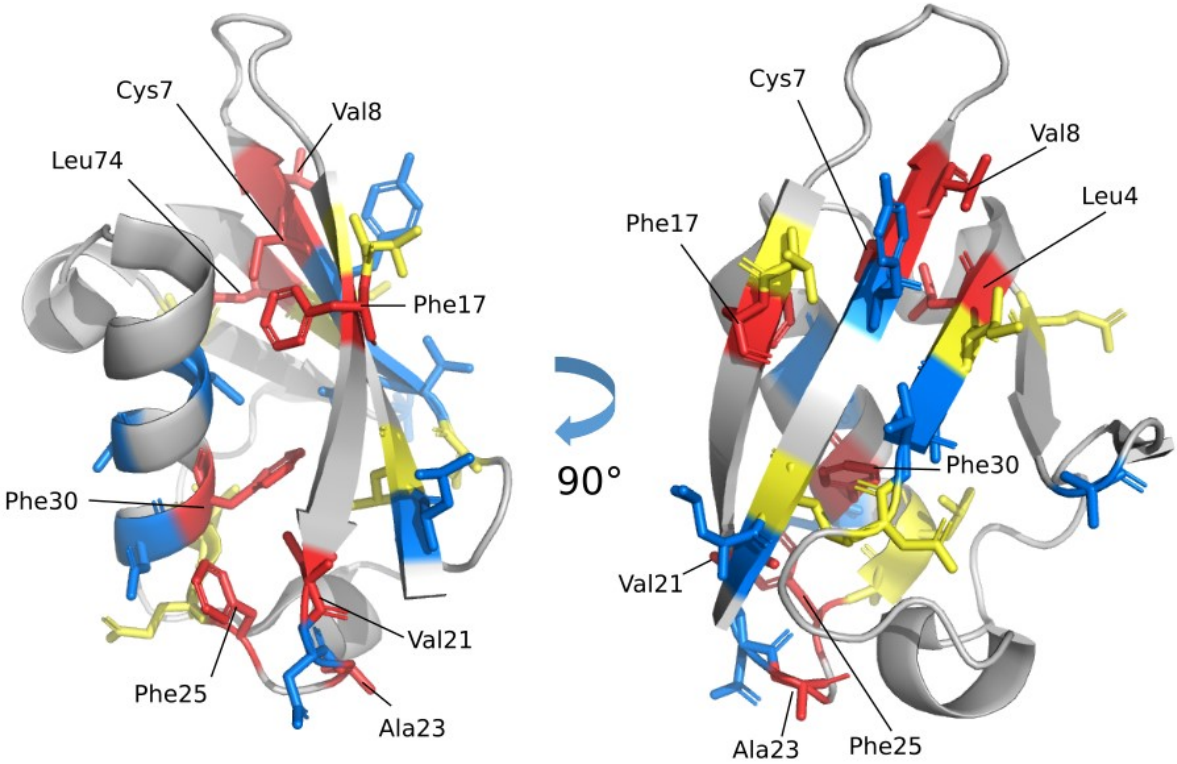

Figure 30: Comparison of involvement of DDI2 UBL (PDB 2N7D) residues in interaction with ubiquitin[1] and with RAD23B$^{1-82}$ UBL. Red – residues with significant perturbations detected for both interactions. Blue – residues with significant perturbations detected only for interaction with RAD23B$^{1-82}$ UBL. Yellow - residues with significant perturbations detected only for interaction with ubiquitin[1]. The picture was created using PyMol software (Schrodinger, LLC).

More information could be obtained by performing inversed titration, with ^{15}N labelled RAD23B UBL being titrated by the $DDI2^{1-212}$ UBL-HDD or $DDI2^{1-81}$ UBL, possibly followed by molecular docking simulation. That, however, is out of the scope of this thesis.

6.3.3. The $DDI2^{1-212}$ UBL-HDD titration by The RAD23B FL

Considering the results of the pull-down experiments, where the interaction was detected even without the presence of the UBL domain of the DDI2 (the $DDI2^{82-399}$ ΔUBL variant), it was suggested that other interaction site must be present within the $DDI2^{1-212}$ UBL-HDD variant. Therefore, we performed titration of the ^{15}N labelled $DDI2^{1-212}$ UBL-HDD with the full-length RAD23B.

Far the most pronounced effect of this titration was a decrease and progressive disappearing of numerous signals observed in the spectra (Figure 27, page 72). This was interpreted as a conclusive sign of the interaction, resulting from a large size (46 kDa) of the RAD23B FL binding partner. Decrease in signal intensity is an effect of an increase of rotational correlation time of the $DDI2^{1-212}$ UBL-HDD as the complex with the large interaction partner is formed. Importantly, the signal attenuation wasn't spread globally, but was concentrated in the UBL domain signals and the assigned linker residues, while decline of the signals of the HDD domain was milder (Figure 28A, page 73). This suggests that while UBL domain became fixed to the large RAD23B FL, the HDD domain retained some rotational independence from the formed complex.

In accordance with this, the observed shifts in the HDD domain, although more significant than in the case of the RAD23B UBL titration (compare Figure 25A, page 70 and 28A, page 73) , do not show any presence of an important binding site. Shifts are spread across the domain but, interestingly, often localised on the junctions of individual helices, so mild rearrangement of the helices during the interaction is possible (Figure 29, page 74).

The widespread disappearing of the signals of the UBL domain prevented the interaction site mapping to be done the same way as for the $RAD23B^{1-81}$ UBL titration (compare Figure 25A, page 70 and 28A, page 73). It is probable that DDI2 UBL and RAD23B UBL domains interact similarly even in presence of the full-length RAD23B. A possibility of additional contacts of the DDI2 UBL with other parts of the RAD23B protein can't however be excluded.

Important information also comes from the behavior of the assigned peaks in the linker region connecting UBL and HDD domains (residues 85-101). While in the RAD23B UBL titration these peaks remained unperturbed, in the RAD23B titration they mostly disappeared (compare Figure 25, page 70 and 28, page 73). While this might be an effect of the nearby UBL domain being immobilised, given the results of the pull-down experiments, it is tempting to hypothesise that this might be additional binding site for the RAD23B protein. The pull-down experiments indicate that other binding site apart from the UBL domain is present within the $DDI2^{1-212}$ UBL-HDD. Titration experiments had shown no significant involvement of the HDD domain in the interaction, meaning that only remaining site available is the linker region. Also, while gradual attenuation of most of the UBL domain signals was observed, the linker signals mostly disappeared already at the 0.5-fold molar addition of the RAD23B FL (Figure 28B, page 73).

In accordance with this, the K_d of the interaction of the $DDI2^{1-212}$ UBL-HDD with RAD23B is 23 ± 3 μM, significantly stronger than for the interaction with $RAD23B^{1-81}$ UBL.

6.3.4. Summary and future prospects

The DDI2 protein interacts with the RAD23B protein via its N-terminal part. The UBL domain of DDI2 was found to interact with the UBL domain of RAD23B. The data also indicate that the linker region connecting the UBL and the HDD domains of DDI2 participates the interaction with RAD23B. Similarly, an additional RAD23B binding site outside of RAD23B UBL domain was suggested.

The fact that DDI2 UBL acts as an interaction surface for another UBL domain is quite unusual among UBLs. It is, however, not so surprising given that the yeast Ddi1 UBL and the human DDI2 UBL interacts with ubiquitin. Nevertheless, this further sets the UBL domains of Ddi1-like proteins as rather atypical members of the UBL domain family.

More experiments need to be performed in order to fully describe the interaction of DDI2 with RAD23B. The DDI2 protein variants carrying mutations in the linker region might help to characterize the importance of this sequence for the interaction. Similarly, a combination of deletion of RAD23B UBL domain with deletions of other parts of RAD23B can be used to observe their effect on the interaction using a pull-down assay. RAD23B UBL domain and potentially other interacting site (if discovered) can be then analysed using NMR titration experiments.

Identification of the binding sites on both proteins is important for the following biological experiments elucidating the role of DDI2-RAD23B interaction in activation of NRF1 transcription factor. So far, the data from the cooperating lab indicate that cells depleted of RAD23B have decreased ability to activate NRF1. To further investigate the biological significance of the interaction for NRF1 activation, different DDI2 and RAD23B variants can be cloned into the plasmids for mammalian expression and their ability to restore NRF1 activiation upon expression can be tested. If the constucts with disrupted interaction sites fail to fully restore NRF1 activation, it will implicate the importance of DDI2-RAD23B interaction for this process.

7 Conclusion

1) The state of knowledge regarding the human DDI2 protein incorporated in the wider context has been provided.

2) The interaction between the human DDI2 and the human protein RAD23B has been characterized:
    a. by pull-down experiments with recombinant DDI2 and RAD23B and the interaction was mapped
    b. by protein NMR titration analysis revealing the interaction interface between the UBL domains of both proteins and suggesting the importance of the linker region (connecting DDI2 UBL-HDD) for the interaction

3) Obtained results were discussed with the literature and directions for further research have been suggested.

8 Bibliography

1. Sivá, M. *et al.* Human DNA-Damage-Inducible 2 Protein Is Structurally and Functionally Distinct from Its Yeast Ortholog. *Sci. Rep.* **6**, 30443 (2016).
2. Kottemann, M. C., Conti, B. A., Lach, F. P. & Smogorzewska, A. Removal of RTF2 from Stalled Replisomes Promotes Maintenance of Genome Integrity. *Mol. Cell* **69**, 24-35.e5 (2018).
3. Koizumi, S. *et al.* The aspartyl protease DDI2 activates Nrfl to compensate for proteasome dysfunction. *Elife* **5**, (2016).
4. Belza, J. *Dna Damage – Inducible Protein : Interakce s proteinovými partnery a potenciální role v proteasomální degradaci proteinů*. (2016).
5. Ibarra-Molero, B., Loladze, V. V., Makhatadze, G. I. & Sanchez-Ruiz, J. M. Thermal versus guanidine-induced unfolding of ubiquitin. An analysis in terms of the contributions from charge-charge interactions to protein stability. *Biochemistry* **38**, 8138–8149 (1999).
6. Makhatadze, G. I., Lopez, M. M., Richardson, J. M. & Thomas, S. T. Anion binding to the ubiquitin molecule. *Protein Sci.* **7**, 689–697 (1998).
7. Lenkinski, R. E., Chen, D. M., Glickson, J. D. & Goldstein, G. Nuclear magnetic resonance studies of the denaturation of ubiquitin. *BBA - Protein Struct.* **494**, 126–130 (1977).
8. Vijay-kumar, S., Bugg, C. E. & Cook, W. J. Structure of ubiquitin refined at 1.8 Å resolution. *J. Mol. Biol.* **194**, 531–544 (1987).
9. Burroughs, A. M., Balaji, S., Iyer, L. M. & Aravind, L. Small but versatile: The extraordinary functional and structural diversity of the β-grasp fold. *Biol. Direct* **2**, 1–28 (2007).
10. Komander, D. & Rape, M. The Ubiquitin Code. *Annu. Rev. Biochem.* **81**, 203–229 (2012).
11. Swatek, K. N. & Komander, D. Ubiquitin modifications. *Cell Res.* **26**, 399–422 (2016).
12. Hershko, A., Heller, H., Elias, S. & Ciechanover, A. Components of Ubiquitin-Protein Ligase System. *J. Biol. Chem.* **258**, 8206–8214 (1983).
13. Husnjak, K. & Dikic, I. Ubiquitin-Binding Proteins: Decoders of Ubiquitin-Mediated Cellular Functions. *Annu. Rev. Biochem.* **81**, 291–322 (2012).
14. Hofmann, K. & Bucher, P. The UBA domain: a sequence motif present in multiple enzyme classes of the ubiquitination pathway. *Trends Biochem. Sci.* **21**, 172–3 (1996).
15. Hofmann, K. & Falquet, L. A ubiquitin-interacting motif conserved in components of the

proteasomal and lysosomal protein degradation systems. *Trends Biochem. Sci.* **26**, 347–350 (2001).

16. Young, P., Deveraux, Q., Beal, R. E., Pickart, C. M. & Rechsteiner, M. Characterization of two polyubiquitin binding sites in the 26 S protease subunit 5a. *J. Biol. Chem.* **273**, 5461–7 (1998).
17. Scott, D., Oldham, N. J., Strachan, J., Searle, M. S. & Layfield, R. Ubiquitin-binding domains: Mechanisms of ubiquitin recognition and use as tools to investigate ubiquitin-modified proteomes. *Proteomics* (2015). doi:10.1002/pmic.201400341
18. Hicke, L., Schubert, H. L. & Hill, C. P. Ubiquitin-binding domains. *Nat. Rev. Mol. Cell Biol.* **6**, 610–621 (2005).
19. Haas, A. L. & Bright, P. M. The immunochemical detection and quantitation of intracellular ubiquitin-protein conjugates. *J. Biol. Chem.* **260**, 12464–12473 (1985).
20. Trempe, J. F. Reading the ubiquitin postal code. *Curr. Opin. Struct. Biol.* **21**, 792–801 (2011).
21. Chai, Y., Berke, S. S., Cohen, R. E. & Paulson, H. L. Poly-ubiquitin Binding by the Polyglutamine Disease Protein Ataxin-3 Links Its Normal Function to Protein Surveillance Pathways. *J. Biol. Chem.* **279**, 3605–3611 (2004).
22. Song, A. X. *et al.* Structural transformation of the tandem ubiquitin-interacting motifs in ataxin-3 and their cooperative interactions with ubiquitin chains. *PLoS One* **5**, (2010).
23. Zhang, N. *et al.* Structure of the s5a:k48-linked diubiquitin complex and its interactions with rpn13. *Mol. Cell* **35**, 280–90 (2009).
24. Sims, J. J. & Cohen, R. E. Linkage-Specific Avidity Defines the Lysine 63-Linked Polyubiquitin-Binding Preference of Rap80. *Mol. Cell* (2009). doi:10.1016/j.molcel.2009.02.011
25. Varadan, R., Assfalg, M., Raasi, S., Pickart, C. & Fushman, D. Structural determinants for selective recognition of a Lys48-linked polyubiquitin chain by a UBA domain. *Mol. Cell* **18**, 687–698 (2005).
26. Trempe, J. F. *et al.* Mechanism of Lys48-linked polyubiquitin chain recognition by the Mud1 UBA domain. *EMBO J.* **24**, 3178–3189 (2005).
27. Zhang, D., Raasi, S. & Fushman, D. Affinity makes the difference: nonselective interaction of the UBA domain of Ubiquilin-1 with monomeric ubiquitin and polyubiquitin chains. *J. Mol. Biol.* **377**, 162–80 (2008).
28. Ohno, A. *et al.* Structure of the UBA domain of Dsk2p in complex with ubiquitin:

Molecular determinants for ubiquitin recognition. *Structure* **13**, 521–532 (2005).

29. Raasi, S., Varadan, R., Fushman, D. & Pickart, C. M. Diverse polyubiquitin interaction properties of ubiquitin-associated domains. *Nat. Struct. Mol. Biol.* **12**, 708–714 (2005).
30. Collins, G. A. & Goldberg, A. L. The Logic of the 26S Proteasome. *Cell* **169**, 792–806 (2017).
31. Koizumi, S., Hamazaki, J. & Murata, S. Transcriptional regulation of the 26S proteasome by Nrf1. *Proc. Japan Acad. Ser. B* **94**, 325–36 (2018).
32. Baumeister, W., Walz, J., Zühl, F. & Seemüller, E. The proteasome: Paradigm of a self-compartmentalizing protease. *Cell* **92**, 367–380 (1998).
33. Kisselev, A. F., Akopian, T. N., Woo, K. M. & Goldberg, A. L. The sizes of peptides generated from protein by mammalian 26 and 20 S proteasomes. Implications for understanding the degradative mechanism and antigen presentation. *J. Biol. Chem.* **274**, 3363–71 (1999).
34. Peth, A., Besche, H. C. & Goldberg, A. L. Ubiquitinated Proteins Activate the Proteasome by Binding to Usp14/Ubp6, which Causes 20S Gate Opening. *Mol. Cell* **36**, 794–804 (2009).
35. Bard, J. A. M. *et al.* Structure and Function of the 26S Proteasome. *Annu. Rev. Biochem.* **87**, 697–724 (2018).
36. Deveraux, Q., Ustrell, V., Pickart, C. & Rechsteiner, M. A 26 S protease subunit that binds ubiquitin conjugates. *J. Biol. Chem.* **269**, 7059–7061 (1994).
37. Husnjak, K. *et al.* Proteasome subunit Rpn13 is a novel ubiquitin receptor. *Nature* (2008). doi:10.1038/nature06926
38. Shi, Y. *et al.* Rpn1 provides adjacent receptor sites for substrate binding and deubiquitination by the proteasome. *Science (80-.).* **351**, (2016).
39. Verma, R. *et al.* Role of Rpn11 metalloprotease in deubiquitination and degradation by the 26S proteasome. *Science (80-.).* **298**, 611–615 (2002).
40. Yao, T. & Cohen, R. E. A cryptic protease couples deubiquitination and degradation by the proteasome. *Nature* (2002). doi:10.1038/nature01071
41. Peth, A., Uchiki, T. & Goldberg, A. L. ATP-Dependent steps in the binding of ubiquitin conjugates to the 26s proteasome that commit to degradation. *Mol. Cell* **40**, 671–681 (2010).
42. Peth, A., Nathan, J. A. & Goldberg, A. L. The ATP costs and time required to degrade ubiquitinated proteins by the 26 S proteasome. *J. Biol. Chem.* **288**, 29215–29222 (2013).

43. Hinnerwisch, J., Fenton, W. A., Furtak, K. J., Farr, G. W. & Horwich, A. L. Loops in the central channel of ClpA chaperone mediate protein binding, unfolding, and translocation. *Cell* **121**, 1029–1041 (2005).

44. Snoberger, A., Anderson, R. T. & Smith, D. M. The Proteasomal ATPases Use a Slow but Highly Processive Strategy to Unfold Proteins. *Front. Mol. Biosci.* **4**, (2017).

45. Matyskiela, M. E., Lander, G. C. & Martin, A. Conformational switching of the 26S proteasome enables substrate degradation. *Nat. Struct. Mol. Biol.* (2013). doi:10.1038/nsmb.2616

46. Peth, A., Kukushkin, N., Bossé, M. & Goldberg, A. L. Ubiquitinated proteins activate the proteasomal ATPases by binding to Usp14 or Uch37 homologs. *J. Biol. Chem.* **288**, 7781–7790 (2013).

47. Pickart, C. M. Targeting of substrates to the 26S proteasome. *FASEB J.* **11**, 1055–66 (1997).

48. Thrower, J. S. Recognition of the polyubiquitin proteolytic signal. *EMBO J.* **19**, 94–102 (2000).

49. Hofmann, R. M. & Pickart, C. M. In Vitro Assembly and Recognition of Lys-63 Polyubiquitin Chains. *J. Biol. Chem.* **276**, 27936–27943 (2001).

50. Hyoung, T. K. *et al.* Certain pairs of ubiquitin-conjugating enzymes (E2s) and ubiquitin-protein ligases (E3s) synthesize nondegradable forked ubiquitin chains containing all possible isopeptide linkages. *J. Biol. Chem.* **282**, 17375–17386 (2007).

51. Grice, G. L. *et al.* The proteasome distinguishes between heterotypic and homotypic lysine-11-Linked polyubiquitin chains. *Cell Rep.* **12**, 545–553 (2015).

52. Xu, P. *et al.* Quantitative Proteomics Reveals the Function of Unconventional Ubiquitin Chains in Proteasomal Degradation. *Cell* **137**, 133–145 (2009).

53. Nathan, J. A., Tae Kim, H., Ting, L., Gygi, S. P. & Goldberg, A. L. Why do cellular proteins linked to K63-polyubiquitin chains not associate with proteasomes? *EMBO J.* **32**, 552–565 (2013).

54. Dimova, N. V. *et al.* APC/C-mediated multiple monoubiquitylation provides an alternative degradation signal for cyclin B1. *Nat. Cell Biol.* **14**, 168–176 (2012).

55. Braten, O. *et al.* Numerous proteins with unique characteristics are degraded by the 26S proteasome following monoubiquitination. *Proc. Natl. Acad. Sci.* **113**, E4639–E4647 (2016).

56. Lee, B.-H. *et al.* USP14 deubiquitinates proteasome-bound substrates that are

ubiquitinated at multiple sites. *Nature* **532**, 398–401 (2016).

57. Lu, Y., Lee, B. H., King, R. W., Finley, D. & Kirschner, M. W. Substrate degradation by the proteasome: A single-molecule kinetic analysis. *Science (80-.).* **348**, 183–184 (2015).

58. Aufderheide, A. *et al.* Structural characterization of the interaction of Ubp6 with the 26S proteasome. *Proc. Natl. Acad. Sci.* **112**, 8626–8631 (2015).

59. Borodovsky, A. *et al.* A novel active site-directed probe specific for deubiquitylating enzymes reveals proteasome association of USP14. *EMBO J.* **20**, 5187–96 (2001).

60. Yao, T. *et al.* Proteasome recruitment and activation of the Uch37 deubiquitinating enzyme by Adrm1. *Nat. Cell Biol.* **8**, 994–1002 (2006).

61. VanderLinden, R. T. *et al.* Structural Basis for the Activation and Inhibition of the UCH37 Deubiquitylase. *Mol. Cell* (2015). doi:10.1016/j.molcel.2015.01.016

62. Goldberg, A. L. & Dice, J. F. Intracellular protein degradation in mammalian and bacterial cells. *Annu. Rev. Biochem.* **43**, 835–69 (1974).

63. van der Lee, R. *et al.* Intrinsically disordered segments affect protein half-life in the cell and during evolution. *Cell Rep.* **8**, 1832–1844 (2014).

64. Lee, C., Schwartz, M. P., Prakash, S., Iwakura, M. & Matouschek, A. ATP-dependent proteases degrade their substrates by processively unraveling them from the degradation signal. *Mol. Cell* **7**, 627–37 (2001).

65. Yu, H. *et al.* Conserved sequence preferences contribute to substrate recognition by the proteasome. *J. Biol. Chem.* **291**, 14526–14539 (2016).

66. Hagai, T. & Levy, Y. Ubiquitin not only serves as a tag but also assists degradation by inducing protein unfolding. *Proc. Natl. Acad. Sci.* **107**, 2001–2006 (2010).

67. Meyer, H., Bug, M. & Bremer, S. Emerging functions of the VCP/p97 AAA-ATPase in the ubiquitin system. *Nat. Cell Biol.* **14**, 117–123 (2012).

68. Buchberger, A. From UBA to UBX: New words in the ubiquitin vocabulary. *Trends Cell Biol.* **12**, 216–221 (2002).

69. Cappadocia, L. & Lima, C. D. Ubiquitin-like Protein Conjugation: Structures, Chemistry, and Mechanism. *Chem. Rev.* **118**, 889–918 (2018).

70. Kiel, C. & Serrano, L. The ubiquitin domain superfold: Structure-based sequence alignments and characterization of binding epitopes. *J. Mol. Biol.* **355**, 821–844 (2006).

71. Faesen, A. C., Luna-Vargas, M. P. A. & Sixma, T. K. The role of UBL domains in ubiquitin-specific proteases. *Biochem. Soc. Trans.* **40**, 539–545 (2012).

72. Jentsch, S. & Pyrowolakis, G. Ubiquitin and its kin: How close are the family ties? *Trends Cell Biol.* **10**, 335–342 (2000).

73. Hartmann-Petersen, R. & Gordon, C. Integral UBL domain proteins: A family of proteasome interacting proteins. *Semin. Cell Dev. Biol.* **15**, 247–259 (2004).

74. van der Veen, A. G. & Ploegh, H. L. Ubiquitin-Like Proteins. *Annu. Rev. Biochem.* **81**, 323–357 (2012).

75. Grabbe, C. & Dikic, I. Functional roles of ubiquitin-like domain (ULD) and ubiquitin-binding domain (UBD) containing proteins. *Chem. Rev.* **109**, 1481–1494 (2009).

76. Watkins, J. F., Sung, P., Prakash, L. & Prakash, S. The Saccharomyces cerevisiae DNA repair gene RAD23 encodes a nuclear protein containing a ubiquitin-like domain required for biological function. *Mol. Cell. Biol.* **13**, 7757–7765 (1993).

77. Baek, G. H. *et al.* Cdc48: A Swiss Army Knife of Cell Biology. *J. Amino Acids* **2013**, 1–12 (2013).

78. Schuberth, C. & Buchberger, A. UBX domain proteins: major regulators of the AAA ATPase Cdc48/p97. *Cell. Mol. Life Sci.* **65**, 2360–71 (2008).

79. Kondo, H. *et al.* P47 Is a Cofactor for P97-Mediated Membrane Fusion. *Nature* **388**, 75–78 (1997).

80. Schuberth, C. & Buchberger, A. Membrane-bound Ubx2 recruits Cdc48 to ubiquitin ligases and their substrates to ensure efficient ER-associated protein degradation. *Nat. Cell Biol.* **7**, 999–1006 (2005).

81. Hänzelmann, P., Stingele, J., Hofmann, K., Schindelin, H. & Raasi, S. The yeast E4 ubiquitin ligase Ufd2 interacts with the ubiquitin-like domains of Rad23 and Dsk2 via a novel and distinct ubiquitin-like binding domain. *J. Biol. Chem.* **285**, 20390–20398 (2010).

82. Chaugule, V. K. *et al.* Autoregulation of Parkin activity through its ubiquitin-like domain. *EMBO J.* **30**, 2853–2867 (2011).

83. Kondapalli, C. *et al.* PINK1 is activated by mitochondrial membrane potential depolarization and stimulates Parkin E3 ligase activity by phosphorylating Serine 65. *Open Biol.* (2012). doi:10.1098/rsob.120080

84. Zhu, X., Ménard, R. & Sulea, T. High incidence of ubiquitin-like domains in human ubiquitin-specific proteases. *Proteins* **69**, 1–7 (2007).

85. Hu, M. *et al.* Structure and mechanisms of the proteasome-associated deubiquitinating enzyme USP14. *EMBO J.* **24**, 3747–3756 (2005).

86. Clerici, M., Luna-Vargas, M. P. A., Faesen, A. C. & Sixma, T. K. The DUSP-Ubl domain of USP4 enhances its catalytic efficiency by promoting ubiquitin exchange. *Nat. Commun.* **5**, 1–11 (2014).

87. Faesen, A. C. *et al.* Mechanism of USP7/HAUSP activation by its C-Terminal ubiquitin-like domain and allosteric regulation by GMP-synthetase. *Mol. Cell* **44**, 147–159 (2011).

88. Braakman, I., Hebert, D. N. & Schwartz, T. Protein Folding in the Endoplasmic Reticulum of the Nuclear Pore and Coat Protein Complexes. (2014). doi:10.1101/cshperspect.a013201

89. Rapoport, T. A. Protein translocation across the eukaryotic endoplasmic reticulum and bacterial plasma membranes. *Nature* **450**, 663–669 (2007).

90. Guerriero, C. J. & Brodsky, J. L. The delicate balance between secreted protein folding and endoplasmic reticulum-associated degradation in human physiology. *Physiol. Rev.* **92**, 537–76 (2012).

91. Hartl, F. U. & Hayer-Hartl, M. Converging concepts of protein folding in vitro and in vivo. *Nat. Struct. Mol. Biol.* **16**, 574–581 (2009).

92. Werner, E. D., Brodsky, J. L. & McCracken, A. A. Proteasome-dependent endoplasmic reticulum-associated protein degradation: An unconventional route to a familiar fate. *Proc. Natl. Acad. Sci.* **93**, 13797–13801 (1996).

93. Tsai, B., Ye, Y. & Rapoport, T. A. Retro-translocation of proteins from the endoplasmic reticulum into the cytosol. *Nat. Rev. Mol. Cell Biol.* **3**, 246–255 (2002).

94. Ye, Y., Meyer, H. H. & Rapoport, T. A. Function of the p97-Ufd1-Npl4 complex in retrotranslocation from the ER to the cytosol: Dual recognition of nonubiquitinated polypeptide segments and polyubiquitin chains. *J. Cell Biol.* **162**, 71–84 (2003).

95. Ruggiano, A., Foresti, O. & Carvalho, P. ER-associated degradation: Protein quality control and beyond. *J. Cell Biol.* **204**, 869–879 (2014).

96. Bordallo, J., Plemper, R. K., Finger, A. & Wolf, D. H. Der3p/Hrd1p Is Required for Endoplasmic Reticulum-associated Degradation of Misfolded Lumenal and Integral Membrane Proteins. *Mol. Biol. Cell* **9**, 209–222 (1998).

97. Carvalho, P., Goder, V. & Rapoport, T. A. Distinct Ubiquitin-Ligase Complexes Define Convergent Pathways for the Degradation of ER Proteins. *Cell* **126**, 361–373 (2006).

98. Bernasconi, R., Galli, C., Calanca, V., Nakajima, T. & Molinari, M. Stringent requirement for HRD1, SEL1L, and OS-9/XTP3-B for disposal of ERAD-LS substrates. *J. Cell Biol.* **188**, 223–235 (2010).

99. Olzmann, J. A., Kopito, R. R. & Christianson, J. C. The mammalian endoplasmic reticulum-associated degradation system. *Cold Spring Harb. Perspect. Biol.* **5**, (2013).

100. Wu, X. & Rapoport, T. A. Mechanistic insights into ER-associated protein degradation. *Curr. Opin. Cell Biol.* **53**, 22–28 (2018).

101. Bays, N. W., Gardner, R. G., Seelig, L. P., Joazeiro, C. A. & Hampton, R. Y. Hrd1p/Der3p is a membrane-anchored ubiquitin ligase required for ER-associated degradation. *Nat. Cell Biol.* **3**, 24–29 (2001).

102. Schoebel, S. *et al.* Cryo-EM structure of the protein-conducting ERAD channel Hrd1 in complex with Hrd3. *Nature* **548**, 352–355 (2017).

103. Carvalho, P., Stanley, A. M. & Rapoport, T. A. Retrotranslocation of a misfolded luminal ER protein by the ubiquitin-ligase hrd1p. *Cell* (2010). doi:10.1016/j.cell.2010.10.028

104. Denic, V., Quan, E. M. & Weissman, J. S. A Luminal Surveillance Complex that Selects Misfolded Glycoproteins for ER-Associated Degradation. *Cell* **126**, 349–359 (2006).

105. Gauss, R., Jarosch, E., Sommer, T. & Hirsch, C. A complex of Yos9p and the HRD ligase integrates endoplasmic reticulum quality control into the degradation machinery. *Nat. Cell Biol.* **8**, 849–854 (2006).

106. Knop, M., Finger, A., Braun, T., Hellmuth, K. & Wolf, D. H. Der1, a novel protein specifically required for endoplasmic reticulum degradation in yeast. *EMBO J.* **15**, 753–763 (1996).

107. Mehnert, M., Sommer, T. & Jarosch, E. Der1 promotes movement of misfolded proteins through the endoplasmic reticulum membrane. *Nat. Cell Biol.* **16**, 77–86 (2014).

108. Bodnar, N. O. & Rapoport, T. A. Molecular Mechanism of Substrate Processing by the Cdc48 ATPase Complex. *Cell* **169**, 722-735.e9 (2017).

109. Ye, Y. Diverse functions with a common regulator: Ubiquitin takes command of an AAA ATPase. *J. Struct. Biol.* **156**, 29–40 (2006).

110. Müller, J. M. M., Deinhardt, K., Rosewell, I., Warren, G. & Shima, D. T. Targeted deletion of p97 (VCP/CDC48) in mouse results in early embryonic lethality. *Biochem. Biophys. Res. Commun.* **354**, 459–465 (2007).

111. Fröhlich, K. U. *et al.* Yeast cell cycle protein CDC48p shows full-length homology to the mammalian protein VCP and is a member of a protein family involved in secretion, peroxisome formation, and gene expression. *J. Cell Biol.* **114**, 443–53 (1991).

112. Confalonieri, F. & Duguet, M. A 200-amino acid ATPase module in search of a basic function. *BioEssays* (1995). doi:10.1002/bies.950170710

113. Dai, R. M. & Li, C. C. H. Valosin-containing protein is a multi-ubiquitin chain-targeting factor required in ubiquitin-proteasome degradation. *Nat. Cell Biol.* **3**, 740–744 (2001).

114. Zhang, X. *et al.* Structure of the AAA ATPase p97. *Mol. Cell* **6**, 1473–1484 (2000).

115. Tang, W. K. *et al.* A novel ATP-dependent conformation in p97 N-D1 fragment revealed by crystal structures of disease-related mutants. *EMBO J.* **29**, 2217–2229 (2010).

116. Deshaies, R. J. *et al.* 2.3 A resolution cryo-EM structure of human p97 and mechanism of allosteric inhibition. *Science (80-.).* **351**, 871–875 (2016).

117. Ye, Y., Meyer, H. H. & Rapoport, T. A. The AAA ATPase Cdc48/p97 and its partners transport proteins from the ER into the cytosol. *Nature* **414**, 652–656 (2001).

118. Bodnar, N. O. *et al.* Structure of the cdc48 atpase with its ubiquitin-binding cofactor ufd1–npl4. *Nat. Struct. Mol. Biol.* **25**, (2018).

119. Meyer, H. H., Shorter, J. G., Seemann, J., Pappin, D. & Warren, G. A complex of mammalian ufd1 and npl4 links the AAA-ATPase, p97, to ubiquitin and nuclear transport pathways. *EMBO J.* **19**, 2181–92 (2000).

120. Isaacson, R. L. *et al.* Detailed structural insights into the p97-Npl4-Ufd1 interface. *J. Biol. Chem.* **282**, 21361–21369 (2007).

121. Park, S., Isaacson, R., Kim, H. T., Silver, P. A. & Wagner, G. Ufd1 exhibits the AAA-ATPase fold with two distinct ubiquitin interaction sites. *Structure* **13**, 995–1005 (2005).

122. Tsuchiya, H. *et al.* In Vivo Ubiquitin Linkage-type Analysis Reveals that the Cdc48-Rad23/Dsk2 Axis Contributes to K48-Linked Chain Specificity of the Proteasome. *Mol. Cell* **66**, 488-502.e7 (2017).

123. Stein, A., Ruggiano, A., Carvalho, P. & Rapoport, T. A. Key Steps in ERAD of Luminal ER Proteins Reconstituted with Purified Components. *Cell* **158**, 1375–1388 (2014).

124. Ernst, R., Mueller, B., Ploegh, H. L. & Schlieker, C. The otubain YOD1 is a deubiquitinating enzyme that associates with p97 to facilitate protein dislocation from the ER. *Mol. Cell* **36**, 28–38 (2009).

125. Sivaraman, T., Arrington, C. B. & Robertson, A. D. Kinetics of unfolding and folding from amide hydrogen exchange in native ubiquitin. *Nat. Struct. Biol.* **8**, 331–333 (2001).

126. Wang, Q. *et al.* A ubiquitin ligase-associated chaperone holdase maintains polypeptides in soluble states for proteasome degradation. *Mol. Cell* **42**, 758–70 (2011).

127. Payapilly, A. & High, S. BAG6 regulates the quality control of a polytopic ERAD substrate. *J. Cell Sci.* **127**, 2898–2909 (2014).

128. Bohm, S., Lamberti, G., Fernandez-Saiz, V., Stapf, C. & Buchberger, A. Cellular

Functions of Ufd2 and Ufd3 in Proteasomal Protein Degradation Depend on Cdc48 Binding. *Mol. Cell. Biol.* **31**, 1528–1539 (2011).

129. Koegl, M. *et al.* A novel ubiquitination factor, E4, is involved in multiubiquitin chain assembly. *Cell* **96**, 635–644 (1999).

130. Richly, H. *et al.* A series of ubiquitin binding factors connects CDC48/p97 to substrate multiubiquitylation and proteasomal targeting. *Cell* **120**, 73–84 (2005).

131. Allen, M. D., Buchberger, A. & Bycroft, M. The PUB domain functions as a p97 binding module in human peptide N-glycanase. *J. Biol. Chem.* **281**, 25502–25508 (2006).

132. Hirayama, H., Hosomi, A. & Suzuki, T. Physiological and molecular functions of the cytosolic peptide: N-glycanase. *Semin. Cell Dev. Biol.* **41**, 110–120 (2015).

133. Medicherla, B., Kostova, Z., Schaefer, A. & Wolf, D. H. A genomic screen identifies Dsk2p and Rad23p as essential components of ER-associated degradation. *EMBO Rep.* **5**, 692–697 (2004).

134. Goh, A. M. *et al.* Components of the ubiquitin-proteasome pathway compete for surfaces on Rad23 family proteins. *BMC Biochem.* (2008). doi:10.1186/1471-2091-9-4

135. Lee, J.-H., Choi, J. M., Lee, C., Yi, K. J. & Cho, Y. Structure of a peptide:N-glycanase-Rad23 complex: Insight into the deglycosylation for denatured glycoproteins. *Proc. Natl. Acad. Sci.* **102**, 9144–9149 (2005).

136. Kim, I. *et al.* The Png1-Rad23 complex regulates glycoprotein turnover. *J. Cell Biol.* **172**, 211–219 (2006).

137. Katiyar, S., Li, G. & Lennarz, W. J. A complex between peptide:N-glycanase and two proteasome-linked proteins suggests a mechanism for the degradation of misfolded glycoproteins. *Proc. Natl. Acad. Sci.* **101**, 13774–13779 (2004).

138. Wang, G. -h. Ataxin-3, the MJD1 gene product, interacts with the two human homologs of yeast DNA repair protein RAD23, HHR23A and HHR23B. *Hum. Mol. Genet.* **9**, 1795–1803 (2000).

139. Wang, Q., Li, A. & Ye, Y. Regulation of retrotranslocation by p97-associated deubiquitinating enzyme ataxin-3. *J. Cell Biol.* **174**, 963–971 (2006).

140. Doss-Pepe, E. W., Stenroos, E. S., Johnson, W. G. & Madura, K. Ataxin-3 Interactions with Rad23 and Valosin-Containing Protein and Its Associations with Ubiquitin Chains and the Proteasome Are Consistent with a Role in Ubiquitin-Mediated Proteolysis. *Mol. Cell. Biol.* **23**, 6469–6483 (2003).

141. Zhong, X. & Pittman, R. N. Ataxin-3 binds VCP/p97 and regulates retrotranslocation of

ERAD substrates. *Hum. Mol. Genet.* **15**, 2409–20 (2006).

142. Baek, G. H., Kim, I. & Rao, H. The Cdc48 ATPase modulates the interaction between two proteolytic factors Ufd2 and Rad23. *Proc. Natl. Acad. Sci.* **108**, 13558–13563 (2011).

143. Lee, D. Y. & Brown, E. J. Ubiquilins in the crosstalk among proteolytic pathways. *Biological Chemistry* (2012). doi:10.1515/hsz-2012-0120

144. Kim, B., Ryu, K. S., Kim, H. J., Cho, S. J. & Choi, B. S. Solution structure and backbone dynamics of the XPC-binding domain of the human DNA repair protein hHR23B. *FEBS J.* (2005). doi:10.1111/j.1742-4658.2005.04667.x

145. Lässle, M., Blatch, G. L., Kundra, V., Takatori, T. & Zetter, B. R. Stress-inducible, murine protein mSTI1. Characterization of binding domains for heat shock proteins and in vitro phosphorylation by different kinases. *J. Biol. Chem.* **272**, 1876–84 (1997).

146. Elsasser, S. *et al.* Proteasome subunit Rpn1 binds ubiquitin-like protein domains. *Nat. Cell Biol.* **4**, 725–30 (2002).

147. Bertolaet, B. L. *et al.* UBA domains of DNA damage-inducible proteins interact with ubiquitin. *Nat. Struct. Biol.* **8**, 417–22 (2001).

148. Dantuma, N. P., Heinen, C. & Hoogstraten, D. The ubiquitin receptor Rad23: At the crossroads of nucleotide excision repair and proteasomal degradation. *DNA Repair (Amst).* **8**, 449–460 (2009).

149. Kang, Y. *et al.* UBL/UBA ubiquitin receptor proteins bind a common tetraubiquitin chain. *J. Mol. Biol.* **356**, 1027–1035 (2006).

150. Chen, L. & Madura, K. Rad23 promotes the targeting of proteolytic substrates to the proteasome. *Mol. Cell. Biol.* **22**, 4902–13 (2002).

151. Chen, L., Shinde, U., Ortolan, T. G. & Madura, K. Ubiquitin-associated (UBA) domains in Rad23 bind ubiquitin and promote inhibition of multi-ubiquitin chain assembly. *EMBO Rep.* **2**, 933–8 (2001).

152. Raasi, S. & Pickart, C. M. Rad23 ubiquitin-associated domains (UBA) inhibit 26 S proteasome-catalyzed proteolysis by sequestering lysine 48-linked polyubiquitin chains. *J. Biol. Chem.* **278**, 8951–9 (2003).

153. Madura, K. *et al.* The DNA repair protein rad23 is a negative regulator of multi-ubiquitin chain assembly. *Nat. Cell Biol.* (2000). doi:10.1038/35023547

154. Verma, R., Oania, R., Graumann, J. & Deshaies, R. J. Multiubiquitin chain receptors define a layer of substrate selectivity in the ubiquitin-proteasome system. *Cell* **118**, 99–110 (2004).

155. McKnight, G. L., Cardillo, T. S. & Sherman, F. An extensive deletion causing overproduction of yeast iso-2-cytochrome c. *Cell* **25**, 409–19 (1981).

156. Guzder, S. N., Sung, P., Prakash, L. & Prakash, S. Affinity of yeast nucleotide excision repair factor 2, consisting of the Rad4 and Rad23 proteins, for ultraviolet damaged DNA. *J. Biol. Chem.* **273**, 31541–6 (1998).

157. Guzder, S. N., Bailly, V., Sung, P., Prakash, L. & Prakash, S. Yeast DNA repair protein RAD23 promotes complex formation between transcription factor TFIIH and DNA damage recognition factor RAD14. *J. Biol. Chem.* **270**, 8385–8 (1995).

158. Gong, F., Fahy, D. & Smerdon, M. J. Rad4-Rad23 interaction with SWI/SNF links ATP-dependent chromatin remodeling with nucleotide excision repair. *Nat. Struct. Mol. Biol.* **13**, 902–7 (2006).

159. Anantharaman, V. Peptide-N-glycanases and DNA repair proteins, Xp-C/Rad4, are, respectively, active and inactivated enzymes sharing a common transglutaminase fold. *Hum. Mol. Genet.* (2001). doi:10.1093/hmg/10.16.1627

160. Schauber, C. *et al.* Rad23 links DNA repair to the ubiquitin/proteasome pathway. *Nature* **391**, 715–718 (1998).

161. Masutani, C. *et al.* Purification and cloning of a nucleotide excision repair complex involving the xeroderma pigmentosum group C protein and a human homologue of yeast RAD23. *EMBO J.* **13**, 1831–43 (1994).

162. Yokoi, M. & Hanaoka, F. Two mammalian homologs of yeast Rad23, HR23A and HR23B, as multifunctional proteins. *Gene* **597**, 1–9 (2017).

163. Van Der Spek, P. J. *et al.* Cloning, comparative mapping, and RNA expression of the mouse homologues of the Saccharomyces cerevisiae nucleotide excision repair gene RAD23. *Genomics* (1996). doi:10.1006/geno.1996.0004

164. Huang, X. *et al.* Expression of a Novel RAD23B mRNA Splice Variant in the Human Testis. *J. Androl.* (2004). doi:10.1002/j.1939-4640.2004.tb02801.x

165. Ng, J. M. Y. *et al.* Developmental Defects and Male Sterility in Mice Lacking the Ubiquitin-Like DNA Repair Gene mHR23B. *Mol. Cell. Biol.* (2002). doi:10.1128/mcb.22.4.1233-1245.2002

166. Ng, J. M. Y. *et al.* A novel regulation mechanism of DNA repair by damage-induced and RAD23-dependent stabilization of xeroderma pigmentosum group C protein. *Genes Dev.* (2003). doi:10.1101/gad.260003

167. Okuda, Y. *et al.* Relative levels of the two mammalian Rad23 homologs determine

composition and stability of the xeroderma pigmentosum group C protein complex. *DNA Repair (Amst).* (2004). doi:10.1016/j.dnarep.2004.06.010

168. Sugasawa, K. *et al.* HHR23B, a human Rad23 homolog, stimulates XPC protein in nucleotide excision repair in vitro. *Mol. Cell. Biol.* **16**, (1996).

169. Sugasawa, K. *et al.* Two human homologs of Rad23 are functionally interchangeable in complex formation and stimulation of XPC repair activity. *Mol. Cell. Biol.* (1997). doi:10.1128/mcb.17.12.6924

170. Masutani, C. *et al.* Identification and characterization of XPC-binding domain of hHR23B. *Mol. Cell. Biol.* **17**, (1997).

171. Hiyama, H. *et al.* Interaction of hHR23 with S5a. The ubiquitin-like domain of hHR23 mediates interaction with S5a subunit of 26 S proteasome. *J. Biol. Chem.* (1999). doi:10.1074/jbc.274.39.28019

172. Chen, X. *et al.* Structures of Rpn1 T1:Rad23 and hRpn13:hPLIC2 Reveal Distinct Binding Mechanisms between Substrate Receptors and Shuttle Factors of the Proteasome. *Structure* (2016). doi:10.1016/j.str.2016.05.018

173. Scanlon, T. C. *et al.* Isolation of human proteasomes and putative proteasome-interacting proteins using a novel affinity chromatography method. *Exp. Cell Res.* (2009). doi:10.1016/j.yexcr.2008.10.027

174. Besche, H. C., Haas, W., Gygi, S. P. & Goldberg, A. L. Isolation of mammalian 26S proteasomes and p97/VCP complexes using the ubiquitin-like domain from HHR23B reveals novel proteasome-associated proteins. *Biochemistry* (2009). doi:10.1021/bi802198q

175. Walters, K. J., Lech, P. J., Goh, A. M., Wang, Q. & Howley, P. M. DNA-repair protein hHR23a alters its protein structure upon binding proteasomal subunit S5a. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 12694–9 (2003).

176. Ryu, K.-S. *et al.* Binding surface mapping of intra- and interdomain interactions among hHR23B, ubiquitin, and polyubiquitin binding site 2 of S5a. *J. Biol. Chem.* **278**, 36621–7 (2003).

177. Elsasser, S. *et al.* Proteasome subunit Rpn1 binds ubiquitin-like protein domains. *Nat. Cell Biol.* **4**, 725–30 (2002).

178. Funakoshi, M., Sasaki, T., Nishimoto, T. & Kobayashi, H. Budding yeast Dsk2p is a polyubiquitin-binding protein that can interact with the proteasome. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 745–50 (2002).

179. Matiuhin, Y. *et al.* Extraproteasomal Rpn10 restricts access of the polyubiquitin-binding protein Dsk2 to proteasome. *Mol. Cell* **32**, 415–25 (2008).

180. Lu, K., den Brave, F. & Jentsch, S. Receptor oligomerization guides pathway choice between proteasomal and autophagic degradation. *Nat. Cell Biol.* **19**, 732–739 (2017).

181. Kaye, F. J. *et al.* A family of ubiquitin-like proteins binds the ATPase domain of Hsp70-like Stch. *FEBS Lett.* **467**, 348–55 (2000).

182. Biggins, S., Ivanovska, I. & Rose, M. D. Yeast ubiquitin-like genes are involved in duplication of the microtubule organizing center. *J. Cell Biol.* **133**, 1331–1346 (1996).

183. Rao, H. & Sastry, A. Recognition of specific ubiquitin conjugates is important for the proteolytic functions of the ubiquitin-associated domain proteins Dsk2 and Rad23. *J. Biol. Chem.* **277**, 11691–5 (2002).

184. Sasaki, T., Funakoshi, M., Endicott, J. A. & Kobayashi, H. Budding yeast Dsk2 protein forms a homodimer via its C-terminal UBA domain. *Biochem. Biophys. Res. Commun.* **336**, 530–5 (2005).

185. Zientara-Rytter, K. & Subramani, S. The Roles of Ubiquitin-Binding Protein Shuttles in the Degradative Fate of Ubiquitinated Proteins in the Ubiquitin-Proteasome System and Autophagy. *Cells* (2019). doi:10.3390/cells8010040

186. Marín, I. The ubiquilin gene family: Evolutionary patterns and functional insights. *BMC Evol. Biol.* (2014). doi:10.1186/1471-2148-14-63

187. Kleijnen, M. F. *et al.* The hPLIC proteins may provide a link between the ubiquitination machinery and the proteasome. *Mol. Cell* **6**, 409–419 (2000).

188. Li, X., Su, V., Kurata, W. E., Jin, C. & Lau, A. F. A novel connexin43-interacting protein, CIP75, which belongs to the UbL-UBA protein family, regulates the turnover of connexin43. *J. Biol. Chem.* **283**, 5748–59 (2008).

189. Walters, K. J., Kleijnen, M. F., Goh, A. M., Wagner, G. & Howley, P. M. Structural studies of the interaction between ubiquitin family proteins and proteasome subunit S5a. *Biochemistry* (2002). doi:10.1021/bi011892y

190. Heir, R. *et al.* The UBL domain of PLIC-1 regulates aggresome formation. *EMBO Rep.* **7**, 1252–8 (2006).

191. Lim, P. J. *et al.* Ubiquilin and p97/VCP bind erasin, forming a complex involved in ERAD. *J. Cell Biol.* **187**, 201–17 (2009).

192. Kim, T.-Y., Kim, E., Yoon, S. K. & Yoon, J.-B. Herp enhances ER-associated protein degradation by recruiting ubiquilins. *Biochem. Biophys. Res. Commun.* **369**, 741–6 (2008).

193. Kang, Y., Zhang, N., Koepp, D. M. & Walters, K. J. Ubiquitin receptor proteins hHR23a and hPLIC2 interact. *J. Mol. Biol.* **365**, 1093–101 (2007).

194. N'Diaye, E.-N. *et al.* PLIC proteins or ubiquilins regulate autophagy-dependent cell survival during nutrient starvation. *EMBO Rep.* **10**, 173–9 (2009).

195. Lee, D. Y., Arnott, D. & Brown, E. J. Ubiquilin4 is an adaptor protein that recruits Ubiquilin1 to the autophagy machinery. *EMBO Rep.* **14**, 373–81 (2013).

196. Rothenberg, C. *et al.* Ubiquilin functions in autophagy and is degraded by chaperone-mediated autophagy. *Hum. Mol. Genet.* (2010). doi:10.1093/hmg/ddq231

197. Chuang, K.-H., Liang, F., Higgins, R. & Wang, Y. Ubiquilin/Dsk2 promotes inclusion body formation and vacuole (lysosome)-mediated disposal of mutated huntingtin. *Mol. Biol. Cell* **27**, 2025–36 (2016).

198. Hjerpe, R. *et al.* UBQLN2 Mediates Autophagy-Independent Protein Aggregate Clearance by the Proteasome. *Cell* (2016). doi:10.1016/j.cell.2016.07.001

199. Krylov, D. M. & Koonin, E. V. A novel family of predicted retroviral-like aspartyl proteases with a possible key role in eukaryotic cell cycle control. *Curr. Biol.* **11**, R584-7 (2001).

200. Sirkis, R., Gerst, J. E. & Fass, D. Ddi1, a Eukaryotic Protein With the Retroviral Protease Fold. *J. Mol. Biol.* **364**, 376–387 (2006).

201. Trempe, J. F. *et al.* Structural studies of the yeast DNA damage-inducible protein Ddi1 reveal domain architecture of this eukaryotic protein family. *Sci. Rep.* **6**, (2016).

202. Perteguer, M. J. *et al.* Ddi1-like protein from Leishmania major is an active aspartyl proteinase. *Cell Stress Chaperones* **18**, 171–81 (2013).

203. Svoboda, M., Konvalinka, J., Trempe, J.-F. & Grantz Saskova, K. The yeast proteases Ddi1 and Wss1 are both involved in the DNA replication stress response. *DNA Repair (Amst).* (2019). doi:10.1016/j.dnarep.2019.06.008

204. Kožíšek, M. *et al.* Thermodynamic and structural analysis of HIV protease resistance to darunavir - Analysis of heavily mutated patient-derived HIV-1 proteases. *FEBS J.* **281**, 1834–1847 (2014).

205. Nowicka, U. *et al.* DNA-damage-inducible 1 protein (Ddi1) contains an uncharacteristic ubiquitin-like domain that binds ubiquitin. *Structure* **23**, 542–557 (2015).

206. Gomez, T. A., Kolawa, N., Gee, M., Sweredoski, M. J. & Deshaies, R. J. Identification of a functional docking site in the Rpn1 LRR domain for the UBA-UBL domain protein Ddi1. *BMC Biol.* **9**, 33 (2011).

207. Rosenzweig, R., Bronner, V., Zhang, D., Fushman, D. & Glickman, M. H. Rpn1 and Rpn2 coordinate ubiquitin processing factors at proteasome. *J. Biol. Chem.* **287**, 14659–71 (2012).

208. Saeki, Y., Saitoh, A., Toh-e, A. & Yokosawa, H. Ubiquitin-like proteins and Rpn10 play cooperative roles in ubiquitin-dependent proteolysis. *Biochem. Biophys. Res. Commun.* **293**, 986–92 (2002).

209. Liu, Y. & Xiao, W. Bidirectional regulation of two DNA-damage-inducible genes, MAG1 and DDI1, from Saccharomyces cerevisiae. *Mol. Microbiol.* **23**, 777–89 (1997).

210. Serbyn, N. *et al.* The Aspartic Protease Ddi1 Contributes to DNA-Protein Crosslink Repair in Yeast. *bioRxiv* 575860 (2019). doi:10.1101/575860

211. Kaplun, L. *et al.* The DNA damage-inducible UbL-UbA protein Ddi1 participates in Mec1-mediated degradation of Ho endonuclease. *Mol. Cell. Biol.* **25**, 5355–62 (2005).

212. Kaplun, L., Ivantsiv, Y., Kornitzer, D. & Raveh, D. Functions of the DNA damage response pathway target Ho endonuclease of yeast for degradation via the ubiquitin-26S proteasome system. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 10077–82 (2000).

213. Ivantsiv, Y., Kaplun, L., Tzirkin-Goldin, R., Shabek, N. & Raveh, D. Turnover of SCFUfo1 complexes requires the UbL-UbA motif protein, Ddi1. *Mol Cell Biol* **26**, 1579–1588 (2006).

214. Clarke, D. J. *et al.* Dosage suppressors of pds1 implicate ubiquitin-associated domains in checkpoint control. *Mol. Cell. Biol.* **21**, 1997–2007 (2001).

215. Gabriely, G., Kama, R., Gelin-Licht, R. & Gerst, J. E. Different domains of the UBL-UBA ubiquitin receptor, Ddi1/Vsm1, are involved in its multiple cellular roles. *Mol. Biol. Cell* **19**, 3625–37 (2008).

216. Kama, R., Gabriely, G., Kanneganti, V. & Gerst, J. E. Cdc48 and ubiquilins confer selective anterograde protein sorting and entry into the multivesicular body in yeast. *Mol. Biol. Cell* mbc.E17-11-0652 (2018). doi:10.1091/mbc.e17-11-0652

217. Marash, M. & Gerst, J. E. Phosphorylation of the autoinhibitory domain of the Sso t-SNAREs promotes binding of the Vsm1 SNARE regulator in yeast. *Mol. Biol. Cell* **14**, 3114–25 (2003).

218. Lustgarten, V. & Gerst, J. E. Yeast VSM1 encodes a v-SNARE binding protein that may act as a negative regulator of constitutive exocytosis. *Mol. Cell. Biol.* **19**, 4480–94 (1999).

219. White, R. E., Dickinson, J. R., Semple, C. A. M., Powell, D. J. & Berry, C. The retroviral proteinase active site and the N-terminus of Ddi1 are required for repression of protein

secretion. *FEBS Lett.* **585**, 139–42 (2011).

220. Lee, C. S. *et al.* Loss of nuclear factor E2-related factor 1 in the brain leads to dysregulation of proteasome gene expression and neurodegeneration. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 8408–13 (2011).

221. Lee, C. S., Ho, D. V & Chan, J. Y. Nuclear factor-erythroid 2-related factor 1 regulates expression of proteasome genes in hepatocytes and protects against endoplasmic reticulum stress and steatosis in mice. *FEBS J.* **280**, 3609–20 (2013).

222. Radhakrishnan, S. K. *et al.* Transcription Factor Nrf1 Mediates the Proteasome Recovery Pathway after Proteasome Inhibition in Mammalian Cells. *Mol. Cell* **38**, 17–28 (2010).

223. Baird, L. *et al.* A Homeostatic Shift Facilitates Endoplasmic Reticulum Proteostasis through Transcriptional Integration of Proteostatic Stress Response Pathways. *Mol. Cell. Biol.* **37**, (2017).

224. Radhakrishnan, S. K., den Besten, W. & Deshaies, R. J. p97-dependent retrotranslocation and proteolytic processing govern formation of active Nrf1 upon proteasome inhibition. *Elife* **2014**, 1–15 (2014).

225. Tomlin, F. M. *et al.* Inhibition of NGLY1 Inactivates the Transcription Factor Nrf1 and Potentiates Proteasome Inhibitor Cytotoxicity. *ACS Cent. Sci.* **3**, 1143–1155 (2017).

226. Konvalinka, J., Kräusslich, H. G. & Müller, B. Retroviral proteases and their roles in virion maturation. *Virology* **479–480**, 403–417 (2015).

227. Chowdhury, A. M. M. A. *et al.* Multiple regulatory mechanisms of the biological function of NRF3 (NFE2L3) control cancer cell proliferation. *Sci. Rep.* **7**, 1–14 (2017).

228. Okonechnikov, K. *et al.* Unipro UGENE: A unified bioinformatics toolkit. *Bioinformatics* **28**, 1166–1167 (2012).

229. Lee, W., Tonelli, M. & Markley, J. L. NMRFAM-SPARKY: Enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **31**, 1325–1327 (2015).

230. Waudby, C. A., Ramos, A., Cabrita, L. D. & Christodoulou, J. Two-Dimensional NMR Lineshape Analysis. *Sci. Rep.* **6**, (2016).

231. Gasteiger, E. *et al.* Protein Analysis Tools on the ExPASy Server. *Proteomics Protoc. Handb. Protein Identif. Anal. Tools ExPASy Serv.* 571–607 (2005). doi:10.1385/1592598900

232. Bateman, A. UniProt: A worldwide hub of protein knowledge. *Nucleic Acids Res.* **47**, D506–D515 (2019).

233. Chen, Y. W., Tajima, T. & Agrawal, S. The crystal structure of the ubiquitin-like (UbL)

domain of human homologue A of Rad23 (hHR23A) protein. *Protein Eng. Des. Sel.* **24**, 131–138 (2011).

234. Van der Spek, P. J. *et al.* XPC and human homologs of RAD23: Intracellular localization and relationship to other nucleotide excision repair complexes. *Nucleic Acids Res.* (1996). doi:10.1093/nar/24.13.2551

235. Frueh, D. P. Practical aspects of NMR signal assignment in larger and challenging proteins. *Progress in Nuclear Magnetic Resonance Spectroscopy* (2014). doi:10.1016/j.pnmrs.2013.12.001