Substrate Recognition by the Proteasome

Jonas Boehringer

Department of Biochemistry

and

Linacre College

University of Oxford

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy at the University of Oxford

Trinity 2010

Ubiquitin Recognition by the Proteasome

Jonas Boehringer

Department of Biochemistry and Linacre College, University of Oxford
Submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy at the University of Oxford

Trinity 2010

Abstract

The ubiquitin proteasome system targets proteins to the proteasome where they are degraded. Substrate recognition and processing prior to degradation take place at the 19S regulatory particle of the proteasome. A polyubiquitin chain, linked through isopeptide bonds formed between the C-terminal G76 and K48, is the signal responsible for delivery to the proteasome. Because chains linked via any of the seven lysine residues of ubiquitin exist *in vivo* and encode signals unrelated to protein degradation it is crucial for cells to avoid crosstalk between these different pathways. Several ubiquitin receptors related to proteasomal degradation have been identified but the selectivity between the different ubiquitin chains has not been assessed quantitatively while avoiding artefacts attributed to GST-dimerisation. By employing isothermal titration calorimetry, analytical ultracentrifugation and nuclear magnetic resonance, discrimination between K48- and K63-linked diubiquitin was established for the *S. pombe* proteasomal receptor Rpn10 and the shuttle protein Rhp23. The same methods allowed us to propose a discriminatory model for Rpn10. The crystal structures of the 19S regulatory particle subunits $Rpn10_{1-193}$ and $Rpn12_{1-224}$ have been determined and possible protein-protein interaction sites were identified by surface conservation and electrostatics analysis. Rpn12 surface residues were identified that had a negative effect on Rpn10-binding. This interaction was studied by surface plasmon resonance, fluorescence anisotropy and nuclear magnetic resonance. These experiments revealed a binding site on Rpn10 that is exclusively occupied by either ubiquitin or Rpn12 and for the first time demonstrated the interaction of a ubiquitin interacting motif with a protein other than ubiquitin.

Acknowledgements

I would like to thank my supervisor Jane Endicott for encouraging me to pursue my own ideas. This has taught me to independently approach and develop scientific projects. I am also very grateful to Anthony Watts and Iain Campbell who always gave me valuable advice and support, whenever needed.

Nick Brown, Max Soegaard and Sonja Baumli helped me out in the lab many times and always showed a keen interest in my work. Sonja also always had a sympathetic ear for both good and bad news during our regular runs through Oxford.

Jean-François Trempe started this project and, while preparing for his *viva,* made time to introduce me to the topic and teach me enough for a good start.

I am indebted to several colleagues in the lab for introducing me to a variety of methods. Christiane Riedinger performed all the NMR experiments mentioned in this thesis. Ed Lowe helped with all aspects of protein crystallography and James Parker introduced me to ITC. For creating a good working environment I am thankful to all LMB labmembers.

I would like to thank all the Part II students I supervised during the last three years: Christina Khoudian, Jack Dean, Dominique Smith and Daniel Grabarczyk. Working with them was a satisfying and inspiring experience.

Our collaborator Colin Gordon complemented my work with his expertise in genetics. Working with his postdoc Kostantinos Paraskevopoulos was a great experience and led to the probably most efficient two weeks of my doctorate.

I would like to thank the Wellcome Trust for generous financial support.

I thank Martina who has always shown an interest in my work for critically reading a draft of my thesis.

I am most grateful to my parents who have always encouraged me to follow my own path. My grandmother's insatiable curiosity for life has taught me to always keep on asking questions and exploring.

Finally, I am indebted to Johanna. You are the reason why I came to Oxford and have shown a great deal of patience during the last couple of months. Generally, life is better with you. Our son Paul, even if unconsciously, made our life as first-time parents as easy as possible during the first months of his life. This is a gift one cannot appreciate enough. He is a pool of happiness and a constant reminder of the truly important things in life.

Abbreviations

AAA	ATPase associated with various cellular activities
APC/C	Anaphase-promoting complex / Cyclosome
ATP	Adenosine tri-phosphate
AUC	Analytical ultracentrifugation
C-terminal	Carboxy-terminal
CCP4	Collaborative Computational Project number 4
Cdc34	Cell division cycle 34
CP	Core particle
CSN	COP9 signalosome
Ddi1	DNA-damage inducible-1
DHFR	Dihydrofolate reductase
DNA	Deoxyribonucleic acid
Dsk2	Dominant suppressor of kar1 mutant - 2
DTT	Dithiothreitol
DUB	Deubiquitylating enzyme
E. coli	*Escherichia coli*
EDTA	Ethylene Diamine Tetra-acetic Acid
ESI-MS	Electrospray ionisation mass spectrometry
ESRF	European synchrotron radiation facility
FL	Full-length
GST	Glutathione-S-transferase
HBS	HEPES-buffered saline
HEPES	N-[2-hydroxyethyl] piperazine-N-[2-ethanesulphonic acid]
HSQC	Heteronuclear single-quantum coherence
IPTG	Isopropyl β-thiogalactoside
kDa	kilo Dalton
LB	Luria Bertani Broth
M	Molar
MTG	α-monothioglycerol
Mts	Mutant in temperature-sensitive complementation gene
MWCO	Molecular weight cut-off
N-terminal	Amino-terminal
NMR	Nuclear magnetic resonance
Od_x	Optical density at x nm
PAGE	Polyacrylamide gel electrophoresis
PC	Proteasome-cyclosome repeat
PCI	Proteasome, COP9 signalosome, Initiation factor 3
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PEG	Polyethylene glycol
Pus1	Polyubiquitin-binding subunit-1
Rad23	Radiation sensitive mutant complementation gene 23
rmsd	Root mean square deviations
RP	Regulatory particle
Rpn	Regulatory particle non-ATPase subunit
Rpt	Regulatory particle ATPase subunit
S. cerevisiae	*Saccharomyces cerevisiae*
S. pombe	*Schizosaccharomyces pombe*
SAD	Single-wavelength anomalous dispersion
SDS	Sodium dodecyl sulphate

SEC	Size exclusion chromatography
SOC	Super Optimal Broth with Catabolic repression
SPR	Surface plasmon resonance
TPR	Tetra-trico peptide repeat
Tris	Tris-buffered saline
Ub	Ubiquitin
UBA	Ubiquitin-associated domain
UBD	Ubiquitin-binding domain
UBL	Ubiquitin-like domain
Ub_n	(n)-polyubiquitin chain
UCH	Ubiquitin C-terminal hydrolase
UIM	Ubiquitin-interacting motif
UPS	Ubiquitin-proteasome system
UV	Ultraviolet
VWA	von Willebrand factor domain A
WHD	Winged helix domain
XPC	Xeroderma pigmentosum complementation group C

Table of Contents

List of Figures

List of Tables

1 Introduction

Cells constantly synthesise and degrade proteins. The function of this seemingly wasteful cycle is twofold: (1) the accumulation of misfolded or otherwise damaged proteins could be harmful to the cell and (2) many regulatory proteins are only required under special circumstances or at a precisely defined interval in the life of a cell. The ability to precisely control protein lifespan depending on its current situation is of great value to a cell because it allows for adaptation to many different conditions. If there was no protein flux, fast and specific regulation of protein concentrations would not be possible.

In eukaryotic cells two major protein degradation mechanisms exist: the rather non-selective lysosomal pathway and the highly selective ubiquitin proteasome system (UPS). Lysosomes are organelles containing about 50 proteolytic enzymes that break down the contents of autophagic vacuoles (membrane-enclosed bits of cytoplasma) after fusing with them (reviewed in Saftig et al. 2009). Although there is some degree of regulation (for example the breakdown of only non-essential proteins during nutritional deprivation (Neely et al. 1974), only proteasomal degradation can be considered truly selective allowing the targeted degradation of a single protein when required. This thesis will exclusively discuss the UPS.

1.1 The Ubiquitin Proteasome System

It was initially assumed that all protein degradation occurs via lysosomes. However, when in 1953 M. Simpson demonstrated that protein breakdown requires energy the existence of an alternative degradation pathway was increasingly discussed (Simpson 1953). In the late 1970s, reticulocytes (which lack lysosomes yet are proteolytically highly active) became the model system to elucidate this alternative pathway. Etlinger and Goldberg were the first to establish a cell-free ATP-dependent proteolytic system based on reticulocyte lysate (Etlinger et al. 1977). Aaron Ciechanover, Avram Hershko and Irwin Rose subsequently identified fractions of reticulocyte lysate that contained different activities required for ATP-dependent proteolysis: a heat-stable ~9 kDa polypeptide and an ATP-stabilized high-molecular

weight compound (Ciechanover, Elias, et al. 1980; Hershko et al. 1979; Ciechanover et al. 1978). In subsequent experiments they found that the heat-stable protein was covalently attached via an isopeptide linkage to target proteins in an ATP-dependent manner and released upon their degradation (Hershko et al. 1980; Ciechanover, Heller, et al. 1980). In the following years, the same group identified the three enzymes (E1, E2, E3) required for the attachment of the heat-stable protein (that was later identified as ubiquitin) to the target protein (Hershko et al. 1983; Ciechanover et al. 1981). By working together in a sequential manner, these three classes of enzymes attach ubiquitin chains to a lysine residue of the target protein (Ciechanover 1994). The last component, the ATP-stabilized protease, was identified and characterised by M. Rechsteiner and colleagues and called the 26S proteasome (Hough et al. 1986).

In the first step of ubiquitin chain synthesis, a ubiquitin-activating enzyme (E1) forms a thioester bond between its active-site cysteine and the C-terminal glycine residue of ubiquitin. This activation step requires energy in the form of ATP to form a reactive ubiquitin-adenylate intermediate (Pickart 2001; Schulman et al. 2009). In a transthioesterification reaction the activated ubiquitin molecule is then transferred to the active-site cysteine of a ubiquitin conjugating enzyme (E2) (Kleiger et al. 2009; Ye et al. 2009). This class of enzymes is more diverse as the linkage between the ubiquitin moieties within a ubiquitin chain can be different, and is in most cases determined by the E2s. Finally, a ubiquitin-protein ligase (E3) forms an isopeptide bond between the carboxyl group of the C-terminal glycine residue of ubiquitin and the ε-amino group of a lysine residue of the target protein. E3 enzymes both bind to the E2 and recognise characteristic amino acid sequences on their target proteins called degrons (Alexander Varshavsky 2005) and thus impart specificity to the UPS. Two classes of E3s have been identified, which differ in their mechanism of isopeptide bond formation. RING (Really Interesting New Gene) finger E3s simultaneously recruit an E2 and the target protein thereby assisting the transfer of the activated ubiquitin to a lysine residue on the target protein. In contrast, HECT (Homologous to the E6-AP Carboxyl Terminus) E3s contain a conserved cysteine residue that forms an intermediate thioester bond with the ubiquitin C-terminus.

Therefore HECT E3s can promote the formation of a distinct linkage type even with a nonspecific E2 (M. Wang et al. 2006). While there are only one E1 and 13 E2s in budding yeast, hundreds of different E3s have been identified (reviewed in Ye et al. 2009). Any of the seven lysine residues of ubiquitin itself can be used as substrate resulting in the formation of ubiquitin chains with different linkage types determined by the E2 enzyme. Linkage via lysine 48 (K48) is the major signal for proteasomal degradation and a chain length of four ubiquitin molecules is the minimal degradation signal (Thrower et al. 2000). Finally, while the substrate is degraded, the ubiquitin is recycled by deubiquitylating enzymes (DUBs, reviewed in Reyes-Turcu et al. 2009) at the proteasome, closing the cycle (Figure 1.1).

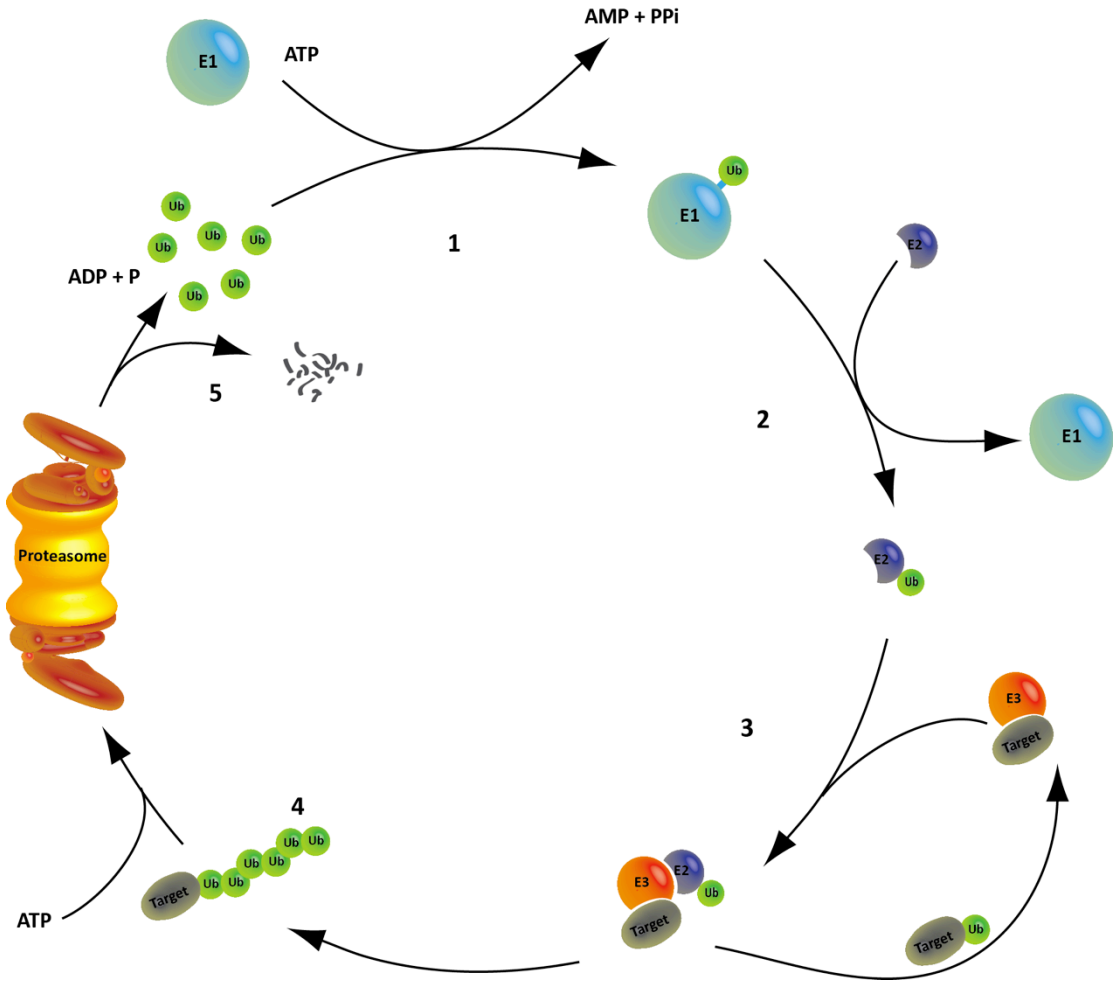


Figure 1.1. The ubiquitin proteasome system targets proteins to the proteasome for their degradation. In a first step (1) ubiquitin is activated by a ubiquitin-activating enzyme (E1). Subsequently, it is transferred to a ubiquitin-conjugating enzyme (E2, (2)). Ubiquitin-protein ligases (E3s) have a specific target protein, and together with E2 transfer the ubiquitin to one of its lysine residues (3). By repeating this last step, a ubiquitin chain is gradually synthesised. Once a K48-linked ubiquitin chain has reached a certain length of four ubiquitin moieties, it can be targeted to the proteasome (4) where the protein is degraded and the ubiquitin is recycled (5).

Ubiquitin

1.1.1 Ubiquitin Structure

Ubiquitin, as the name suggests, is a 76 amino acid protein that is found in every eukaryotic cell and was first isolated in 1975 by Goldstein and co-workers (Schlesinger et al. 1975). With only three amino acid changes between the yeast and human orthologs, it is also the most conserved eukaryotic protein (Figure 1.2A). This remarkable conservation is an indicator of its fundamental importance to the cell. Its functions derive from its use as a posttranslational modifier of proteins, thereby delivering a remarkable array of cellular signals.

Because of its involvement in many different processes and the resulting high cellular concentration, ubiquitin is expressed from several copies of the gene in a head-to-tail manner (Ozkaynak et al. 1984). The resulting N-C-linked polyubiquitin chain is then processed by a specialised protease to yield monoubiquitin. In order to prevent accidental activation, this chain is synthesised with an additional amino acid at the C-terminus.

The crystal structure of ubiquitin was first determined in 1987 at 1.8 Å (Vijay-Kumar et al. 1987, Figure 1.2). It revealed an extremely compact fold with about 87% of the polypeptide chain involved in hydrogen-bonded secondary structure explaining the extraordinary stability of ubiquitin. The only notable exception from this tight hydrogen-bond network is the C-terminus that remains rather flexible. This is e.g. necessary to allow efficient activation of G76. The secondary structure consists of a single α-helix tightly packed against a five-stranded β-sheet that seems to grasp the helix (β-grasp fold). The long loop connecting beta-strands 4 and 5 contains a short 3_{10} helix. L8, I44 and V70 form a solvent-exposed hydrophobic patch that was shown to be essential for proteasomal targeting (Beal et al. 1996) and was later revealed to be the major site involved in all ubiquitin recognition events.

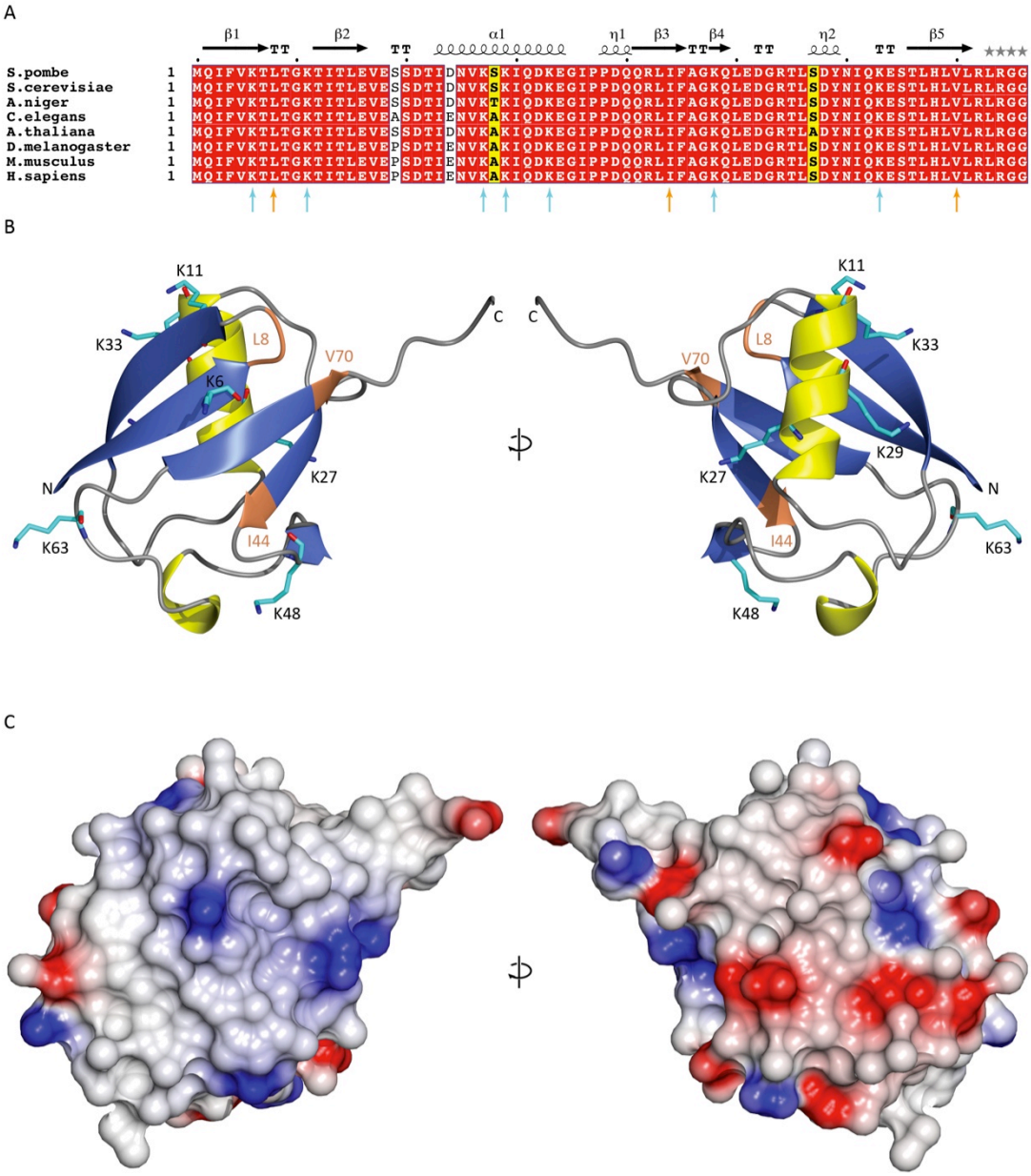


Figure 1.2. The structure of ubiquitin. (A) The sequences of eight ubiquitin orthologs were aligned with the ClustalW2 alignment software (Larkin et al. 2007). Similarity scores were calculated with the Risler matrix (Risler et al. 1988) and coloured accordingly. The alignment reveals the extraordinary conservation of ubiquitin even between yeast and human orthologs. Important residues are highlighted with arrows (lysine=cyan, hydrophobic binding patch=orange). **(B)** Cartoon of the ubiquitin crystal structure (PDBID 1UBQ) to show the tight packing of the secondary structure elements. Ubiquitin consists of a five-stranded β-sheet that folds around a central α-helix. A short 3_{10}-helix connects β-strands 4 and 5. The seven lysine residues are shown as cylinders and the core binding patch is highlighted in orange. **(C)** Ubiquitin surface electrostatics as calculated by CCP4MG (Potterton et al. 2004). The charge distribution is relatively well distributed apart from the hydrophobic patch, which can be easily identified in the left-hand panel. The view of the ubiquitin fold is the same in panels (B) and (C), and in each panel the views are rotated by 180°.

1.1.2 Ubiquitin as a Signalling Molecule

The diversity of the ubiquitin signal is remarkable. A single ubiquitin molecule attached to a target protein (monoubiquitylation) has been shown to play a role in such diverse events as DNA repair (Bergink et al. 2009), histone modification (Kodadek 2010), endocytosis (K Haglund 2003) and viral budding (Martin-Serrano 2007). When ubiquitin itself is ubiquitylated, ubiquitin chains are formed. Any of its seven lysine residues can be utilised resulting in chains again encoding different signals. K48-linked chains are the best-studied ubiquitin signal and target the substrate to the 26S proteasome (Chau et al. 1989). All other ubiquitin chains have been shown to exist *in vivo* but their precise functions are characterised to a much lesser extent (Ikeda et al. 2008).

Of the atypical ubiquitin chains K63-linked chains have been best characterised. They play a major role in NF-κB pathway signalling, endocytosis and DNA repair (Kaisa Haglund et al. 2005; Hayden et al. 2008). K29-linked ubiquitin chains are also involved in the proteasomal degradation of a subset of substrates (Johnson et al. 1995) and lysosomal degradation (Chastagner et al. 2006). The BRCA1-BARD1 complex synthesises K6-linked ubiquitin chains suggesting a role in DNA repair (Nishikawa et al. 2004). Ubiquitin chains linked through K11 have been shown to be critical for cell cycle regulation (Kirkpatrick et al. 2006) and were proposed to promote degradation of the anaphase-promoting complex/cyclosome (APC/C) (Jin et al. 2008). All of the different linkage types have been found *in vivo* and except for K63-linked ubiquitin chains their levels were inversely correlated with proteasomal activity suggesting an involvement in proteasomal targeting (Xu et al. 2009). Combinations of different E2s and E3s can also produce ubiquitin chains that contain several different isopeptide linkages. Substrates linked to these chains cannot be degraded by the proteasome possibly as a result of their reduced affinity to ubiquitin receptors (H. T. Kim et al. 2007).

The various linkage types in polyubiquitin chains restrain the ubiquitin moieties to adopt different orientations towards one another, which in turn allows efficient discrimination (see Section 1.3). The crystal structure of K48-linked

diubiquitin (K48-Ub_2) engages in a “closed” conformation with the two hydrophobic patches facing each other (Cook et al. 1992). NMR studies have confirmed this conformation in solution and at neutral pH (Figure 1.3A). In acidic solutions, however, the diubiquitin molecule adopts a more open conformation (Figure 1.3B, Varadan et al. 2002). This is probably caused by the protonation of H68. The interaction between the two hydrophobic patches is rather weak and dynamic (Varadan et al. 2002) and enables efficient binding to ubiquitin receptors.

Two different K48-linked tetraubiquitin (K48-Ub_4) crystal structures have been determined (Cook et al. 1994; Phillips et al. 2001). In both structures inter-moiety contacts are prevalent supporting the solution structure. This observation supports the hypothesis that the K48-Ub_2 building block is the motif recognised by a K48-specific receptor.

The isopeptide bond in K63-linked polyubiquitin chains restrains the ubiquitin moieties prohibiting contacts of the hydrophobic patches. NMR studies (Tenno et al. 2004; Varadan et al. 2004) and the crystal structures of K63-linked ubiquitin chains (Komander et al. 2009; Datta et al. 2009) revealed that these chains adopt an extended conformation that lacks any interaction between the individual molecules apart from the linkage (Figure 1.3E). The different orientations of the proximal (the moiety with a free C-terminal G76) towards the distal moiety in these structures indicate a high degree of rotational freedom around the isopeptide bond. The conformation of N-C-linked polyubiquitin is virtually equivalent to that of K63-linked polyubiquitin chains (Figure 1.3F). Both molecules form a structure in which each individual ubiquitin moiety can be regarded as an individual unit that is rotationally unrestrained and highly flexible (Komander et al. 2009). Interestingly, even though there is some functional redundancy between these two linkage types, there is evidence for molecular discrimination (Komander et al. 2009).

The structure of K11-linked diubiquitin (K11-Ub_2) is different from both K48- and K63-Ub_2 conformations. It is less extended than that of N-C-linked or K63-linked diubiquitin and in contrast to K48-Ub_2, I44 is solvent-exposed (Figure 1.3D, Bremm et al. 2010). The structure of K6-linked diubiquitin (K6-Ub_2) has recently been released

(PDBID 2XK5) and exhibits yet another relative orientation of the two moieties (Figure 1.3C). The hydrophobic patch of the distal moiety is solvent-exposed and the surface of the region C-terminal to the α-helix and perpendicular to the hydrophobic patch occludes the hydrophobic patch on the proximal ubiquitin moiety. Structural information for the other non-canonical polyubiquitin chains (linked via K27, K29 or K33) is not available. However, models for their structures have been proposed based on molecular modelling approaches (Fushman et al. 2010). Based on these studies, Fushman and colleagues assigned the differently linked polyubiquitin chains to two groups. The first group is characterised by a "closed" conformation as observed in K48-linked polyubiquitin chains and includes chains linked through K6, K11 and K27. Chains in the second group are proposed to adopt an extended structure as reported for K63- and N-C-linked polyubiquitin and also include chains linked through K29 and K33. However, subsequent determination of structures for K11-Ub_2 and K6-Ub_2 suggested that a re-appraisal of the models was required. It remains to be seen whether crystal packing contributes to the observed differences between the experimental structures and the solution conformations proposed by modelling studies. However, given the close agreement between the K63- and K48-Ub_2 crystal and solution structures, this seems less likely.

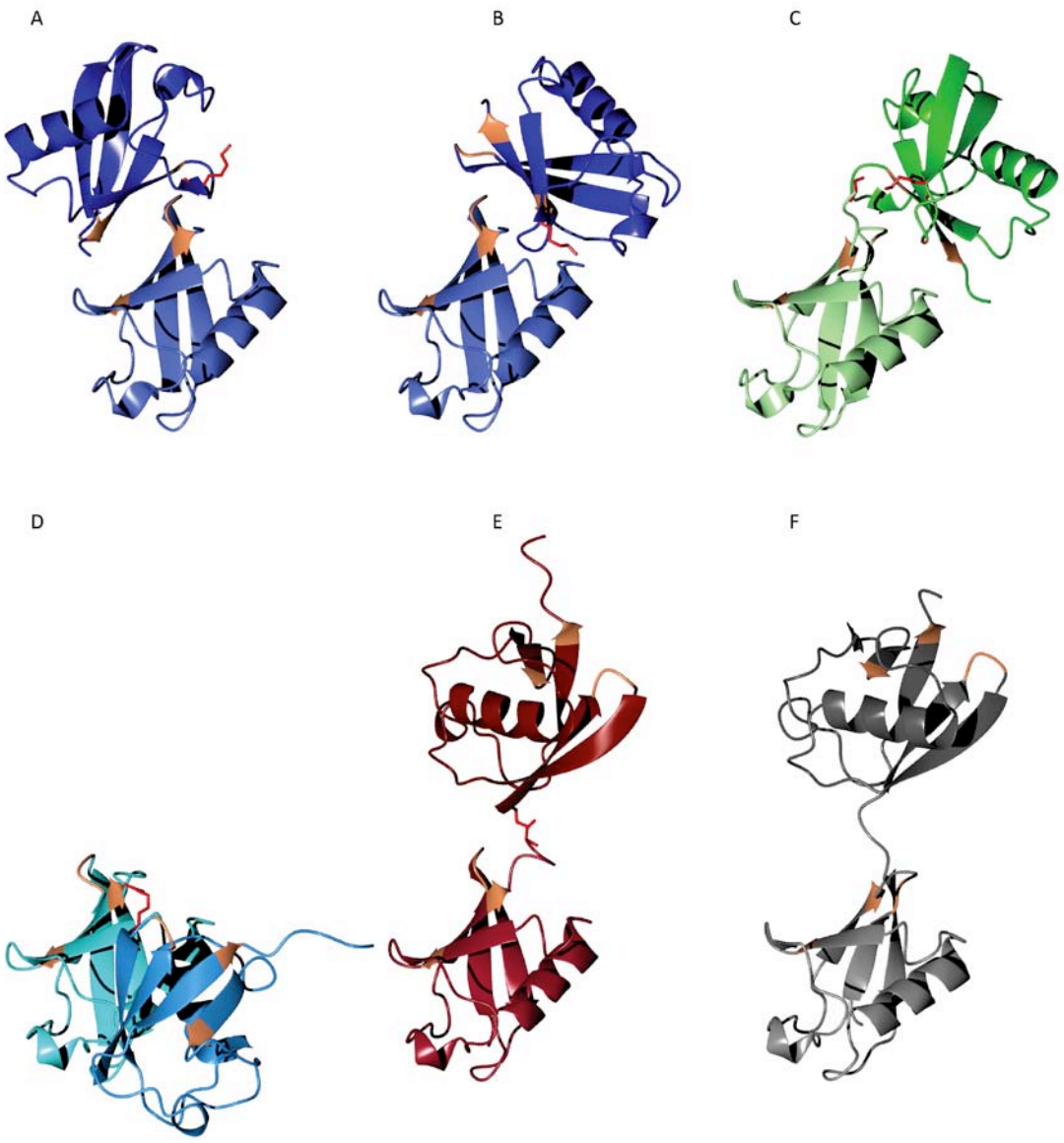


Figure 1.3. The conformations of different ubiquitin chains. The distal ubiquitin moiety is drawn in a lighter shade and in the same orientation in each panel. The hydrophobic patch is coloured orange and the isopeptide bond is highlighted as red cylinders. **(A)** At physiological pH, K48-Ub_2 (blue) adopts a closed conformation (PDBID 2PEA), while at acidic pH **(B),** the inter-domain contacts are lost (PDBID 2PE9). **(C)** In K6-Ub_2, the hydrophobic patch of the proximal moiety is occluded by a region C-terminal to the α-helix perpendicular to the hydrophobic patch (PDBID 2XK5). **(D)** K11-Ub_2 (cyan) has an open structure (PDBID 2XEW) but is much less extended than K63-linked ubiquitin (brown, PDBID 2JF5, **(E)**) and NC-Ub_2 (grey, PDBID 2W9N, **(F)**). In both structures the hydrophobic patches face almost opposite directions.

1.2 The Proteasome

The proteasome was discovered as the ATP-stabilised protease of the ubiquitin proteasome system (Hough et al. 1986). Our knowledge about this 2.5 MDa protein complex has grown considerably and apart from its proteolytic activity it is now known to recognise and unfold substrates and also remodel and recycle ubiquitin chains (Finley 2009). It consists of a 20S core particle (CP) that is activated by association with activator complexes at the ends of the proteolytic chamber (Figure 1.4).

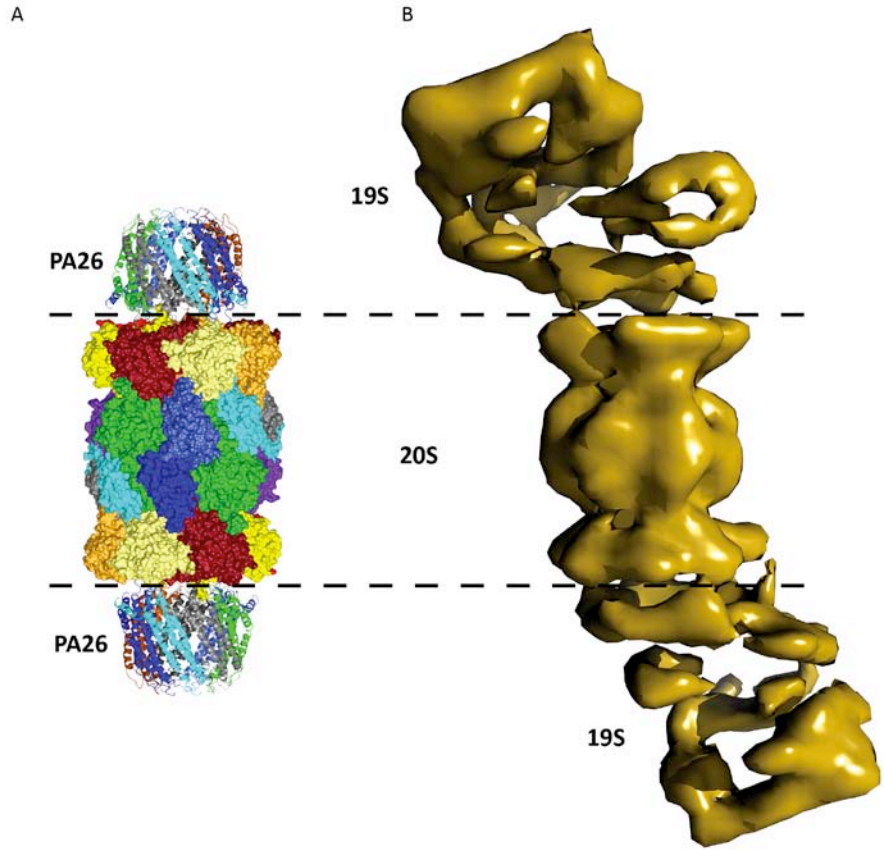


Figure 1.4. The proteasome. (A) The 20S CP in complex with the trypanosomal activator PA26 (PDBID 1FNT). **(B)** Cryo-EM reconstruction of the 26S proteasome. The well-defined 20S CP can be easily identified. Image based on (Nickell et al. 2009).

Four proteasomal activators REGγ, PA200/Blm10, PA28 and the 19S regulatory particle (RP) have been identified to date. These activators can bind at one end of the CP while the 19S RP still occupies the other (Shibatani et al. 2006; Tanahashi et al. 2000; Cascio et al. 2002; Schmidt et al. 2005). REGγ and PA28 are 25% identical.

However, while PA28 is exclusive to vertebrates, REGγ is also found in invertebrates and it has been speculated that PA28 evolved from REGγ (Masson et al. 2001). Hybrid proteasomes containing one 19S RP and one PA28 promote the production of antigenic peptides and it has been shown that peptides generated by the proteasome leave the particle through the same channel as used for entry (Köhler et al. 2001). Mutants that artificially open the CP result in an increased size of the peptides that exit the proteasome. This is the result of a shorter dwell time in the catalytic chamber. Opening of the CP by PA28 might thus further reduce this time by providing a larger exit channel that increases the average released peptide to a size appropriate for antigen presentation (Whitby et al. 2000; Köhler et al. 2001). PA200 is conserved from yeast to humans and, in contrast to the other activators, it is a monomeric ~245 kDa protein that lies like a turban on top of the cylinder end of the CP (Iwanczyk et al. 2006). It is required for optimal proteasomal function and spermatogenesis (Schmidt et al. 2005). The main form of the holoenzyme, the 26S proteasome, is formed by a 19S RP attached to each end of the CP (Figure 1.4B) and will be the subject of this thesis.

1.2.1 The 20S Core Particle

The crystal structure of the 20S core particle (CP) provides a detailed view of its architecture and function (Löwe et al. 1995). It is a barrel-like structure consisting of four seven-membered rings of subunits. The two central rings built of β-subunits are flanked by an α-ring on both sides (Figure 1.5A, C). These rings enclose a cavity of approximately 100 Å axially and 60 Å along the orthogonal axis. Although the sequences of the α- and β-subunits are quite different, they share a similar fold. This fold is also found in the bacterial ClpQ protease that (as a dodecamer), together with ClpY (a hexameric ATPase), forms a compartmentalised protease for the degradation of misfolded proteins. This conservation of fold suggests that they all might have evolved from a common ancestor. The proteolytically active sites are located in the β1, β2 and β5 subunits that face the central chamber and are therefore efficiently shielded from the cytoplasm (Figure 1.5B). They belong to the N-terminal nucleophile (Ntn) class of hydrolases in which the side chain of the N-terminal

residue (in this case threonine) acts as nucleophile while the N-terminus is the proton acceptor. After the formation of the catalytic tetrahedral intermediate by nucleophilic attack of the carbonyl carbon atom on the scissile peptide bond, this bond is cleaved and the intermediate acyl-threonine is hydrolysed by a water molecule. The three active subunits cleave a broad range of peptide substrates with slightly differing specificities. Subunit β1 cleaves on the C-terminal side of acidic residues, β2 after tryptic residues and β5 after hydrophobic residues (Borissenko et al. 2007). These complementary activities ensure the efficient degradation of (probably) any protein that enters the cavity (P. Chen et al. 1996). In mammals subunits β1, 2 and 5 can be substituted by γ-interferon-inducible β-subunits (M Gaczynska et al. 1993). These immuno-subunits allow the proteasome to more efficiently produce antigenic peptides that are then loaded on MHC class 1 proteins to initiate immune responses. This substitution requires the *de-novo* assembly of the proteasome.

The assembly of the 20S CP is not trivial, as four hetero-heptameric rings of α- and β-subunits have to be formed and stacked on top of each other. Five molecular chaperones have been identified that help to effectively achieve this task (reviewed in Murata et al. 2009). Moreover, in order to avoid indiscriminate degradation of cellular proteins, the proteolytic β-subunits must only become active after the successful formation of the mature CP. In a first step the α-ring is assembled with the help of the hetero-dimeric PAC3-PAC4 chaperone that binds to subunit α5 and guides the assembly of the other α-subunits. Before the subsequent addition of β-subunits the PAC3-PAC4 complex has to disassemble. The PAC1-PAC2 complex also assists the proper organisation of the α-ring that, in effect, serves as a scaffold for β-ring assembly. To ensure that the β-subunits are correctly incorporated, the chaperone UMP1 prevents the association of two half-CPs until all seven β-subunits are correctly assembled (Li et al. 2007). The catalytic β-subunits and the non-catalytic β6 and β7 subunits are incorporated as pro-peptides with an N-terminal precursor sequence. These N-termini help with the dimerisation of the half-proteasomes to the 20S CP and are removed after successful assembly to activate the proteolytic activities of subunits β1, β2 and β5.

Access to the CP in the apo state is limited to small peptides and the opening of the entry channel is strictly regulated. The N-termini of the α-subunits converge at the central axis closing the entry channel to the central cavity (Figure 1.5C, Groll et al. 2000). At the RP-facing side, the α-rings contain pockets located at the interfaces between neighbouring α-subunits. Occupation of these pockets by the C-termini of the six Rpt subunits of the RP (see Section 1.2.2) results in the opening of a substrate translocation channel (Figure 1.5D, Whitby et al. 2000). Once opened, substrates must pass through the RP (and be unfolded) in order to gain access to the proteolytic chamber. This compartmentalisation and the finely tuned assembly of the proteasome make sure that only proteins that have been properly processed by the RP can gain access to the proteolytic chamber. ATP-dependent compartmentalised proteases have been characterised in all domains of life. The 26S proteasome is the most complex example but this increased complexity seems to reflect an adaptation to ubiquitin recognition and processing while conserving substrate denaturation and proteolysis.

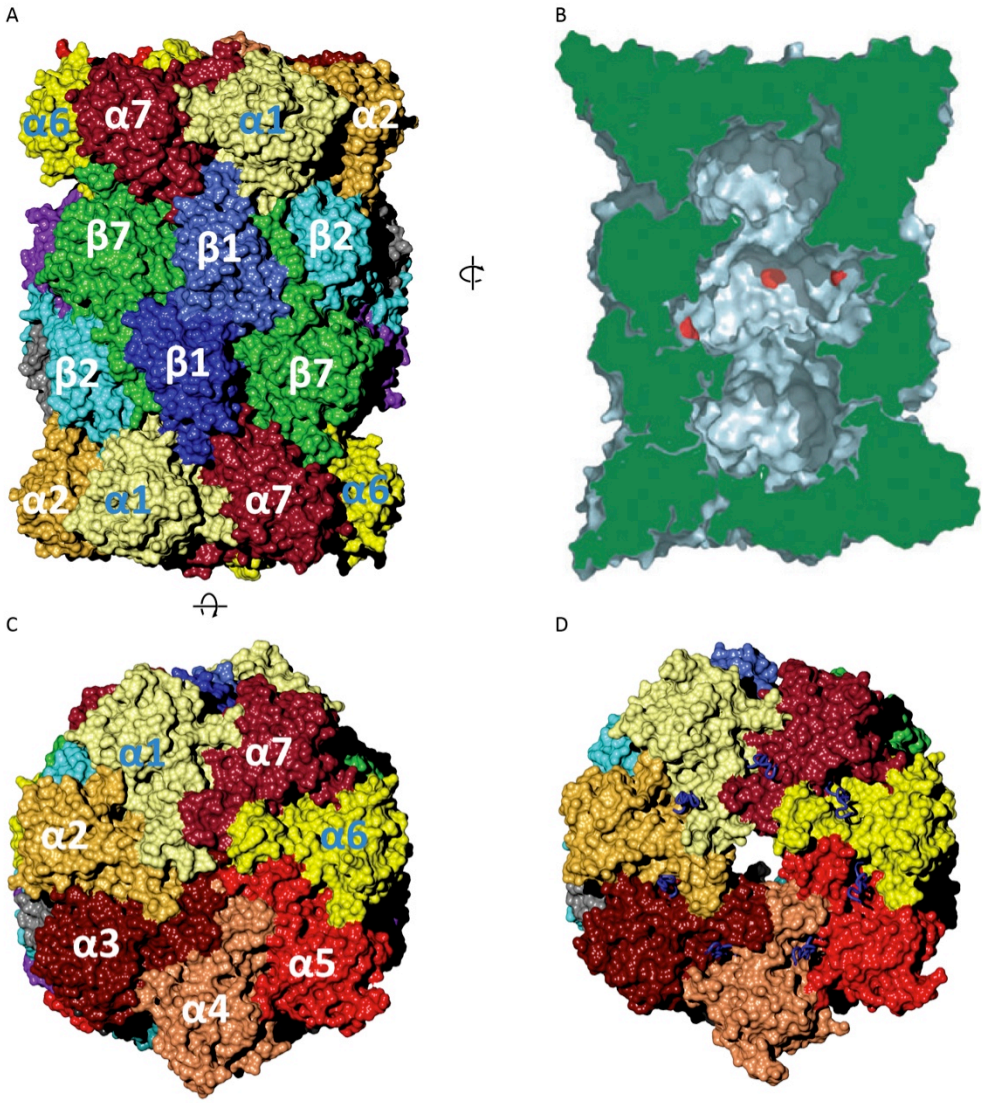


Figure 1.5. The 20S core particle. (A) Surface representation of the *S. cerevisiae* 20S core particle (1RYP). Equivalent subunits are in the same colour. **(B)** Tight packing of the subunits against each other ensures that the sites of proteolytic activity (highlighted in red) are contained within the 20S CP catalytic chamber**. (C)** View along the pseudo sevenfold symmetry axis. The N-termini of the α-subunits converge at the centre closing the channel. **(D)** Binding of an activator (in this case trypanosomal PA26, PDBID 1FNT) opens the channel by insertion of the C-termini (blue worms) of its seven subunits into pockets between adjacent α-subunits.

1.2.2 The 19S Regulatory Particle

The composition of the 19S RP was initially studied by gradient SDS-PAGE following its purification from *S. cerevisiae* (Glickman, Rubin, V A Fried, et al. 1998). 17 proteins that form the RP were identified. Subsequent mass spectrometry of 1-step affinity-purified 26S proteasomes revealed many more proteasome associated proteins that were grouped into four classes (Verma et al. 2000). The first class

contained UPS-associated proteins such as the ubiquitin ligases SCF and APC, the second class the deubiquitylating enzyme (DUB) Ubp6 and the novel proteasomal subunit Rpn13. Chaperones such as Hsc70, Hsc82 and Pac2 were placed in the third class and miscellaneous proteins such as transcription factors and ribosomal proteins into the fourth. The association of such a diverse array of factors with the RP makes it difficult to discriminate between proteasomal and proteasome-associated proteins. A summary of all subunits currently thought to be *bona-fide* 19S subunits can be found in Table 1.1.

19S RPs purified from a yeast strain from which the non-essential subunit Rpn10 had been deleted were found to be unstable, falling apart into two subcomplexes, subsequently called the base and the lid (Glickman et al. 1998). The base contains ten subunits of which six ATPases belonging to the ATPases Associated with diverse cellular Activities (AAA) family (Rpt1-Rpt6) form a ring with pseudo sixfold symmetry. This ring interlocks with the pseudo sevenfold symmetry of the α-subunits of the 20S core particle. Rpn1 and Rpn2 are large (>100 kDa) α-helical proteins containing nine Proteasome/Cyclosome (PC) repeats preceded by two more divergent repeats (Glickman et al. 1998) and they have been predicted to form toroidal structures (Kajava 2002). Atomic force microscopy studies have supported a model in which Rpn1 and Rpn2 form concentric solenoid rings stacked on each other (Rosenzweig et al. 2008). However, cryo electron microscopy (EM) data of 26S proteasomes does not support this model (Nickell et al. 2009). Until higher-resolution cryo EM models are available, this crucial part of the 19S organisation remains contentious. Interestingly, all known Rpn1 and Rpn2 binding partners interact with ubiquitin. For example Rpn13 is a proteasomal ubiquitin receptor located in the base and binds to Rpn2 while also serving as a receptor for the deubiquitylating enzyme Uch37 (Husnjak et al. 2008). Rpn10 has been reported to interact with both Rpn1 (Seeger et al. 2003) and Rpn12 (C. R. M. Wilkinson et al. 2000) placing it at the interface between base and lid. Proteins of the UBL-UBA class (for example the *S. cerevisiae* proteins Dsk2 and Rad23 and their *S. pombe* orthologs Dph1 and Rhp23) also reversibly interact with Rpn1 and Rpn2 via their UBL domain

(Elsasser et al. 2002). These findings support a role for Rpn1/Rpn2 as a platform for ubiquitin processing.

The lid shares its architecture and domain organization with two other large eukaryotic protein complexes, the COP9 signalosome (CSN) and the eukaryotic initiation factor 3 (eIF3), collectively called PCI (Proteasome, COP9, Initiation factor 3) complexes. For every proteasomal lid subunit a paralogous CSN subunit exists. Despite intensive research, however, only very limited structural information about those important regulatory complexes is available. Several yeast-2-hybrid and pull-down studies have mapped protein-protein interactions within the RP (summarised in Fu et al. 2001). However, both of these methods are prone to false-positive results because they do not discriminate well between direct and indirect interactions. By mass spectrometry two major proteasomal lid subcomplexes consisting of Rpn5/6/8/9 and Rpn3/7/Sem1, connected via Rpn3-Rpn5 could be identified (Sharon et al. 2006). In a more recent study employing the same approach, the CSN was also found to be composed of two modules, Csn1/2/3/8 and Csn4/5/6/7, connected by interactions between Csn1 and Csn6 (Sharon et al. 2009). EM studies have compared the three PCI complexes (Enchev et al. 2010) and a reconstruction of the 19S RP revealed the boundaries of the AAA-ATPase ring (Nickell et al. 2009). The location for Rpn10 was also proposed based on its sub-stoichiometric assembly with the proteasome and assembly modelling. According to this model, Rpn10 is located near the mouth of the ATPase ring (Figure 1.6).

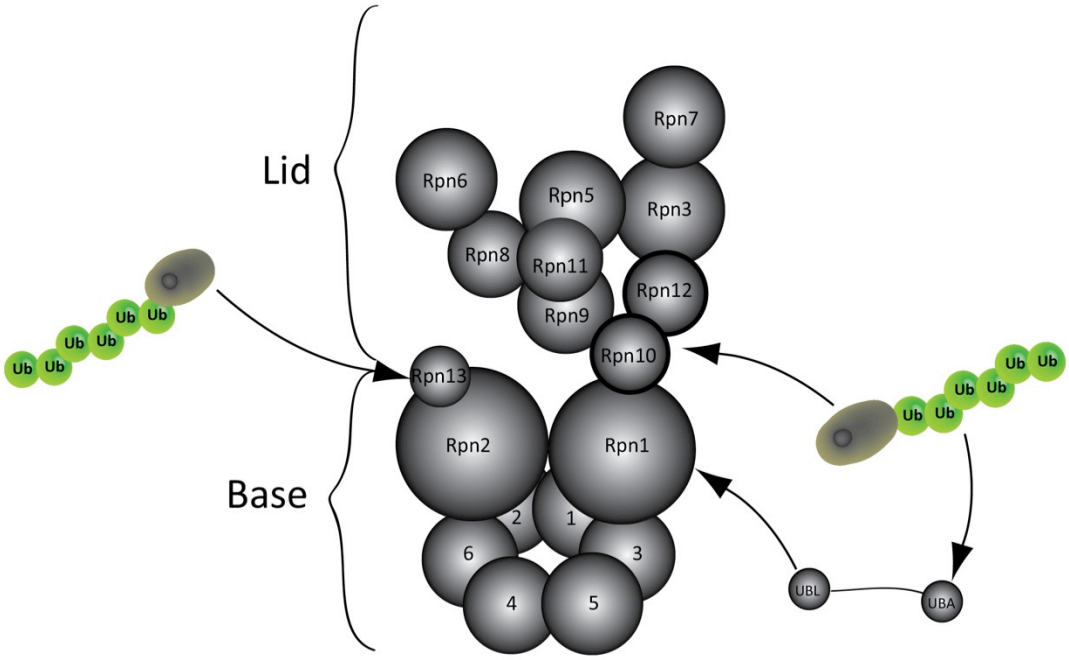


Figure 1.6. Model of the 19S RP subunit organisation based on reported protein-protein interactions. The 19S regulatory particle can be divided into base and lid sub-particles. All ubiquitin receptors identified so far interact with the base. The base also contains the unfolding activity of the proteasome located in the ring of ATPases. The function of the lid remains obscure as the only activity associated with it is that of the DUB Rpn11.

In higher eukaryotes two subunits of the 19S RP contain an MPN (Mpr1-Pad1-N-terminal) domain whereas six comprise the hallmark PCI domain (Scheel et al. 2005). PCI domains serve as scaffolding units allowing the efficient assembly and maintenance of large protein complexes. They consist of a tetratrico-peptide repeat (TPR)-like N-terminus and a C-terminal winged helix domain (WHD).

Rpn11 is the only proteasomal lid subunit with known catalytic function. It is a metalloprotease with specificity for the proximal ubiquitin moiety. Interestingly, unfolding of substrates seems to be tightly coupled to deubiquitylation by Rpn11 and it can be speculated that the sole function of the lid is to support Rpn11 and guarantee the coupling of deubiquitylation and degradation (T. Yao et al. 2002; Verma et al. 2002). Two other deubiquitylating enzymes, Uch37 and Ubp6, associate with the proteasome via Rpn13 and Rpn1, respectively (Guterman et al. 2004; Stone et al. 2004). The association of a deubiquitylating enzyme (Uch37) with a proteasomal ubiquitin receptor is interesting, as they possess opposing activities with respect to degradation. The fact that Rpn13 binds monoubiquitin so strongly

with a preference for the proximal moiety (Husnjak et al. 2008) might indicate that its function is not to bind substrates for degradation but to position them for efficient deubiquitylation by Uch37. Interestingly, full-length Rpn13 shows a 26-fold decreased affinity for monoubiquitin compared to its isolated PH (Pleckstrin Homology) Receptor for Ubiquitin (Pru) domain (see Section 1.3) as a result of intramolecular interactions that prevent efficient ubiquitin binding (X. Chen et al. 2010). Binding to Rpn2 disrupts this intra-molecular "masking" and results in an affinity towards ubiquitin that is identical to that of the isolated Pru domain.

Table 1.1. Subunits of the 19S regulatory particle.

	S. cerevisiae	*S. pombe*	*H. sapiens*	Size [kDa]	Domain	Activity
Base	Rpt1	PRS7	S7	48.6		AAA ATPase
	Rpt2	Mts2	S4	49.2		AAA ATPase
	Rpt3	PRS6B	S6	47.3		AAA ATPase
	Rpt4	PRS10	S10	44.1		AAA ATPase
	Rpt5	SpRpt5	S6'	49.1		AAA ATPase
	Rpt6	Let1	S8	45.7		AAA ATPase
	Rpn1	Mts4	S2	100.2		Scaffold
	Rpn2	SpRpn2	S1	105.9		Scaffold
	Rpn13	SpRpn13		40.7	Pru	Ubiquitin receptor
	Rpn10	Pus1	S5a	42.1	VWA, UIM	Ubiquitin receptor
Lid	Rpn3	SpRpn3	S3	61.0	PCI	
	Rpn5	SpRpn5		52.9	PCI	
	Rpn6	SpRpn6	S9	47.4	PCI	
	Rpn7	SpRpn7	S10a	45.5	PCI	
	Rpn8	SpRpn8	S12	37.1	MPN	
	Rpn9	Mts1	S11	42.9	PCI	
	Rpn11	Pad1	S13	34.6	MPN	deubiquitylation
	Rpn12	Mts3	S14	30.0	PCI	
	Rpn15	SpRpn15	S15	24.6		

1.3 Ubiquitin Recognition

Ubiquitin is a very diverse cellular signal and in order to avoid convolution of the different signalling pathways (mediated by mono- and the different poly-ubiquitylated substrates) ubiquitin receptors need to be highly selective and able to discriminate a particular ubiquitin species against a background of a cellular concentration of free ubiquitin of up to 10 μM (Hicke et al. 2005). The multimeric nature of ubiquitin chains can be exploited by a ubiquitin receptor with several ubiquitin-binding domains (UBDs). Avidity effects arise from an orientation of the UBDs towards each other that allows the simultaneous binding of several ubiquitin moieties. In addition to a higher UBD concentration (compared to a single UBD in a

receptor protein), this also effectively increases the local concentration of ubiquitin moieties around the UBDs. The binding of the first ubiquitin moiety has a positive pseudo-cooperative effect on the binding of the second moiety. This effect allows a polyubiquitin receptor to have weak affinity towards monoubiquitin to avoid being constantly occupied by a non-substrate while at the same time binding strongly to polyubiquitin chains. A similar effect is achieved by double-sided UBDs that contain binding patches on different sides of their surface such as the Rad23 UBA2 domain (Figure 1.7A). These mechanisms have been identified in tandem ubiquitin interacting motifs (UIMs), double-sided UIMs and UBA domains (see below). Because the linkage via different isopeptide bonds in polyubiquitin chains may constrain the ubiquitin moieties in their relative orientations (Figure 1.3), recognition of the relative position of the ubiquitin moieties or of the isopeptide bond to the hydrophobic patch provides another way of discriminating between mono- and polyubiquitin and the different linkage types. For example the DUB AMSH-LP recognises Q62 and E64 of the proximal ubiquitin moiety to specifically cleave K63-linked ubiquitin chains (Y. Sato et al. 2008).

UBDs are a large and diverse group that can bind and often distinguish different ubiquitin modifications. To date, more than 20 different domains have been identified that can be organised into four groups: (i) α-helical, (ii) zinc-finger (ZnF), (iii) plekstrin homology (PH) and (iv) ubiquitin-conjugating-like (Ubc) domains.

UIM, IUIM (inverted UIM) and ubiquitin-binding zinc finger (UBZ) domains contain a single α-helix that contacts the ubiquitin hydrophobic patch in a parallel or antiparallel way. Selectivity of linkage type and chain length has only been demonstrated in this motif by exploiting avidity effects generated as a result of either double-sided UIMs (Figure 1.7B) that bind one ubiquitin hydrophobic patch on each side of the helix as found for example in Hrs (Hirano et al. 2006a) or two sequential UIMs such as in Rap80 (Figure 1.7C) and S5a (Sims and R. E. Cohen 2009; N. Zhang et al. 2009). The exploitation of these avidity effects by the tandem UIMs of Rap80 has been studied in detail (Walters et al. 2009; Y. Sato et al. 2009; Sims and R. E. Cohen 2009). The linker connecting the two UIMs orients them perfectly for avid

interaction with two neighbouring ubiquitin moieties of a K63-linked ubiquitin chain. In contrast, ataxin 3 contains a two-residue spacer that confers K48-linkage selectivity. Rap80 can be made K48-selective by exchanging its linker sequence with that of ataxin 3 demonstrating that it is indeed only the spacing of the two domains that promotes selectivity (Sims and R. E. Cohen 2009).

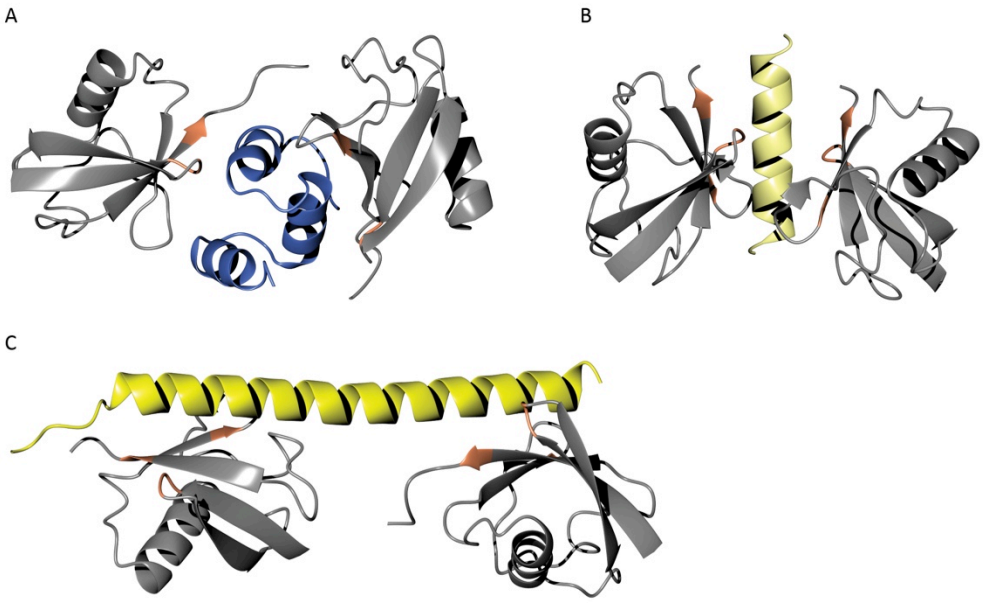


Figure 1.7. Multivalent interactions between UBDs and ubiquitin. Ubiquitin is drawn in grey with the hydrophobic patch highlighted in orange. **(A)** The Rad23 UBA2 domain (blue) exhibits selectivity for K48-linked ubiquitin chains. This is a result of an increased binding surface that includes the surface opposite the canonical ubiquitin-binding site (PDBID 1ZO6). **(B)** The Hrs UIM also uses two sides to interact with ubiquitin (PDBID 2D3G). **(C)** Rap80 contains two UIMs that are perfectly spaced for the simultaneous interaction with K63-Ub_2 resulting in a strong discrimination against K48-Ub_2 (PDPID 3A1Q).

Ubiquitin-associated (UBA) domains contain three short helices that form a tight bundle (Mueller et al. 2002; Dieckmann et al. 1998). Typically helices 1 and 3 each contain a conserved hydrophobic motif of LL and MGF, respectively, that form a hydrophobic patch to which the ubiquitin binds (Figure 1.8A). UBA domains exist in a number of proteins and they exhibit a wide range of affinities towards monoubiquitin and ubiquitin chains. According to their different ubiquitin-binding properties, they have been grouped into four different classes (Raasi et al. 2005). Class 1 exhibits specificity for K48-linked chains (Rad23 UBA2), class 2 is K63-selective (E2-25K, Drm2), class 3 contains the so-called “orphan” UBAs that do not

(or very weakly) bind ubiquitin and class 4 UBA domains bind monoubiquitin very strongly (Dsk2, UQ1). However, as this study employed GST-fusion proteins and surface plasmon resonance (SPR) experiments to determine the affinities, the dimerisation of GST compromises the results by artificially increasing the affinity through avidity effects. Indeed, a later study demonstrated that GST dimerisation positions the UBA domains in a way that matches the spacing of the ubiquitin moieties in K63-linked ubiquitin chains (Sims, A. Haririnia, et al. 2009). Taking into account this later analysis, no K63-selective UBA domains have been identified so far.

Not only α-helices are involved in ubiquitin recognition. The β-sheet of UBCH5C (Figure 1.8D), a deubiquitylating enzyme, interacts with the hydrophobic patch of ubiquitin as does the GRAM-like ubiquitin-binding in EAP45 (GLUE) domain (Figure 1.8C). This domain folds into a split PH domain and, in addition, uses residues from a loop and an α-helix (Hirano et al. 2006b; Alam et al. 2006; Brzovic et al. 2006) to bind ubiquitin. The proteasomal ubiquitin receptor Rpn13 contains a PRU domain that binds ubiquitin in a completely unrelated manner (Figure 1.8E). This domain employs three loops that contact a large part of the ubiquitin surface and form a hydrogen bond with ubiquitin H68 (Schreiner et al. 2008).

These structures show that even UBDs with the same domain organisation can exploit different strategies to bind ubiquitin but that the hydrophobic patch always provides the contact from the ubiquitin side. This might be a mechanism to limit crosstalk between different ubiquitin signalling pathways (Dikic et al. 2009). A notable exception to this rule are the various zinc-finger (ZnF) domains that bind ubiquitin at three different sites (Figure 1.8B): NZF domains bind the hydrophobic patch, the A20-type ZnF domain of RAB5 guanine nucleotide exchange factor (RABEX5) recognizes a polar surface of ubiquitin that is centred on D58 (S. Lee et al. 2006) and the DUB isopeptidase T (IsoT) binds to the C-terminus of ubiquitin (Reyes-Turcu et al. 2006).

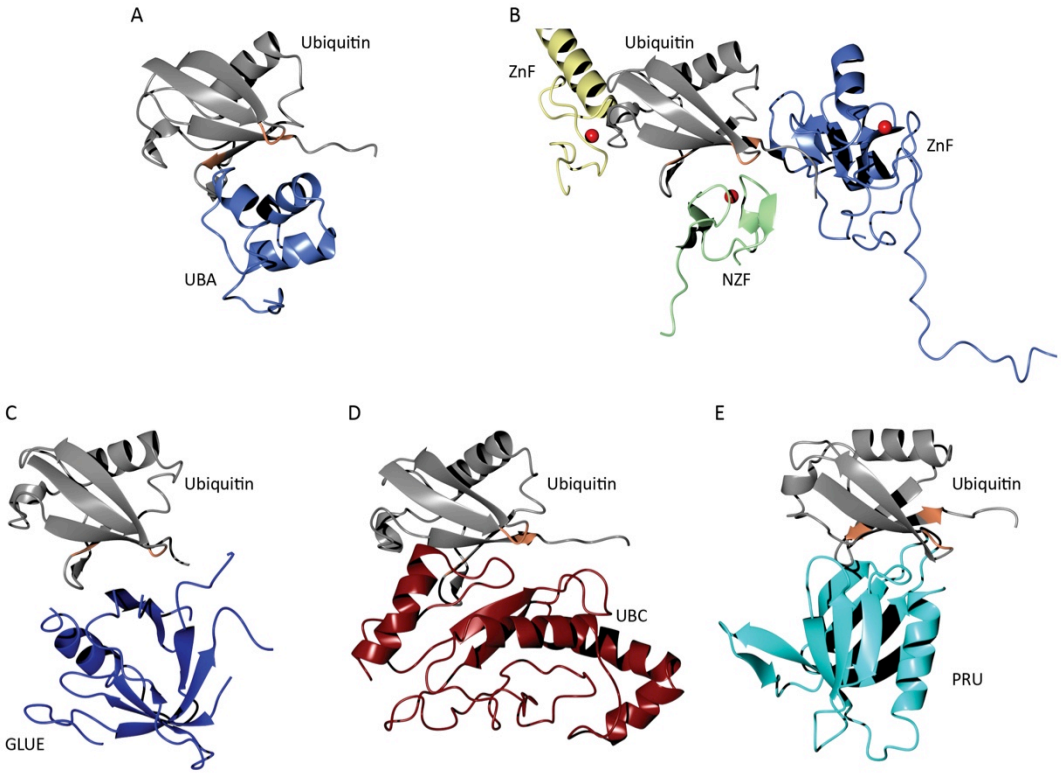


Figure 1.8. Ubiquitin is recognised by a range of structurally different domains. Cartoon representation of the interactions of various UBDs with ubiquitin. In all but one case UBDs interact with the hydrophobic patch (orange) of ubiquitin (grey). **(A)** UBA domain of protein linking IAP with cytoskeleton 1 (PLIC1; light blue, PDBID 2JY6) in complex with ubiquitin. **(B)** The A20-type ZnF domain of RAB5 guanine nucleotide exchange factor (yellow, PDBID 2FIF), nuclear protein localization 4 (NPL4) NZF domain (green, PDBID 1Q5W) and the ZnF domain of the DUB isopeptidase T (light blue, PDBID 2G45) contact all different surfaces of the ubiquitin molecule. **(C)** GRAM-like ubiquitin-binding in EAP45 (GLUE) domain of ELL-associated protein of 45 kDa (EAP45; also known as VPS36) (blue, PDBID 2DX5). **(D)** E2 ubiquitin-conjugating enzyme UBCH5C (brown, PDBID 2FUH). **(E)** Pru domain of Rpn13 (cyan, PDBID 2Z59).

Ubiquitin recognition by the proteasome was first studied by competition experiments using synthesised K48-linked ubiquitin chains of defined length (Thrower et al. 2000). The K_i for the inhibition of the degradation of the model substrate Ub_5-DHFR decreased about 100-fold from >15 μM for Ub_2 to 0.17 μM for Ub_4 but only sevenfold from there to 26 nM for Ub_8. This result suggested that the minimal degradation signal is K48-Ub_4. For these experiments the deubiquitylating enzyme Uch37 was inhibited specifically so that only the effect of ubiquitin chain length on substrate association with the proteasome and not the combined effects of binding and deubiquitylation was measured. This is important as the degradation rate is dependant on the stability of the substrate clearly demonstrating that the

rate-limiting step during proteasomal degradation is unfolding. Experiments that artificially located substrates at the proteasome by fusing chemically inducible affinity tags to the substrate and 19S subunits demonstrated that localisation at the 19S is sufficient for degradation (Janse et al. 2004). Thus, the proteasome employs an apparently futile cycle by containing ubiquitin receptors that capture substrates for degradation and deubiquitylation enzymes that trim ubiquitin chains. Furthermore, the ubiquitin ligase, Hul5, that can elongate K48-linked chains has been shown to associate with Rpn1 (Crosas et al. 2006). Collectively, these activities provide a dynamic environment for the modification of ubiquitin chains that could allow bound substrates to be differentially regulated. Therefore, it is currently thought that substrate degradation is not triggered by a single binding event but that it is the overall time a substrate spends at the proteasome that increases the likelihood of its degradation.

1.4 Aims of the Thesis

Ubiquitin signalling has attracted a lot of attention during the last decade and the number of studied ubiquitin receptors has increased substantially. In case of K48-linked polyubiquitin signalling it has become apparent that the process of targeting substrates to the proteasome does not involve a single recognition event that inevitably leads to substrate degradation but it is instead composed of a multi-layered network of ubiquitin receptors that allows different routes to be taken from the ubiquitylation event to the proteasome. This is illustrated by the number of different ubiquitin receptors and existing “shortcuts” via the direct association of ubiquitin ligases with the proteasome. This thesis aims to address ubiquitin recognition by the proteasome and 19S RP organisation by structural and protein-protein interaction studies on *S. pombe* orthologs Rpn10, Rpn12 and Rhp23.

The first proteasomal ubiquitin receptor to be identified was Rpn10. Early studies also found that it plays an important role in maintaining the structural integrity of the 19S RP. These two functions are located on different domains of the protein. The von-Willebrand-factor A (VWA) domain anchors Rpn10 to the proteasome while the UIM binds ubiquitin. K48-linked ubiquitin chains are

acknowledged as the signal for proteasomal degradation but the discrimination of ubiquitin receptors against other linkage types has not been studied in great detail. Here we assess the ubiquitin-binding properties of Rpn10 and Rhp23 in Chapters 2 and 4, respectively.

Little is known about the structural organisation of the 19S RP and, apart from the ATPase subunits, the locations of other subunits have not been unambiguously determined. Rpn10 is a mediator between the 19S RP base and lid subcomplexes. Its VWA domain is thought to be responsible for these interactions and its crystal structure is described in Chapter 2. The interaction of Rpn10 with the lid-subunit Rpn12 and the crystal structure of the PCI domain of Rpn12 will be the subject of Chapter 3.

Relatively weak and transient interactions seem to be an inherent feature of ubiquitin receptors. Therefore, a broad range of structural, biophysical and biochemical methods have to be used to probe this system. X-ray crystallography was employed to gain structural insights into two proteasomal subunits, and biophysical methods such as SPR, ITC, AUC and fluorescence anisotropy were used to characterise protein-protein interactions. To produce sufficient quantities of reagents, the production of labelled and unlabelled Rpn10, Rpn12, Rhp23 and ubiquitin chains had to be established and optimised.

2 Structure of Rpn10 and its Interactions with Ubiquitin Chains

2.1 Introduction

The human ortholog of Rpn10 (S5a) was the first ubiquitin receptor to be identified (Deveraux et al. 1994). The gene is essential in mouse (Hamazaki et al. 2007), but has been found to be non-essential (van Nocker et al. 1996; C. R. M. Wilkinson et al. 2000) in yeast. However, a triple deletion of Rpn10, Rhp23 and Dph1 is synthetically lethal in *S. pombe* (Hartmann-Petersen et al. 2004).

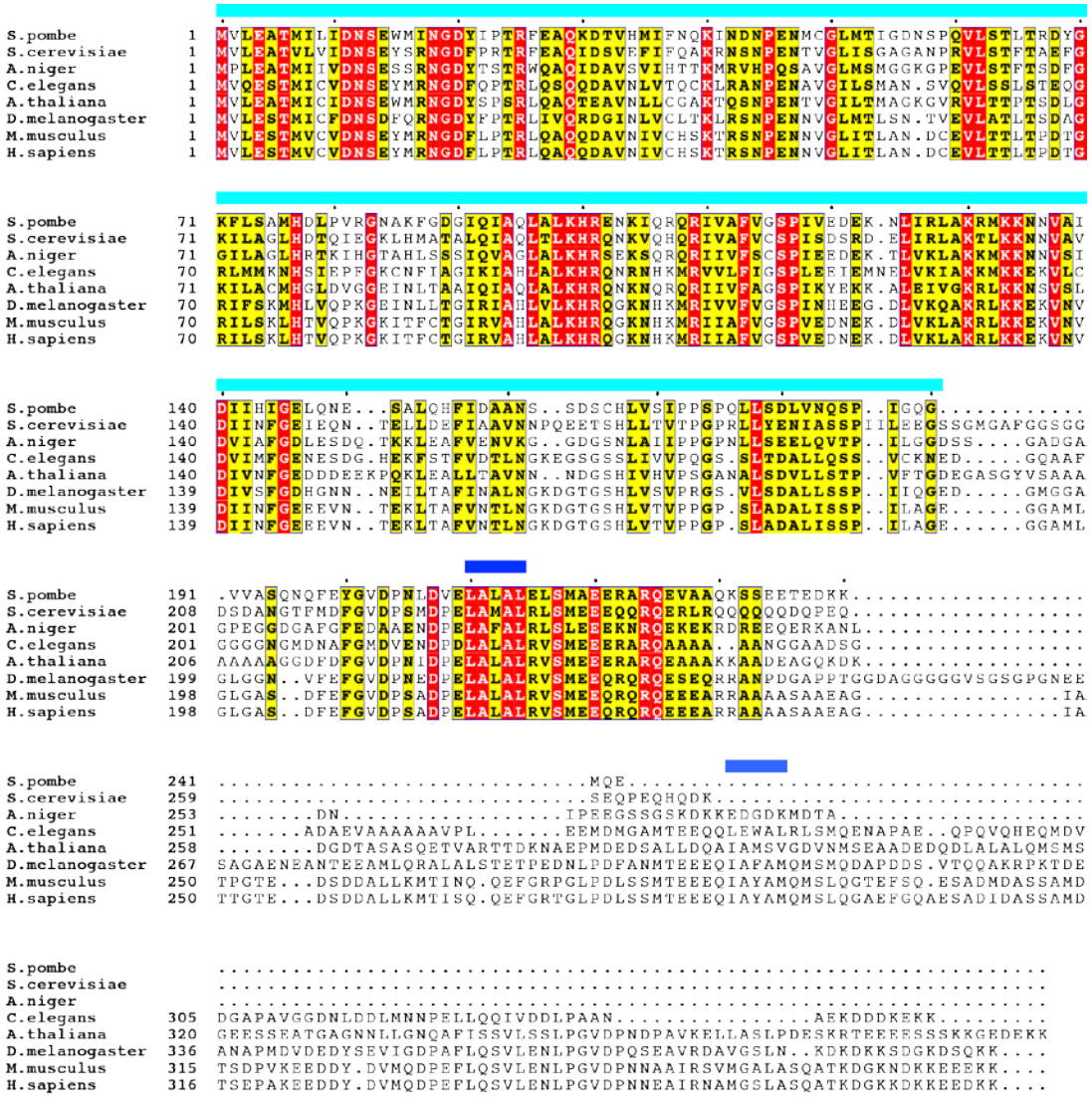


Figure 2.1. Sequence alignment of Rpn10 orthologs. The sequences of Rpn10 orthologs were aligned with ClustalW2 alignment software (Larkin et al. 2007). Similarity scores were calculated with the Risler matrix (Risler et al. 1988) and coloured accordingly. The N-terminal VWA domain and the helix containing the UIM (identified by the cyan and dark blue bars respectively above the sequences) are very well conserved amongst Rpn10 orthologs. The second UIM found in higher eukaryotes is indicated by a light blue bar.

Rpn10 consists of a conserved 193 residue N-terminal domain and a single C-terminal UIM (Figure 2.1). Higher eukaryotes contain a second C-terminal UIM. The N-terminal domain is weakly homologous to the A-domain of the von-Willebrand-factor (VWA), a plasma protein that mediates cellular adhesion (Whittaker et al. 2002). The structures of several VWA domains have been solved (reviewed in Springer 2006). They all consist of a central β-sheet with one antiparallel edge strand, sandwiched between three α-helices on each side (Figure 2.2). Despite their similarity they do vary in terms of binding sites and binding mechanisms. The anthrax toxin binds to the loops following the β-sheet of lamin receptors (Santelli et al. 2004), collagen interacts with the bottom of the VWF-A3 domain (Bienkowska et al. 1997) and platelet glycoprotein Ib binds along the side of β-strand 3 with the von Willebrand factor A1 domain (Huizinga et al. 2002). Some VWA domains have a metal ion coordinated by loop residues that contribute to a motif known as the metal-ion-dependent adhesion site (MIDAS). This motif consists of the DXSXS sequence in the first loop, a threonine in the second and an arginine in the third loop. The coordinated metal ion is always implicated in binding. In integrins VWA domains undergo a large conformational change upon binding as the coordination of the Mg^{2+} is altered by the threonine moving in and the aspartate moving out. This local movement is transduced to the following C-terminal domain (Springer 2006).

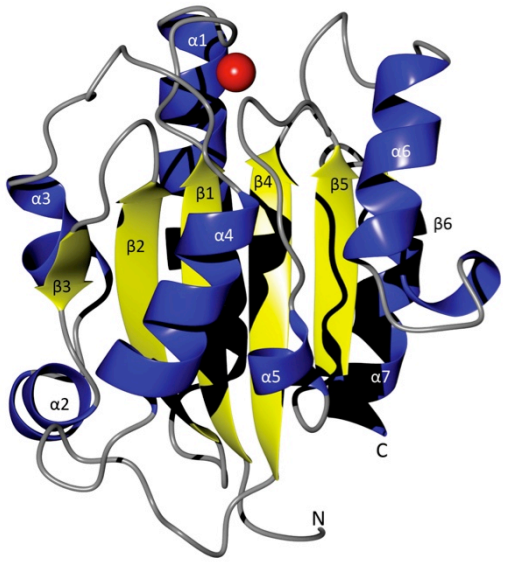


Figure 2.2. The canonical VWA-domain fold. The structure of the VWA/I domain of the integrin αL subunit (PDBID 1T0P). β-strands are highlighted in yellow, α-helices in blue and the coordinated Mg^{2+} ion is coloured red.

Several phenotypes and functions are associated with the VWA domain of Rpn10 that are well conserved across its orthologs (Figure 2.1). Proteasomes purified from *Δrpn10* yeast strains led to the discovery of the base and lid sub-complexes of the 19S RP and demonstrated the role of the Rpn10 VWA domain as a mediator between base and lid (Glickman et al. 1998). Later *in vitro* assays showed that this domain interacts with Rpn12 (C. R. M. Wilkinson et al. 2000) and the base subunit Rpn1 (Seeger et al. 2003) establishing this role. Indeed, deletion of residues 1-84 abolishes the ability of Rpn10 to bind to PC-repeat modules from Rpn1, Rpn2 or the APC/C component Cut4 (Seeger et al. 2003). However, Rpn10 truncated at the N-terminus to residue 62 was still able to bind to the base of the 19S particle but not to the lid (Glickman et al. 1998). Mutation of D11 to arginine also resulted in the dissociation of Rpn3 and other lid subunits, possibly (as the authors suggested) by disrupting the fold (Fu et al. 2001).

Sequence alignments revealed two highly conserved motifs towards the C-terminus of Rpn10 (Haracska et al. 1997). Lower eukaryotes only possess the more N-terminal one of the two. Truncated recombinant versions of the *D. melanogaster* ortholog were subsequently tested for binding radiolabelled ubiquitin conjugates. These experiments confirmed the hypothesis that ubiquitin binding is accomplished by one or two conserved hydrophobic stretches with the consensus sequence e-e-x-x-ϕ-x-x-A-ϕ-x-e/ϕ-S-z-x-e, where e is an acidic, x a helix-favouring, ϕ a bulky hydrophobic and z a bulky hydrophobic or polar residue with large aliphatic content (Swanson et al. 2003). This motif is now called the Ubiquitin Interacting Motif (UIM). Additional support for this model came from a mutational study in *S. pombe* in which the central LALAL sequence was converted to NNNNN resulting in complete loss of ubiquitin binding (C. R. M. Wilkinson et al. 2000).

The structure of the C-terminal tail of the human ortholog S5a in complex with ubiquitin has been determined by NMR and consists of three non-interacting helices connected by short unstructured linker-sequences (Wang et al. 2005, Figure 2.3). Within the UIMs, the hydrophobic LALAL/IAYAM stretch forms the core of the UIM and it is the residues at positions 1 (L216/I287) and 4 (A219/A290) that form the

strongest and most extensive contacts with ubiquitin. In addition to these hydrophobic interactions, S223 in UIM-1 and S294 in UIM-2 form a hydrogen bond with the backbone of G47 of ubiquitin. The large hydrophobic residues of this stretch also contact this residue in addition to I44. UIM-2 of S5a forms a more extensive interface with ubiquitin (Figure 2.3A, B), and this is reflected in a fivefold higher affinity (350 μM vs. 75 μM). Several factors contribute to this difference. Y289 can form more extensive van-der-Waals interactions with L8 of ubiquitin than L218, the comparable residue in UIM-1. The N-terminal region of UIM-2 comprised of P276-T282 also caps the ubiquitin hydrophobic surface formed by L8, V70 and L71. In more recent work, K. Walters et al. extended this study to the binding of K48-Ub_2 and demonstrated that both UIMs are involved and form a 1:1-complex in which UIM-2 is folded back to position itself next to UIM-1 to bind an open K48-Ub_2 molecule (Figure 2.3C, D, Zhang, Wang, et al. 2009). Interestingly, UIM-2 has a threefold preference for binding the proximal rather than the distal ubiquitin moiety. In addition to the interactions observed upon monoubiquitin binding, contacts are made by the conserved ^{206}FGVDPS^{211} sequence located N-terminal to the UIM. Overall, these factors result in an increased affinity of Rpn10 for K48-linked Ub_2 of 8.9 μM.

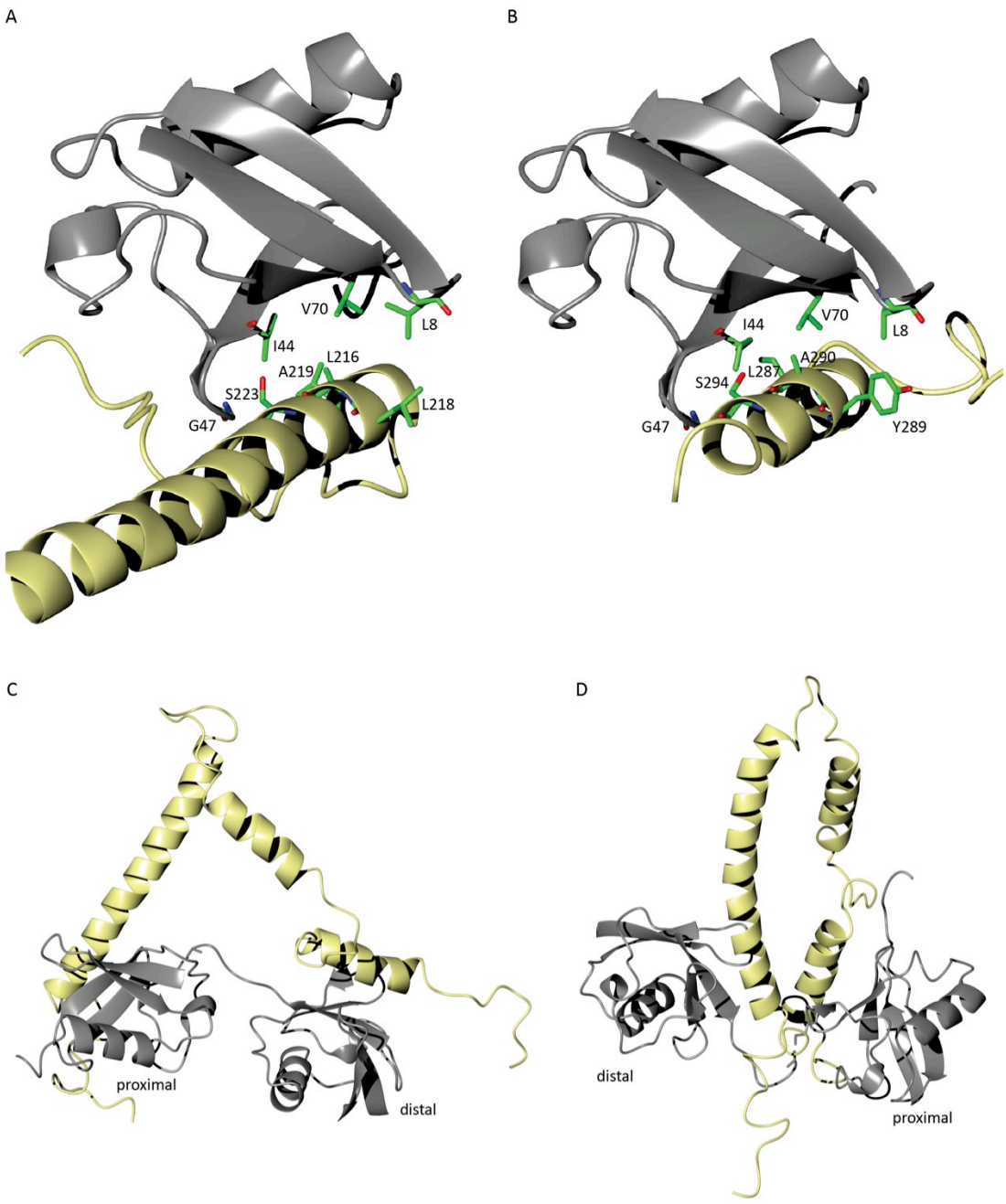


Figure 2.3. Interaction of S5a with ubiquitin and K48-Ub_2. Both UIM-1 and UIM-2 of S5a bind to monoubiquitin via the hydrophobic patch (L8, I44, V70) and a hydrogen bond between G47 of ubiquitin and S223/S294 of S5a. Ubiquitin is drawn in grey and the S5a UIM helices in yellow. **(A)** UIM-1 contributes mainly L216, L218 and A219 (1YX5). **(B)** The interaction with UIM-2 is stronger because Y289 can form stronger van-der-Waals interactions with L8 of ubiquitin than L218 (PDBID 1YX6). Both UIMs are involved in the interaction with K48-Ub_2 to form a 1:1 complex. **(C)** In a minor species, the proximal ubiquitin moiety interacts with UIM-1 (PDBID 2KDF). **(D)** The interaction between UIM-1 and the distal ubiquitin moiety is more favourable because multiple contacts between the proximal moiety and UBA-2 become more compact (PDBID 2KDE).

Interestingly, there seems to exist a large pool of Rpn10 that is not associated with the proteasome (Deveraux et al. 1995; Haracska et al. 1995; C. R. M. Wilkinson et al. 2000). Apart from ubiquitin, UIM-containing proteins can also bind to the UBL domains of hHR23 and Dsk2 (Matiuhin et al. 2008; Y. Kang et al. 2007). This has been proposed to have functional consequences. For example, it has been suggested that by masking the UBL domain of UBL-UBA domain ubiquitin receptors Rpn10 can limit the flow of ubiquitylated proteins to the proteasome (Matiuhin et al. 2008). A more recent and extensive model proposes that these interactions can be exploited *in vivo* to make Rpn10 effectively a ubiquitin chain-length sensor (D. Zhang et al. 2009). This is possible because Rpn10 binds the UBL of Dsk2 tighter than ubiquitin while the Dsk2 UBA domain has a higher affinity for ubiquitin than for its own UBL domain. Taken together these interactions would result in a ternary complex in which the UBA domain is bound to ubiquitin while the Rpn10 UIM masks the Dsk2 UBL domain making it impossible for Dsk2 to bind to the proteasome (Figure 2.4A). Only when the ubiquitin chain consists of at least four ubiquitin moieties can the complex remodel in a way that the ubiquitin chain is shared between Rpn10 and Dsk2 exposing the UBL domain (Figure 2.4B).

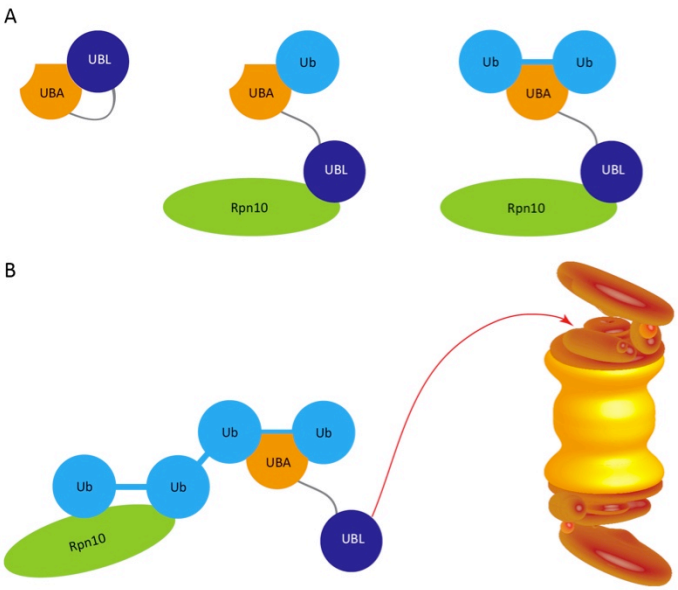


Figure 2.4. Rpn10 can act as a chain-length sensor. Because Rpn10 has a higher affinity for the Dsk2 UBL domain than for ubiquitin and the Dsk2 UBA domain binds ubiquitin more tightly than its own UBL domain, the UBL domain is only exposed (and thus able to bind to the proteasome) if a ubiquitin chain is long enough to displace Rpn10 from the UBL domain. Model redrawn from (Zhang, Chen, et al. 2009).

Although the selectivity of S5a for K48-Ub_2 over monoubiquitin has been demonstrated and the structural mechanism revealed, the discrimination over K63-linked polyubiquitin has only been addressed qualitatively (Haririnia et al. 2007; Q. Wang et al. 2005). The question as to whether a single UIM as found in the yeast orthologs of Rpn10 has inherent selectivity for chains of particular length and linkage type remains to be answered.

2.2 Expression and Purification of the Rpn10 VWA Domain

Untagged $Rpn10_{1-193}$ was expressed recombinantly from *E. coli* and purified by a strategy combining hydroxyapatite, ion exchange and size exclusion chromatography (SEC). A detailed description of the purification protocol can be found in Section 6.3.3. The first hydroxyapatite chromatography step was very efficient and resulted in ~50% pure protein (Figure 2.5A). Subsequent size exclusion chromatography (Sephadex 75) further purified the protein to ~80% (Figure 2.5B) and exchanged the buffer to HBS, suitable for anion-exchange chromatography. This third chromatography step resulted in almost pure protein and – even more importantly – removed all DNA fragments bound to $Rpn10_{1-193}$ that eluted between 40 and 60 mS/cm (Figure 2.5C). Finally, remaining minor impurities were removed with a Sephadex 30 column (Figure 2.5D). The protein was concentrated to 18 mg/ml and buffer-exchanged into 10 mM HEPES, 12 mM NaCl, pH 7.5 resulting in circa 2.5 mg of pure protein from a 1-l culture.

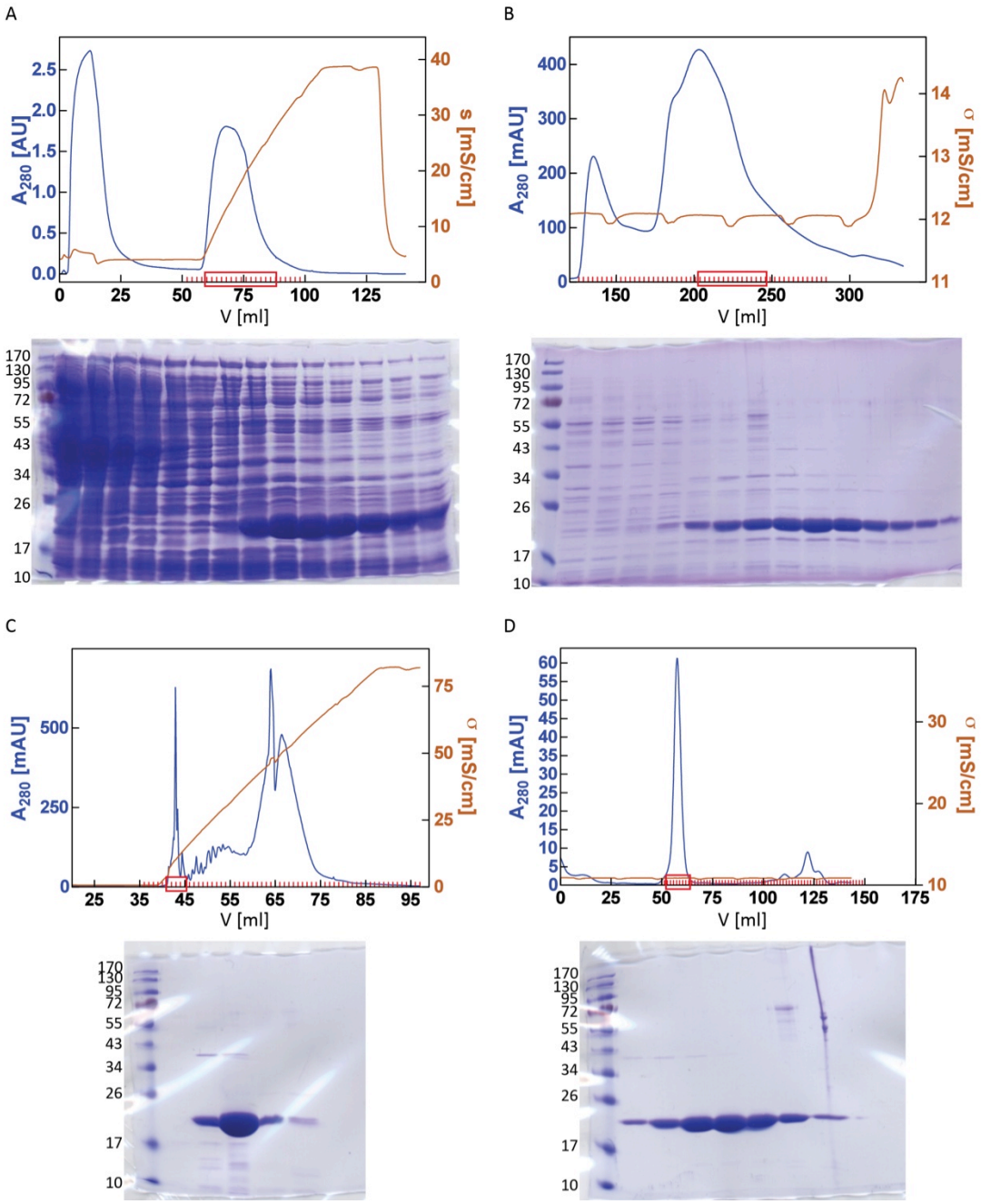


Figure 2.5. Purification of the untagged VWA domain. The upper panel depicts the chromatogram. Fractions analysed by SDS-PAGE (lower panel) are identified with a red square. The untagged Rpn10 VWA domain was purified by the use of hydroxyapatite **(A)**, Sephadex 75 **(B)**, MonoQ **(C)** and Sephadex 30 **(D)** columns.

2.3 Crystallisation and Structure Determination of the Rpn10 VWA Domain

An extensive screen of the Rpn10 VWA domain purified following the above strategy in 96-well-plate format against a mixture of commercially available sparse matrix and systematic grid screens yielded several crystals in different conditions (Figure 2.6). Crystals were all cryo-protected with 25% glycerol and tested for X-ray diffraction. Remarkably, the thin sheets grown in Tris/HCl pH 8.5, 1.5 M $(NH_4)_2SO_4$, 25 mM $MgSO_4$ (Figure 2.6C) were by far the best diffracting crystals and a dataset was collected to 1.3 Å at the ESRF beamline ID14.2 (Table 2.1). Previous attempts to obtain well diffracting crystals failed despite the use of numerous constructs, conditions and crystallisation techniques (J.-F. Trempe, unpublished data). We speculate that this might have been caused by the VWA domain non-specifically but quite tightly binding DNA fragments, probably via a large basic patch formed by an accumulation of lysine residues on the surface (Figure 2.8). In the previous purification procedures affinity-tagged constructs had been used that did not necessitate the use of ion-exchange chromatography, the only chromatography technique that removed the DNA from the VWA domain.

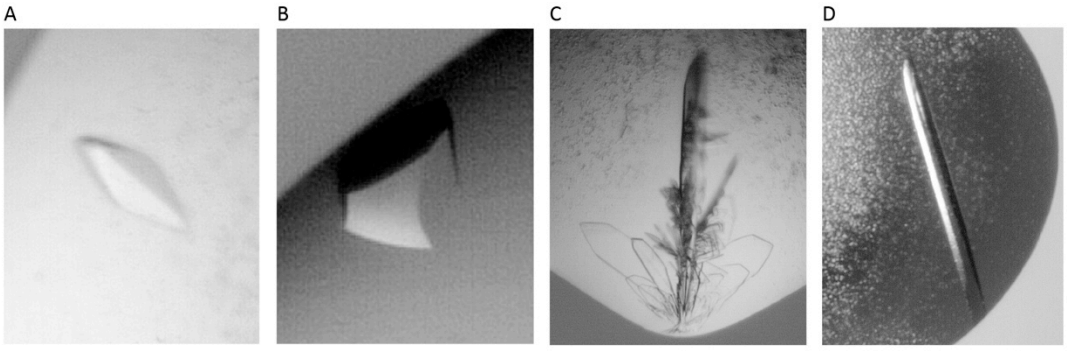


Figure 2.6. VWA crystals. Several conditions yielded VWA domain crystals. **(A)** 0.1 M HEPES pH 7.5, 1.3 M $(NH_4)_2SO_4$. **(B)** 0.1 M HEPES pH 7.5, 2.0 M $(NH_4)_2SO_4$. **(C)** Tris/HCl pH 8.5, 1.5 M $(NH_4)_2SO_4$, 25 mM $MgSO_4$. **(D)** 0.1 M citric acid pH 5.0, 20% PEG-6000**.** The crystals obtained from (C) diffracted best and a single crystal was used to collect native and sulphur-SAD datasets.

The VWA domain structure was successfully solved by Sulphur Single-Wavelength Anomalous Dispersion (S-SAD). Native and S-SAD anomalous datasets, respectively, were collected from the same crystal at the ESRF beamline ID14-2 and in-house.

Datasets were integrated, using MOSFLM (Leslie 1992) and scaled using SCALA (Collaborative Computational Project 1994), and then converted to SHELX format (hkl) with MTZ2VARIOUS (Collaborative Computational Project 1994). Both datasets were prepared for further analysis with SHELXC (Sheldrick 2008). Ten anomalous sites were identified by SHELXD and phasing calculations and solvent flattening were performed in SHELXE. At this point an electron density map with single atoms resolved allowed model building in COOT (Emsley et al. 2004). The model was subsequently improved by alternating building – refinement cycles in COOT and PHENIX.REFINE (Adams et al. 2010), respectively. The final model comprises residues 2-147 and 151-193. 264 water molecules and three sulphate ions, supported by peaks in the difference electron density map could be modelled into the density. It possesses a crystallographic R-factor of 0.1228 and an R_{free} of 0.1646. Data collection and refinement statistics are given in Table 2.1. A more detailed description of the structure solution can be found in Section 6.4.1.

Table 2.1. Rpn10 VWA crystallographic data and refinement statistics.

	SAD	Native
Data Collection		
Space Group	C2	
Unit Cell dimensions		
a, b, c [Å]	80.08, 35.28, 60.16	79.99, 35.23, 60.14
α, β, γ [°]	90.00, 111.23, 90.00	90.00, 111.2, 90.00
Wavelength [Å]	1.54	0.933
Resolution [Å]	30.0-2.1 (2.2-2.1)	30.0-1.3 (1.4-1.3)
Completeness [%]	100 (100)	99.4 (98.7)
Multiplicity	43.14 (38.82)	7.00 (6.80)
I/σ	59.07 (33.02)	19.37 (7.34)
R_{sym}	0.009 (0.019)	0.036 (0.134)
Anom. Multiplicity	20.21 (11.67)	1.70 (1.70)
Anom. I/σ (Inf.-2.2 Å)	1.45	
Anom. CC (Inf.- 2.2 Å)	0.44	
Refinement		
Resolution Range [Å]	30.0-1.3	
R_{Work}/R_{free} [%]	12.28/16.46	
Average B factor [$Å^2$]	15.73	

2.4 The Rpn10 VWA Crystal Structure

The Rpn10 VWA structure consists of a central six-stranded β-sheet of five parallel strands and one anti-parallel edge-strand (β3) sandwiched between two triplets of α-helices (Figure 2.7A). As expected, the fold and topology is very similar to that of other VWA and I-domains despite their low sequence identities. Between β1 and α1 the Rpn10 VWA domain has a five-amino-acid insertion that is almost invariant amongst Rpn10 orthologs and is absent in most other VWA domains (Figure 2.1). Notably the insertion disrupts the metal-ion-dependent adhesion site (MIDAS) that binds Mg^{2+} ions in a number of VWA domains (Springer 2006). As described above, this part of the VWA fold is essential for maintaining the structural integrity of the 19S proteasome and for conferring resistance to amino acid analogues (Fu et al. 2001). The structure reveals that the substitution of D11 at the end of β1 to an arginine that results in the dissociation of Rpn3 and other lid subunits would not only prevent the formation of a hydrogen bond between the side chain of D11 and the backbone nitrogen of G115 (located at the end of β4), but might also disrupt the hydrogen bond between S13 and S116 (Figure 2.7B). Together, these changes could result in loss of the loop orientations necessary for interactions with the lid.

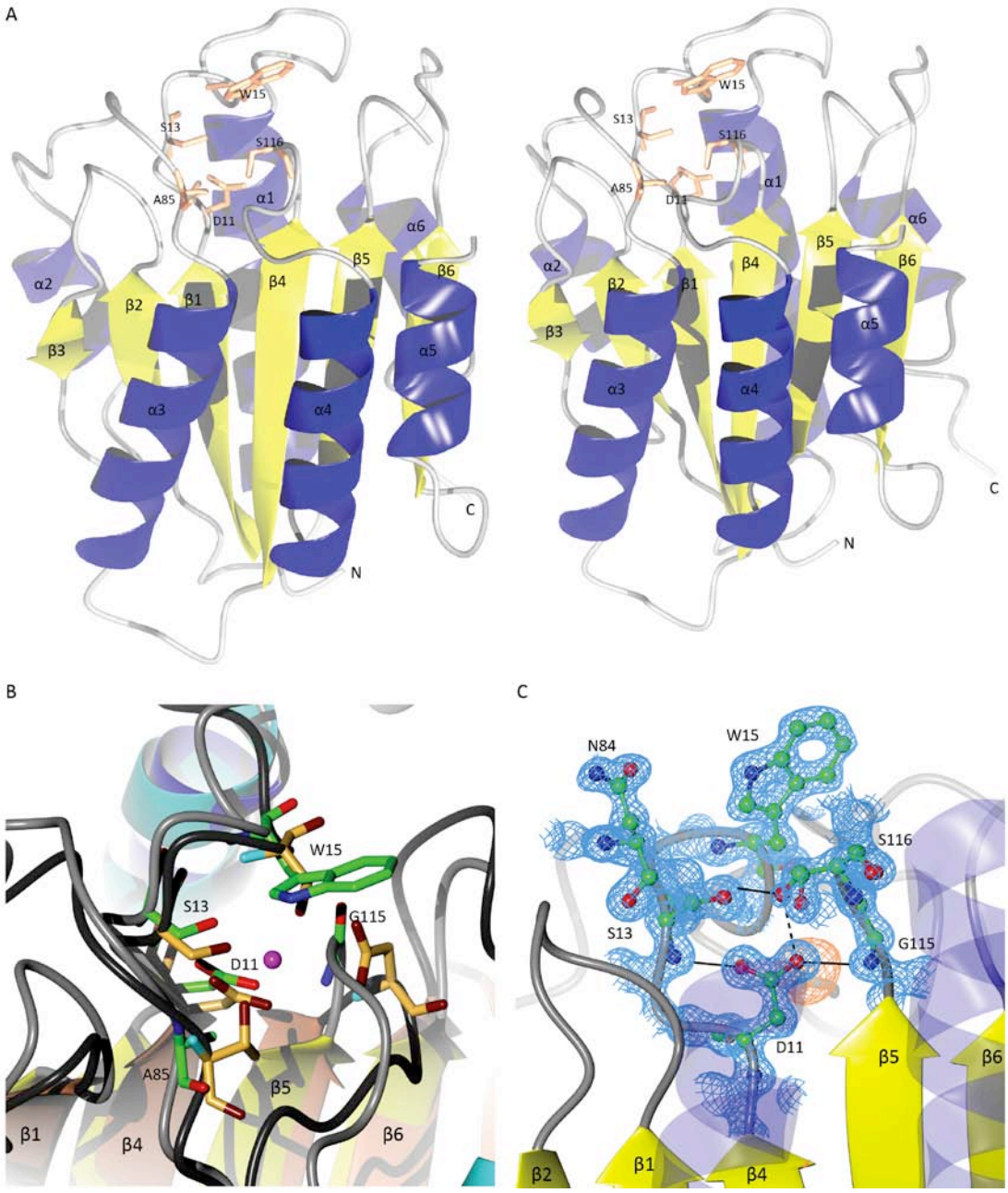


Figure 2.7. The Rpn10 VWA structure. **(A)** Cross-view stereo pair of the structure of the Rpn10 VWA domain, drawn with secondary structure elements coloured yellow (β-sheets β1-β6) and blue (α-helices α1-α6). Amino acids that constitute the pseudo-MIDAS at the end of β1 are drawn as cylinders, labelled and coloured in coral. **(B)** Secondary structure alignment of the region around the MIDAS of integrin αL with the same region of Rpn10. Rpn10 is coloured identical to **(C)** and the Mg^{2+} ion coordinated by integrin αL is highlighted in cyan. The MIDAS is characterised by the sequence DXSXS together with threonine and aspartate residues in the second (between α2 and α3) and third (between β4 and α4) loops, respectively. In the Rpn10 VWA domain, these residues are replaced by the sequence DNSEW (of which the side chains of D11, S13 and W15 are drawn) and amino acids A85 and G115. Despite low sequence conservation the region is structurally highly conserved. **(C)** displays the disrupted MIDAS of Rpn10 from a different point of view. The hydrogen bonds formed by D11 show the importance of interactions in this region for the structural integrity of the "top loop" region of the fold. The $2F_o$-F_c electron density map is contoured at 0.81 $e^-/Å^3$ and drawn in blue, the anomalous difference map at 10.01 $e^-/Å^3$ and coloured orange.

Analysis of the surface electrostatics revealed a large basic patch covering a substantial part of one side of the VWA domain (Figure 2.8A). The orientation of the VWA domain in the top panel is similar to that of the secondary structure overview in Figure 2.7. Conserved surface residues were identified by colouring the surface according to conservation scores obtained from CONSURF (Landau et al. 2005) from white to green and residues with scores >8 being bright green (Figure 2.8B). This analysis revealed a large patch on the top end of the domain that is formed by extended loops involved in the pseudo-MIDAS (Figure 2.7A). We used the PISA software (Evgeny Krissinel et al. 2007) to identify interfaces and surfaces of the crystal structure that may be possible protein-protein interaction sites. One interface can be formed by a symmetry operation generating a twofold axis on a surface formed by α1 and α2. This interaction would involve an area of 561.5 $Å^2$. The same interface was involved in crystal contacts in another crystal form of the Rpn10 VWA domain (J-F Trempe, unpublished data). The solvation free energy gain upon formation of the interface is -6.7 kcal/mol reflecting the hydrophobic nature of the interaction. A second interface formed by crystal contacts involves the pseudo-MIDAS interacting with the side formed by the parallel β6 edge-strand. This interaction was also predicted to be of hydrophobic nature with a free energy gain of -6.0 kcal/mol.

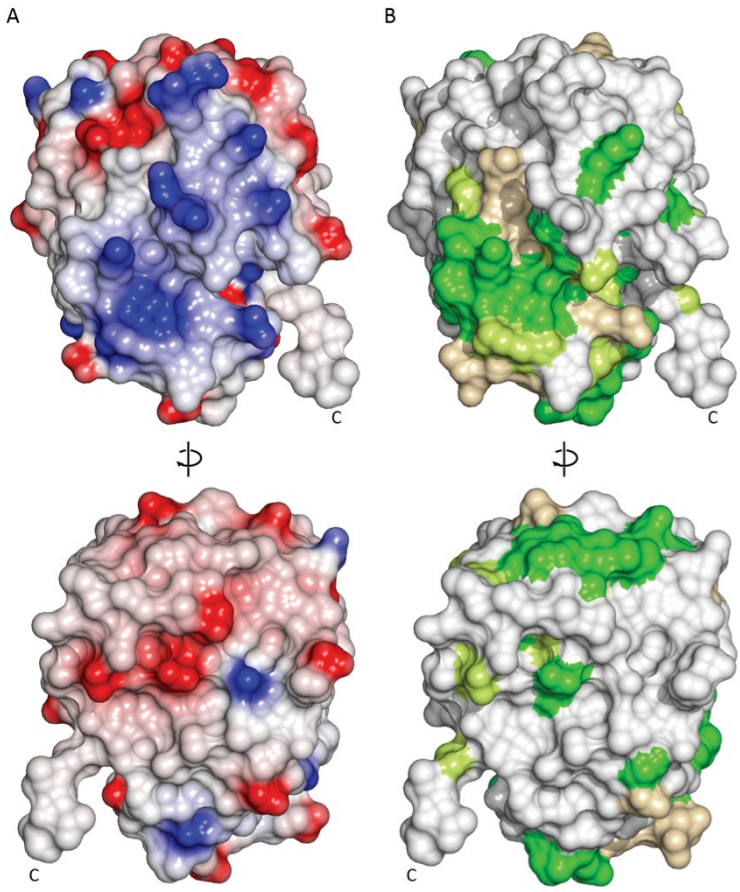


Figure 2.8. Rpn10 surface properties and mutagenesis. (A) Surface electrostatics analysis reveals a striking basic patch on one side of the VWA domain (top panel). **(B)** Conserved surface residues were identified with the software CONSURF (Landau et al. 2005) and calculated conservation scores ranging from 1 (low) to 9 (high) were used to highlight the most conserved residues. Residues scoring <5 are coloured in white, those scoring 5-8 are highlighted with a gradient from white to green and those scoring >8 in bright green.

2.5 The VWA Structure Identifies Potential Sites of Rpn10 Interactions

VWA domains have no conserved protein-protein interaction sites. In VWA domains with a MIDAS, the coordinated metal ion is always implicated in binding. Other binding sites have also been identified along the antiparallel β3 edge-strand and near the bottom-side of the domain (Springer 2006). Based on the analysis of the solvent-exposed surface of the Rpn10 VWA domain, we designed patches of alanine mutations in order to identify potential protein-protein binding sites (Figure 2.9 and Table 2.2). Although Rpn10 does not contain a MIDAS, the equivalent region is structurally very well conserved (Figure 2.7C) and contains the solvent-exposed W15 residue indicative of a potential binding site. These mutations in the context of the

full-length protein were tested by our collaborator C. Gordon for their ability to rescue a *Δrpn10, Δrhp23* conditional lethal *S. pombe* strain. Previous experiments had shown that Rpn10ΔUIM was able to restore wild-type functionality in this strain whereas the UIM on its own failed to do so. All of the VWA mutations we generated were able to rescue this yeast strain suggesting that our mutations did not sufficiently interrupt the interaction site.

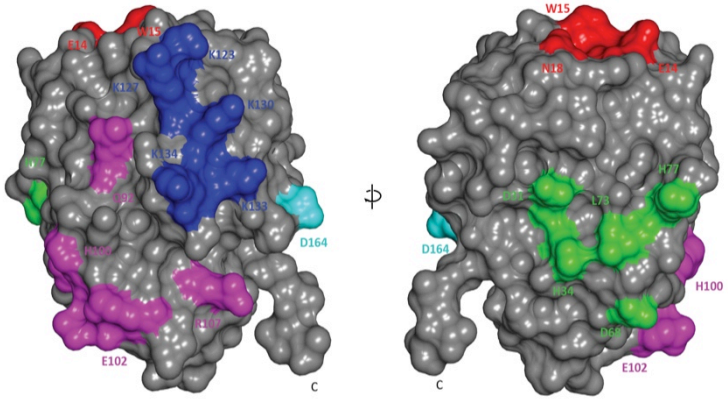


Figure 2.9. Rpn10 mutagenesis. Based on the analysis of the solvent-exposed surface of the Rpn10 VWA domain, five patches were mutated to alanine to identify potential binding sites.

Table 2.2. Rpn10 VWA mutants.

Region	Mutations
Pseudo-MIDAS	E14A, W15A, N18A
PISA-Dimer	D31A, H34A, D68A, H77A
β3-α3	T64A, Q92A, H100A, E102A, K104A, R107A
Basic patch	K123A, K127A, K130A, K133A, K134A
β6-α6	D164A, H167A, D180A

2.6 The Rpn10 UIM Binds Selectively to K48-Ub_2

It has previously been shown that tandem UIMs can discriminate between mono- and polyubiquitin (N. Zhang et al. 2009) or exhibit specificity for K63-Ub_2 (Y. Sato et al. 2009). In both cases avidity effects are exploited. Yeast Rpn10 cannot take advantage of these effects because it contains only a single single-sided UIM. It was unknown whether other mechanisms of discrimination are used by yeast Rpn10 or whether the chain selectivity of this protein was a later evolutionary acquirement. To determine the selectivity of Rpn10 for K48-Ub_2 we characterised the interactions

between Rpn10 and ubiquitin by Isothermal Titration Calorimetry (ITC) and sedimentation velocity analytical ultracentrifugation (AUC) to avoid artificial avidity effects often seen in surface-based methods such as SPR. Initial ITC experiments were used to determine approximate affinities and then the experiments were repeated using concentrations ~10× K_d in the cell and ~100× K_d in the syringe. These experiments required very high protein concentrations (up to 6.9 mM) and only the three ubiquitin species tested (monoubiquitin, K48-Ub_2 and K63-Ub_2) proved soluble at these concentrations. For this reason, only the ubiquitin species were used in the syringe. Heats of dilution were determined by titrating the ubiquitin species into buffer and were subsequently subtracted. Rpn10 bound to both monoubiquitin and K48-Ub_2 with 1:1 stoichiometry and exhibited a ten-fold preference for binding K48-Ub_2 (K_d= 13.9 μM) over monoubiquitin (K_d= 139 μM). The binding of K63-Ub_2 to Rpn10 showed 1:2 stoichiometry and could be modelled by K63-Ub_2 having two equivalent binding sites for SpRpn10 each with an affinity similar to that of monoubiquitin (134 μM). AUC was used to confirm the stoichiometries and affinities determined by ITC by another method. Analysis of the interaction of Rpn10 with K48- and K63-linked Ub_2 by sedimentation velocity AUC unambiguously confirmed that K48-Ub_2 forms a 1:1 complex with Rpn10. AUC analysis of the K63-Ub_2/Rpn10 complex yielded a broader profile that was consistent with the predominant species in solution being a 1:1 complex of relatively weakly binding proteins. It is likely that the 2:1 Rpn10:K63-Ub_2 species was not observed by sedimentation velocity AUC because the lifetime of this complex is too short for detection by this non-equilibrium method.

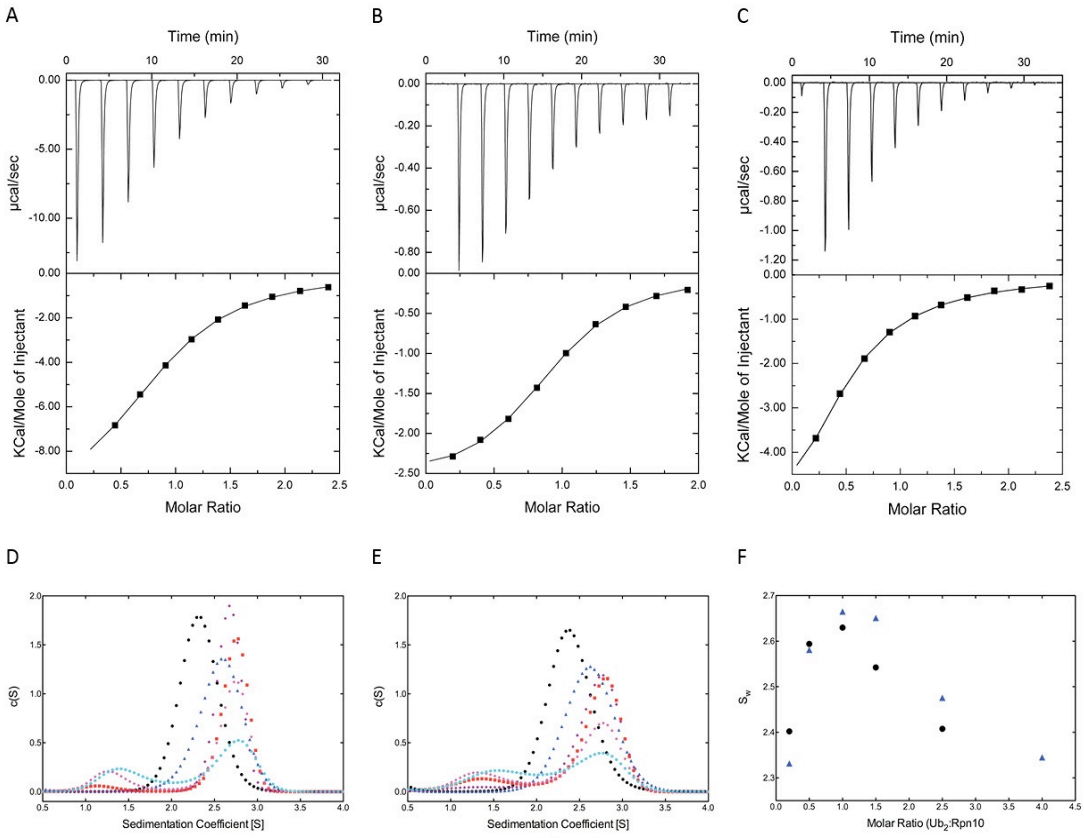
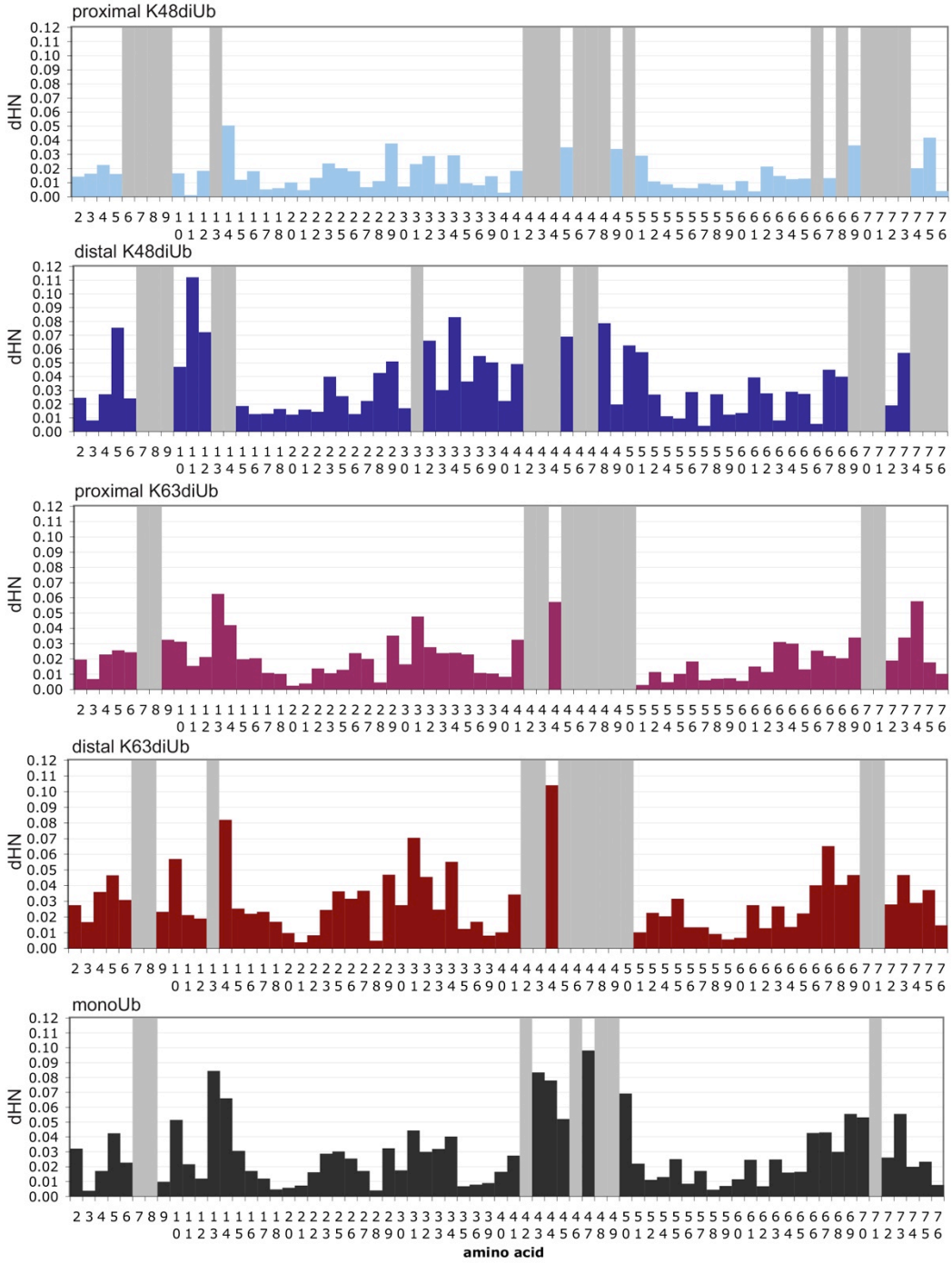


Figure 2.10. Rpn10 interactions with various ubiquitin chains. (A-C) ITC analysis. Raw ITC traces with the respective integrated and normalized isotherms of Rpn10 with monoubiquitin **(A)**, K48-Ub_2 **(B)** and K63-Ub_2 **(C)**. Monoubiquitin and K48-Ub_2 show 1:1 binding characteristics with affinities of 139 µM and 13.9 µM respectively. The interaction of K63-Ub_2 with Rpn10 can be modelled by a two-site model in which each ubiquitin moiety binds Rpn10 with a K_d (134 µM) equivalent to that of monoubiquitin. **(D-F)**. Sedimentation velocity analytical ultracentrifugation. Size distribution profiles for each Ub_2 titration are shown in **(D)** for Rpn10 K48-Ub_2 and **(E)** for Rpn10 and K63-Ub_2. The total protein concentration was kept constant at 300 µM and the Ub_2:Rpn10 ratio was varied from 0.2 (black), 0.5 (blue), 1 (red), 1.5 (purple), 2.5 (magenta) to 4.0 (cyan). **(F)** Plots of the experimentally determined weight-average sedimentation coefficients against the molar ratio indicate in each case formation of a 1:1 complex at the given concentrations. Values for the K48-and K63-linked Ub_2 titrations are plotted as blue triangles and black circles respectively.

To characterise the binding and discrimination mechanism we used $^{1}H/^{15}N$ heteronuclear single quantum coherence (HSQC) NMR titrations to identify residues in the UIM and the different ubiquitin species that mediate the UIM-ubiquitin interaction. The resulting two-dimensional spectra contain a peak for each proton attached to ^{15}N and thus at least one peak for each amino acid (except proline). Because in a folded reasonably sized protein the chemical environment of each of these protons is usually diverse, a well-dispersed spectrum can be expected. Ligand

binding results in a change in the chemical environment of the residues comprising the binding sites that manifests itself in shifts of the corresponding peaks. With $^{1}H/^{15}N$ HSQC assignments of both ubiquitin and Rpn10 spectra available from the Biological Magnetic Resonance Data Bank and previous work by J-F Trempe, respectively, we titrated monoubiquitin, K48-Ub_2 and K63-Ub_2 into ^{15}N-labelled Rpn10 (and vice-versa) and plotted the resulting chemical shift perturbations (CSP) against the corresponding residue number (Figure 2.11). Grey bars indicate peak attenuation. Attenuation of specific peaks (in contrast to broadening of most of the peaks in a spectrum) normally reflects intermediate or slow (on the NMR chemical shift timescale) exchange between free and ligand-bound states due to slow on/off kinetics. All NMR titrations and data analysis were performed by C. Riedinger. Because the C-terminus of Rpn10 is largely unstructured and very flexible, it was possible to only view peaks from this part of the molecule by choosing a higher contour level. All three ubiquitin species were found to bind to almost identical amino acids of Rpn10 centred on the canonical UIM (Figure 2.11B). A notable exception is K48-Ub_2 that seems to weakly interact with the sequence preceding the UIM (193-198). To be able to discriminate between the proximal and distal ubiquitin moieties in the diubiquitin species, we used segmentally ^{15}N-labelled Ub_2 in which either the proximal or the distal moiety was isotope-labelled. Remarkably, only minor differences between both proximal-distal moieties and between the different ubiquitin species could be detected (Figure 2.11A). In all cases, amino acids around the core hydrophobic patch (residues L8, I44 and V70) display major changes in their chemical environment characterised by slow on/off kinetics. Both K63-Ub_2 moieties experience essentially identical CSP and signal attenuation, which are also extremely similar to the ones observed with monoubiquitin. K48-Ub_2 differs slightly. Its distal moiety experiences strong perturbations at the very C-terminus that were not observed in any of the other ubiquitin species used. This might indicate some direct participation of the isopeptide bond. However, these results exclude a discrimination model in which the UIM directly recognises structural features of the isopeptide bond linkage because the NMR data demonstrates that both ubiquitin moieties are involved in interactions with the UIM that are very similar between all tested ubiquitin species.

A


proximal K48diUb
distal K48diUb
proximal K63diUb
distal K63diUb
monoUb
dHN
amino acid

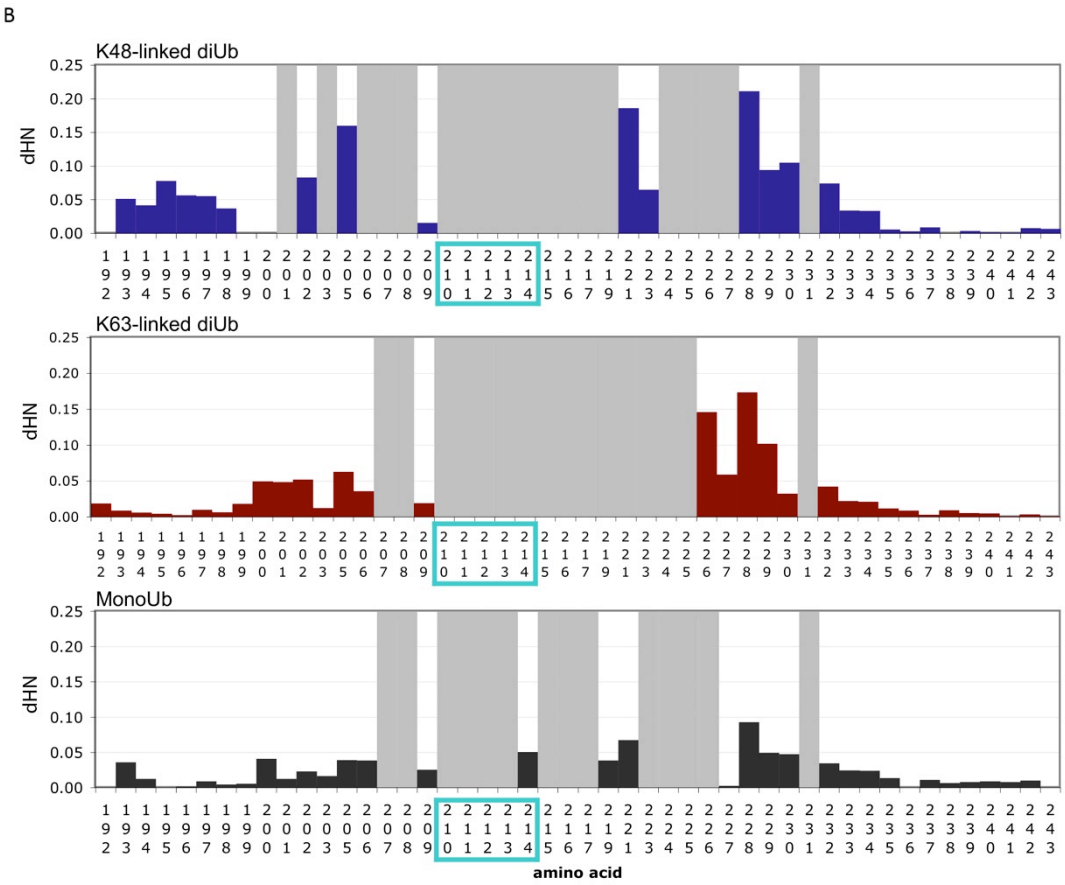


Figure 2.11. Rpn10 and ubiquitin bind via the hydrophobic patch and the UIM. (A) NMR analysis of the interaction between ^{15}N-labeled ubiquitin chains and Rpn10. Chemical shift changes observed in ^{15}N-labeled proximal and distal K48-Ub_2, K63-Ub_2 and monoubiquitin at 1.44 fold molar excess of Rpn10. **(B)** Weighted ^{1}H/^{15}N chemical shift perturbations in backbone amide HSQC peaks of ^{15}N-Rpn10 after addition of monoubiquitin, K63-Ub_2 or K48-Ub_2. Residues for which the peak position in the bound state could not be determined are highlighted in grey. The position of the LALAL motif that forms the core of the Rpn10-ubiquitin binding site is boxed in cyan.

Because all attempts to crystallise the full-length protein were unsuccessful we decided to use a previously determined NMR structure of the C-terminal tail (residues 196-306) of S5a (Q. Wang et al. 2005) to generate a homology model using the software ESYPRED3D (Lambert et al. 2002). The resulting structure is in agreement with CαCβ chemical shifts obtained from NMR assignments performed previously by C. Riedinger and has been used in our models describing the interactions between Rpn10 and monoubiquitin, K63-Ub_2 and K48-Ub_2 (see below).

2.7 Discussion

Analysis of the surface properties of the Rpn10 VWA domain allowed us to identify patches that are possibly involved in protein-protein interactions. In all analyses (electrostatics, conservation, interface prediction) the pseudo-MIDAS was highlighted as a potential binding site. This is also the only site identified that corresponds with known VWA interactions sites. No intriguing properties were found at the antiparallel edge-strand or the bottom side of the domain that are reported protein-protein interaction sites in other VWA domains. Unfortunately, none of our mutants displayed any effect in a rescue experiment in the background of a *Δrpn10*, *Δrhp23 S. pombe* strain. Although the VWA was needed in this rescue experiment it is possible that, instead of being involved in any interaction, it is necessary to stabilise the UIM that might be prone to degradation due to its largely unstructured nature. Another interesting *in vivo* background to test the mutants in would be a temperature-sensitive strain in which Rpn12 is truncated. This strain has previously been shown to be rescued by over-expression of Rpn10 (Wilkinson et al. 2000).

Selectivity for K48-linked ubiquitin chains by a UIM has previously been reported only for two adjacent UIMs exploiting avidity effects or double-sided UIMs (Sims et al. 2009; Zhang et al. 2009; Hirano et al. 2006a). It was questionable whether a single α-helix could discriminate between different polyubiquitin chains. By using ITC to determine the affinities for K48-Ub_2, K63-Ub_2 and monoubiquitin, we avoided the avidity effects from GST-dimerisation that have led to previous false claims of Rpn10 not being selective for K48-linked ubiquitin chains (N. Zhang et al. 2009). Our results clearly demonstrate an almost tenfold higher affinity for K48-Ub_2 over both monoubiquitin and K63-Ub_2 (Figure 2.10). At the same time these experiments revealed that the different complexes also have different stoichiometries: whereas monoubiquitin and K48-Ub_2 form a 1:1 complex, K63-Ub_2 can bind two Rpn10 molecules with an affinity identical to that of monoubiquitin. Our AUC experiments support both of these findings: the weaker affinity of the K63-Ub_2:Rpn10 complex is reflected in broader c(s) distribution peaks and the lack of a 2:1 complex can be explained by the fact that at the concentrations used, this non-equilibrium method

could not resolve this species. Unfortunately, we were not able to determine the affinities of these complexes by AUC. Global analysis of the sedimentation profiles with the software SEDPHAT (Dam et al. 2005) failed because of an apparently very poorly defined isotherm resulting in many possible solutions. This could possibly be overcome by keeping the concentration of one protein constant over the titration thus effectively only sampling one slice of the 3-dimensional isotherm.

Our segmental labelling NMR experiments demonstrated that both K48-linked ubiquitin moieties seem to bind equally well to the UIM. Because the Rpn10 UIM is not a double-sided UIM, this means that there are either two populations (with Rpn10 bound to either the proximal or the distal ubiquitin) or there is a dynamic equilibrium between binding to one moiety, dissociation and rebinding to the other one. Because the affinity is not very high (making dissociation a frequent event) and the two hydrophobic patches are facing each other we propose a model in which sterically favoured rebinding to either the same or the other ubiquitin moiety results in the tenfold discrimination over monoubiquitin and K63-Ub_2 (Figure 2.12A, B). K63-Ub_2, however, has a very elongated conformation and the two hydrophobic patches are on almost opposite sides on the Ub_2 molecule (Figure 1.3D). This would allow two Rpn10 molecules to engage with one hydrophobic patch each at the same time albeit with much lower affinity (Figure 2.12C). This model is supported by our findings that both ubiquitin moieties experience the same chemical shift perturbations and that the ITC data could be fit to a two-site binding model. This tenfold selectivity should be attributed to the conformation of K48-Ub_2 rather than Rpn10 and any ubiquitin receptor with a single α-helix should have this discriminatory effect if it does not possess other features that favour binding to a different linkage type such as the Rap80 tandem UIMs. It is interesting to note that the affinity between K48-Ub_2 and Rpn10 as determined by us is slightly stronger than the K_i determined by competition experiments with 26S proteasomes (Thrower et al. 2000) even though the recently discovered proteasomal ubiquitin receptor Rpn13 binds K48-Ub_2 about 150 times stronger (Husnjak et al. 2008). The already very high affinity of Rpn13 for monoubiquitin might suggest a role distinct from recruiting substrates to the proteasome. Its ability to bind the DUB Uch37 could indicate a role in the proper deubiquitylation of proteasomal substrates or the recycling of ubiquitin moieties from ubiquitin chains.

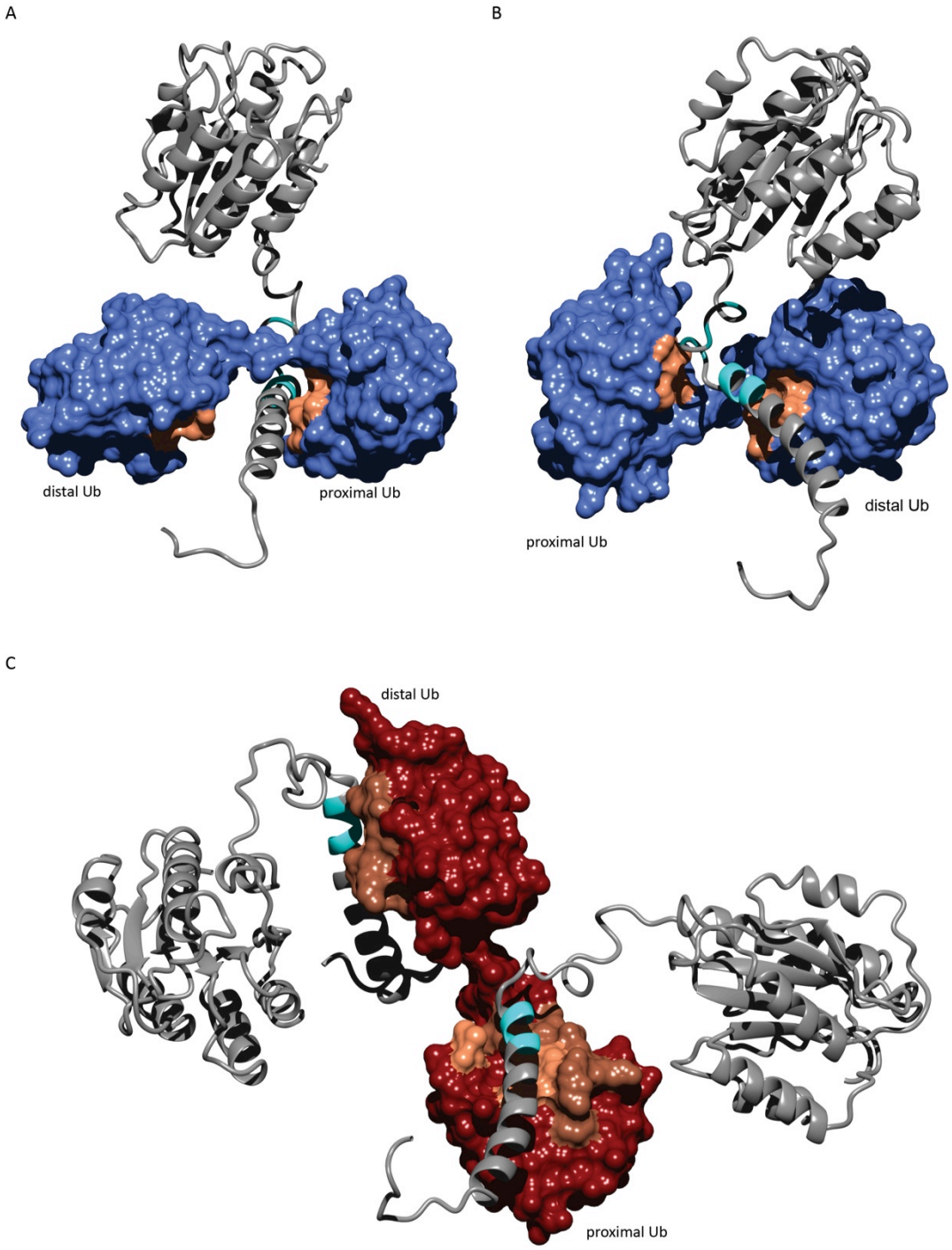


Figure 2.12. Ubiquitin binding to Rpn10. Models for K48- and K63-Ub_2 binding to Rpn10. **(A)** The two hydrophobic patches (orange) of K48-Ub_2 are facing each other and form a groove that can only accommodate one UIM. With the proximal moiety bound, the distal one would be positioned on the opposite side of the UIM-helix. **(B)** Binding of the distal ubiquitin moiety to the UIM places the proximal in close proximity to the Rpn10 linker region. The groove formed between the two moieties "traps" the UIM between two UIM binding sites allowing for efficient rebinding. **(C)** K63-Ub_2 is quite extended and the two hydrophobic UIM binding sites are located on opposite sides of the molecule. This relative disposition permits independent binding of two Rpn10 molecules. Residues within the Rpn10 linker sequence that undergo significant chemical shifts upon K48-Ub_2 binding are coloured dark cyan while the canonical LALAL motif is coloured cyan. Ubiquitin residues L8, I44 H68 and V70 that contribute to the UIM binding site are coloured in orange. Additional K63-Ub_2 residues that experience peak broadening upon Rpn10 addition are coloured brown.

3 Rpn12 Competes With Ubiquitin for Binding to Rpn10

3.1 Introduction

Rpn12 is an essential gene, however, a genetic study in yeast could not identify a specific function for the wild-type protein (Nisogi et al. 1992). The synthetic lethality of the *Δrpn10* and *mts3-1* (a temperature-sensitive Rpn12 mutation) mutations in *S. pombe* was the first genetic indication that there might be an interaction between the two proteins (C. R. M. Wilkinson et al. 2000). Further biochemical characterisation confirmed this finding and demonstrated that they physically associate. This interaction is not dependant on the Rpn10 UIM as Rpn12 was shown to bind equally well to wild–type Rpn10 and to an Rpn10 mutant in which the LALAL sequence had been mutated to NNNNN. Interestingly, however, an intact UIM is required for Rpn10 to rescue the *mts3-1* phenotype at the restrictive temperature by over-expression (C. R. M. Wilkinson et al. 2000).

Rpn12 consists of an N-terminal PCI domain and therefore by analogy to other PCI domains it is likely to serve as a scaffold for the assembly of other proteasomal lid components. Indeed, biochemical studies and mass spectrometry have demonstrated that Rpn12 is part of a sub-complex of the proteasomal lid (Glickman, Rubin, Oliver Coux, et al. 1998; Sharon et al. 2006). It belongs to the atypical Rpn12-like class of PCI proteins that is defined by (i) the lack of any preceding sequence, (ii) a poorly conserved TPR-like N-terminus and (iii) a C-terminal WHD (Pick et al. 2009). An analysis of reported structures has highlighted the extreme versatility of TPR domains: amino acid insertions, loop modifications and changes of the helix length can all be accommodated while still maintaining the TPR fold (D'Andrea et al. 2003). This structural diversity is reflected in the fact that no single binding site is conserved. Protein-protein interactions have been identified at a groove formed by the TPRs of the protein Hop (Scheufler et al. 2000) but also at the loops connecting the helices of TPRs 2 and 3 of $p67^{phox}$ (Lapouge et al. 2000). Interestingly, in this structure the aforementioned groove is still utilised to form an intra-molecular interaction with the unstructured $p67^{phox}$ C-terminus. In *A. thaliana*

CSN7, K144 within the β-sheet of its WHD and E44/H71 located on the second TPR have been shown to contribute to the interaction with CSN8 (Dessau et al. 2008). It has therefore been proposed that one face of the TPR domain and another from the WHD form contact surfaces with CSN8. This structure demonstrates that TPR domains are poorly conserved and very versatile protein-protein interaction units that have adapted to a variety of cellular environments. It is therefore not surprising that the primary sequence among Rpn12 orthologs is much less (31.9% vs. 50.7% pairwise identity) conserved than that of Rpn10 (Figure 3.1).

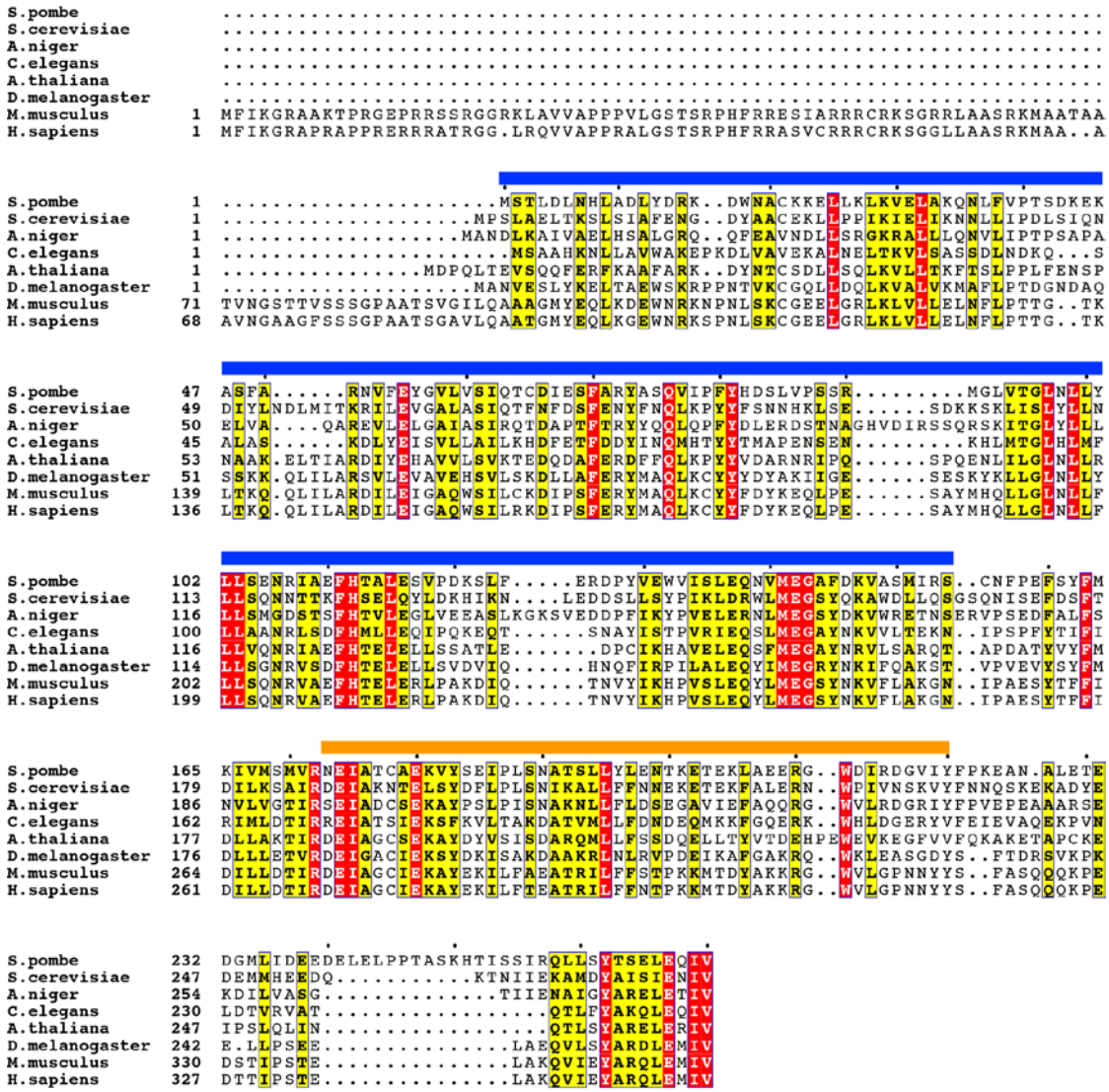


Figure 3.1. Sequence alignment of Rpn12 orthologs. The sequences of Rpn10 orthologs were aligned with ClustalW2 alignment software (Larkin et al. 2007). Similarity scores were calculated with the Risler matrix (Risler et al. 1988) and coloured accordingly. Compared with Rpn10 the sequence is much less conserved. Blue and orange bars indicate the TPR and WH domains, respectively.

3.2 Expression, Purification and Crystallisation of Rpn12

S. pombe Rpn12 was expressed recombinantly as a 3C-protease-cleavable GST-fusion protein in *E. coli* BL21(DE3) from the vector pGEX-6P-1. Because the full-length protein purified as two species suggesting partial degradation, limited proteolysis with subtilisin A was used to produce proteolytically stable constructs. N-terminal sequencing (A. Willis, Department of Biochemistry, Oxford) and mass spectrometry (J. Nettleship, OPPF, Oxford) identified two C-terminal truncations terminating at residues 228 and 250. The two sequences $Rpn12_{1\text{-}228}$ and $Rpn12_{1\text{-}250}$ were subsequently cloned into pGEX-6P-1 and expressed in *E. coli* BL21(DE3) cells. Briefly, both Rpn12 constructs were purified from the soluble fraction by a combination of affinity chromatography exploiting the GST tag and size-exclusion chromatography (Figure 3.2). A more detailed description of the purification protocol can be found in Section 6.3.4. Selected fractions from the size-exclusion column were pooled and concentrated to 12 mg/ml. Rpn12 was then screened using a selection of sparse matrix crystallisation screens (Section 6.4.2) at 20 °C. The cloning, purification and crystallisation of the two Rpn12 constructs was carried out under my supervision by C. Khoudian, a Part II M. Biochem. Student.

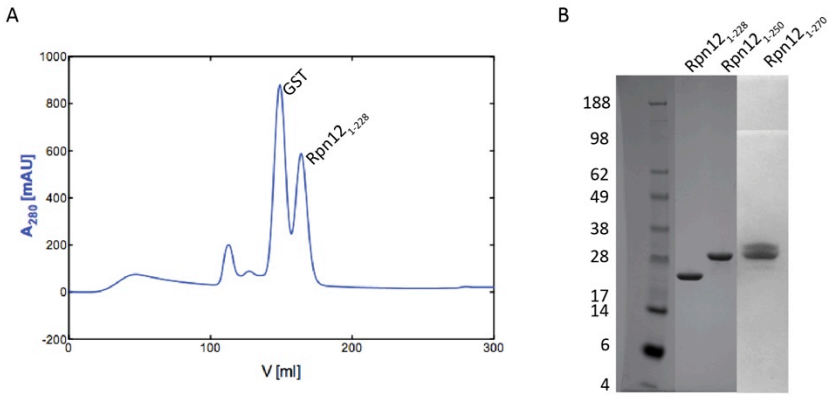


Figure 3.2. Rpn12 purification. (A) The final step of the Rpn12 purification is exemplified by a Sephadex 75 size-exclusion chromatogram of $Rpn12_{1\text{-}228}$. **(B)** SDS PAGE demonstrating the achieved purity of the different Rpn12 constructs. The double-band resulting from C-terminal degradation during expression of the full-length protein is clearly visible.

Only the shorter Rpn12 construct containing residues 1-228 produced crystals of good quality in several conditions. These conditions were characterised by being

buffered in the range of pH 6.5 to pH 8.5 and containing added salt and circa 20% high molecular weight polyethylene glycols (Figure 3.3). By varying the pH and precipitant concentration in the customised screens, these general conditions were optimised and produced large crystals. A native dataset of a crystal grown in 0.2 M $NaNO_3$, 0.1 M Bis-Tris propane pH 7.5, 22.5% PEG-3350 was collected at the ESRF beamline ID14.2 to 1.6 Å.

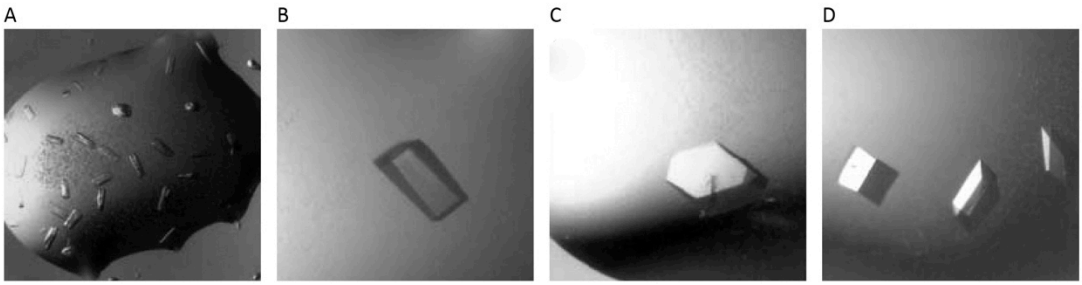


Figure 3.3. Rpn12 crystals obtained from different conditions. (A) 0.2 M $MgCl_2$, 0.1 M Tris pH 8.5, 20% PEG 8000. **(B)** 0.2 M NaI, 0.1 M Bis-Tris propane pH 6.5, 20% PEG 3350. **(C)** 0.2 M $NaNO_3$, 0.1 M Bis-Tris propane pH 6.5, 20% PEG 3350. **(D)** 0.2 M $NaNO_3$, 0.1 M Bis-Tris propane pH 7.5, 22.5% PEG 3350.

3.3 Rpn12 Structure Determination

Initial attempts to phase the data using a sulphur-SAD dataset collected on a rotating anode source were unsuccessful as was heavy atom and iodine soaking. To overcome these difficulties, further crystals were grown from selenomethionine-containing protein and then used to collect a SAD dataset at the ESRF on beamline ID14.4 at a wavelength of 0.979 Å (Table 3.1). The datasets were integrated with MOSFLM (Leslie 1992) and scaled with SCALA (Collaborative Computational Project 1994). For the structure calculation the SHELX suite was used (Sheldrick 2008). Both datasets were prepared by SHELXC. The anomalous signal to noise (reported by SHELXC as d''/σ) dropped below 1.5 at 2.3 Å and therefore the maximal resolution for SHELXD to identify anomalous sites was limited to 2.4 Å. This search resulted in 14 anomalous sites with an occupancy above 0.3, consistent with 14 selenomethionine residues per asymmetric unit. Based on this substructure, SHELXE calculated phases for all reflections and optimised these by applying connectivity restraints and using solvent-flattening with a solvent content of 54%. This resulted in an electron density map of sufficient quality to manually build most of the molecule.

Several rounds of alternating cycles of refinement in PHENIX.REFINE (Paul D Adams et al. 2010) and manual building in COOT (Emsley et al. 2004) resulted in a model with a crystallographic R-factor of 18.80% and an R_{free} of 22.48%. It contains two polypeptide chains (A and B), 364 water, four nitrate, three glycerol and two monothioglycerol molecules.

3.4 The Rpn12 crystal structure

The Rpn12 crystals contained two molecules in the asymmetric unit referred to as chain A and chain B, connected by a disulphide bridge formed by their respective C178 residues. The model for chain A contains residues 1-222 and the N-terminal GPLGS remainder of the restriction/3C-cleavage site, that for chain B residues 5-224. In both proteins, a monothioglycerol molecule was also found to form a disulphide bond with C65. As the Rpn12 construct encodes the first 228 residues, amino acids 225-228 were not visible in the electron density maps and were presumed to be disordered.

Table 3.1. Rpn12 crystallographic data and refinement statistics.

	SAD	Native
Data Collection		
Space Group	$P2_12_12_1$	
Unit Cell dimensions		
a, b, c [Å]	41.45, 91.32, 142.73	41.78, 91.38, 143.34
α, β, γ [°]	90, 90, 90	90, 90, 90
Wavelength [Å]	0.979	0.9685
Resolution [Å]	43.48-1.88 (1.98-1.88)	42.34-1.59 (1.67-1.59)
Completeness [%]	99.4 (99.3)	99.7 (99.7)
Multiplicity	7.0 (7.0)	3.9 (3.9)
I/σ	21.9 (4.1)	17.1 (2.5)
R_{sym}	0.057 (0.389)	0.050 (0.497)
R_{pim}	0.032 (0.170)	0.030 (0.285)
Anom. Multiplicity	3.7 (3.7)	
Anom. CC (Inf.-2.3 Å)	0.75	
Anom. I/σ (Inf.-2.3 Å)	3.17	
Refinement		
Resolution Range [Å]	40.11-1.59	
Number of Reflections	73327	
R_{Work}/R_{free} [%]	18.89 / 23.67	
Average B factor [$Å^2$]	31.97	

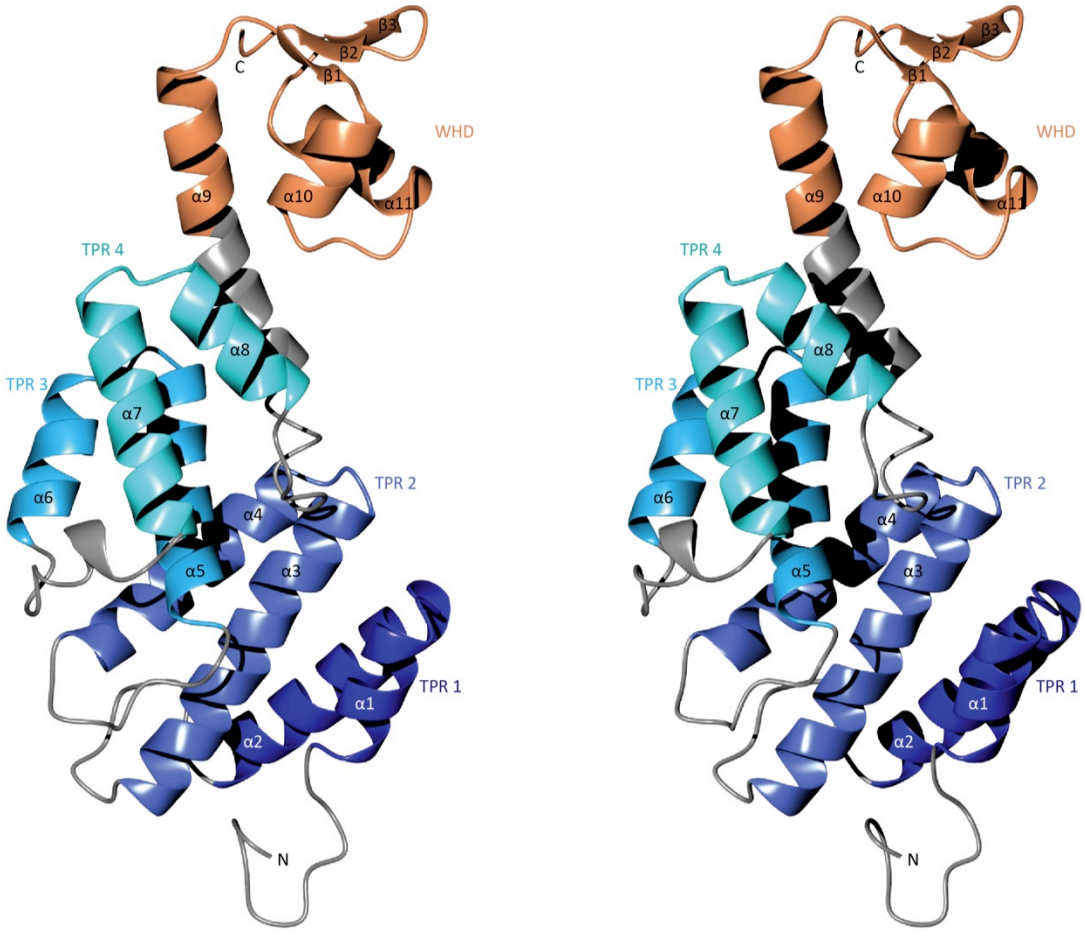


Figure 3.4. Cross-view Stereo Pair of Rpn12. (A) Rpn12 consists of four N-terminal tetratrico peptide repeats (TPRs, blue) connected via a long helix to a winged helix domain (WHD, orange).

As predicted from sequence analysis, the Rpn12 PCI domain consists of a TPR-like N-terminal domain connected to a WHD by a long helix. This helix is often referred to as the “capping” and/or “solubility helix” in the general PCI domain architecture (Figure 3.4). The TPR domain itself contains four repeats forming an elongated superhelix in which each TPR motif is rotated about 20° anticlockwise around a common axis. The long capping helix is packed against the last repeat and connects to the WHD giving the whole molecule an extended shape with a concave and a convex side. This structure has been previously described for the closely related elongation factor eIF3k (Wei et al. 2004). Comparing the two structures reveals the typical PCI domain organization (Figure 3.5). However, the TPRs in Rpn12 are less densely packed resulting in an overall more extended shape.

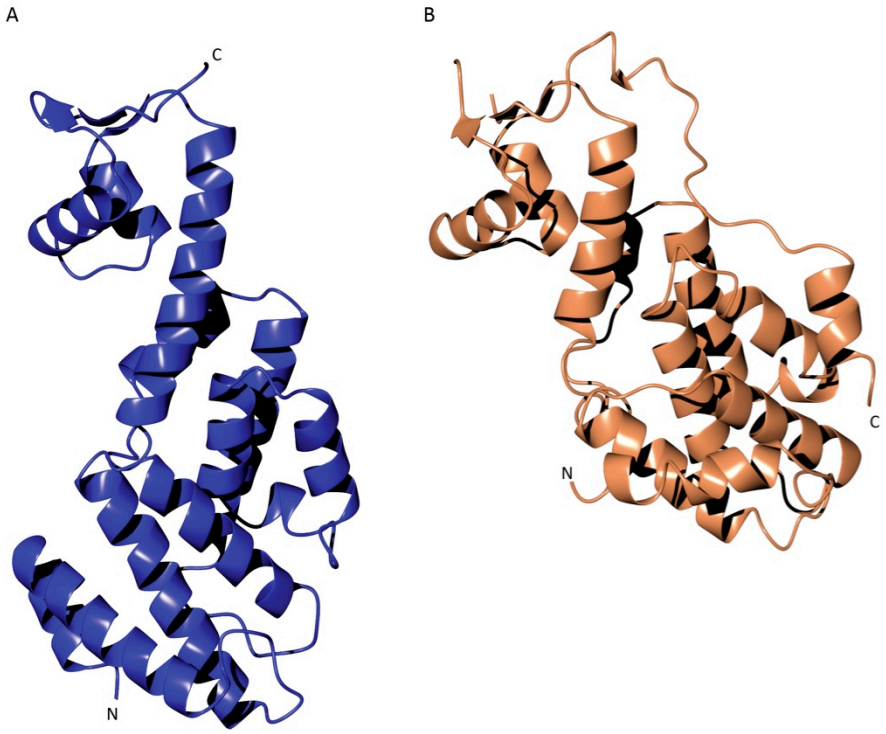


Figure 3.5. The Rpn12 TPR domain differs from that of eIF3k. The Rpn12 **(A)** and eIF3K (PDBID 1RZ4) **(B)** crystal structures were aligned on their WHDs with CCP4MG (Potterton et al. 2004). Although the secondary structures are very similar the TPR domains of these two proteins differ substantially in their tertiary structures. The Rpn12 α-helices are generally longer and TPRs 3 and 4 are parallel to the capping helix. TPRs 1 and 2 are rotated by about 40 °. In contrast, all of the TPRs in eIF3k are rotated with respect to each other and the capping helix. This gives the molecule a less extended shape than Rpn12.

The WHDs of Rpn12 and eIF3 are structurally similar with an rmsd of 1.34 Å over 49 residues despite only 15% sequence identity (Figure 3.6). It seems likely that the Rpn12 C-terminus following the WHD that is missing in this structure might well extend along the concave face as a rather unstructured peptide.

Figure 3.6. The WHDs of Rpn12 and eIF3k are structurally conserved. Secondary structure alignment of the WHDs of eIF3k (blue, PDBID 1RZ4) with that of Rpn12 (orange). Structures were aligned using CCP4MG (Potterton et al. 2004). Despite 15% sequence identity the three-dimensional structure is almost perfectly conserved.

The surface electrostatics are evenly distributed around the molecule with the only noteworthy exception being a single glutamate residue (E185) that is surrounded by a hydrophobic patch surface on the surface formed by α9 and α10 (Figure 3.7A). Using the Rpn12 sequence alignment (Figure 3.1) the most conserved surface residues were identified with CONSURF (Landau et al. 2005) and visualised using CCP4MG (Potterton et al. 2004), (Figure 3.7B). Interestingly, the vast majority of the highly conserved amino acids are located on one side of Rpn12. A large patch is found on the convex face at the end of the TPR motif. This patch involves helices α6 of TPR3 and α7 of TPR4 extending into the WHD via the long capping helix. In addition a much smaller conserved patch is located between TPRs 1 and 2.

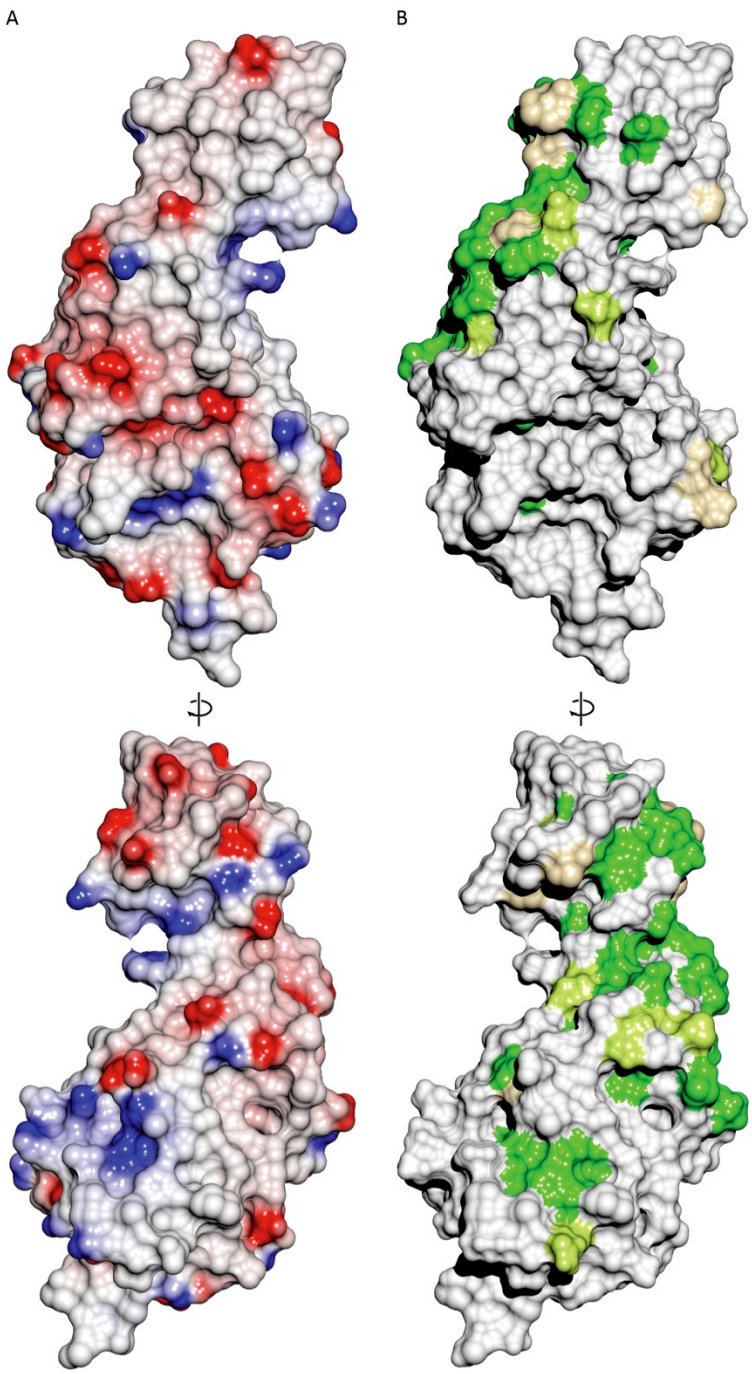


Figure 3.7. Rpn12 surface analysis. (A) Surface electrostatics analysis reveals evenly distributed charges throughout the whole molecule. Notable is a single glutamate residue surrounded by hydrophobic amino acids (top panel). **(B)** Conserved surface residues were identified with the software CONSURF (Landau et al. 2005) and the calculated conservation scores ranging from 1 (low) to 9 (high) were used to highlight the most conserved residues. Residues scoring <5 are coloured in white, those scoring 5-8 are highlighted with a gradient from white to green and those scoring >8 in bright green.

3.5 The interaction between Rpn12 and Rpn10

Genetic and biochemical approaches have previously demonstrated an interaction between Rpn10 and Rpn12 (C. R. M. Wilkinson et al. 2000). To confirm the interaction using purified components we carried out a series of SPR titrations immobilising Rpn12 on the chip and flowing over increasing concentrations of full-length Rpn10 and the Rpn10 VWA domain as the analytes (Figure 3.8). To our surprise using this technique, the interaction between the Rpn10 VWA domain and Rpn12 was much weaker (88 ± 76 μM) than that of the full-length protein (1.9 ± 0.1 μM) (Figure 3.8B). To assess the nature of the interaction the dissociation rates were recorded after short (150 s), medium (300 s) and long (1200 s) analyte injections (Figure 3.8C). Decreasing dissociation rates after long association times are an indicator for a slow conformational change occurring that stabilises the interaction. The decreased dissociation rate that could be detected, however, was too subtle to draw any further conclusions. Because of the very high affinity of full-length Rpn10 for Rpn12 and the unusually slow dissociation rates (Figure 3.8A) we set out to confirm our findings by another, solution-based method. Unfortunately, no signal could be detected using ITC even though titrations were carried out at elevated concentrations of Rpn10 (100 μM) and Rpn12 (1000 μM). ITC measures the enthalpy of a binding event. Rarely, very little or no heat is absorbed or released during such an event in which case this method cannot be used. The interaction between Rpn10 and Rpn12 seems to be one of these cases.

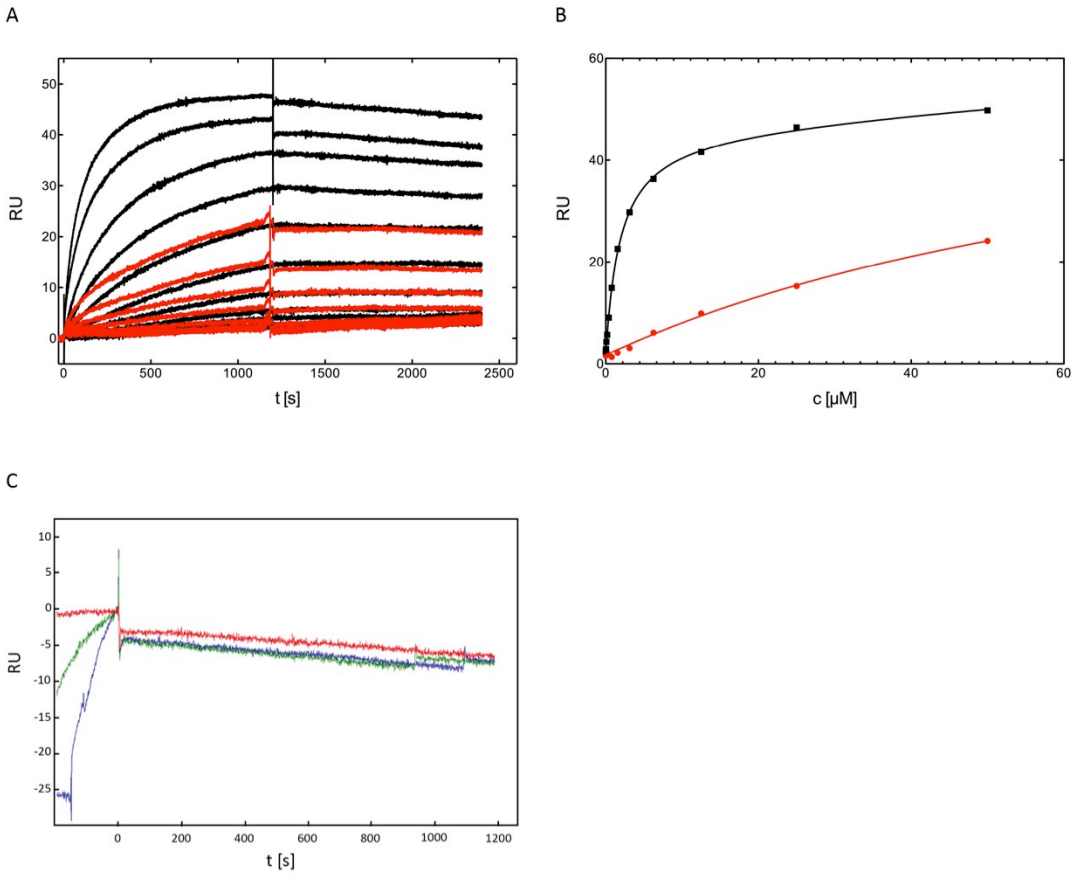


Figure 3.8. The Rpn12-Rpn10 interaction as measured by SPR. Analysis of the Rpn10-Rpn12 interaction by SPR demonstrated that the sequence towards the Rpn10 C-terminus that includes the UIM makes a major contribution to the binding of Rpn10 to Rpn12. **(A)** Raw SPR traces of full-length Rpn10 (black) and the Rpn10 VWA domain (red) flown over immobilised Rpn12. **(B)** The equilibrium binding values were plotted against the respective analyte concentration and the resulting binding curves were fitted with a 1:1 binding model. **(C)** Dissociation traces after 150 s (blue), 300 s (green) and 1200 s (red) association times.

$^{1}H/^{15}N$ HSQC NMR titrations were performed by C. Riedinger to confirm our findings from the SPR experiments and to further characterise the interaction. Addition of both Rpn12 and ubiquitin to ^{15}N-labelled Rpn10 resulted in substantial chemical shift perturbations in the sequence towards the C-terminus of Rpn10 (Figure 3.9A). The same residues that were implicated in the recognition of ubiquitin were also found to be involved in Rpn12 binding, notably D207, V208, L210, L212 and A213. This is the first report to our knowledge of a protein other than ubiquitin interacting with a UIM. To determine whether this interaction with Rpn12 excludes the binding of ubiquitin the $^{1}H/^{15}N$ HSQC spectra of Rpn10 in the apo state, after Rpn12, K48-Ub_2 and Rpn12 and K48-Ub_2 addition were compared (Figure 3.9B). The peak position in

the latter experiment was in almost all cases identical to that of the K48-Ub_2 titration but most importantly distinct from that of the Rpn10-Rpn12 complex. This experiment clearly demonstrates that binding to the UIM is an exclusive event and K48-Ub_2 can effectively displace Rpn12 at the given concentrations.

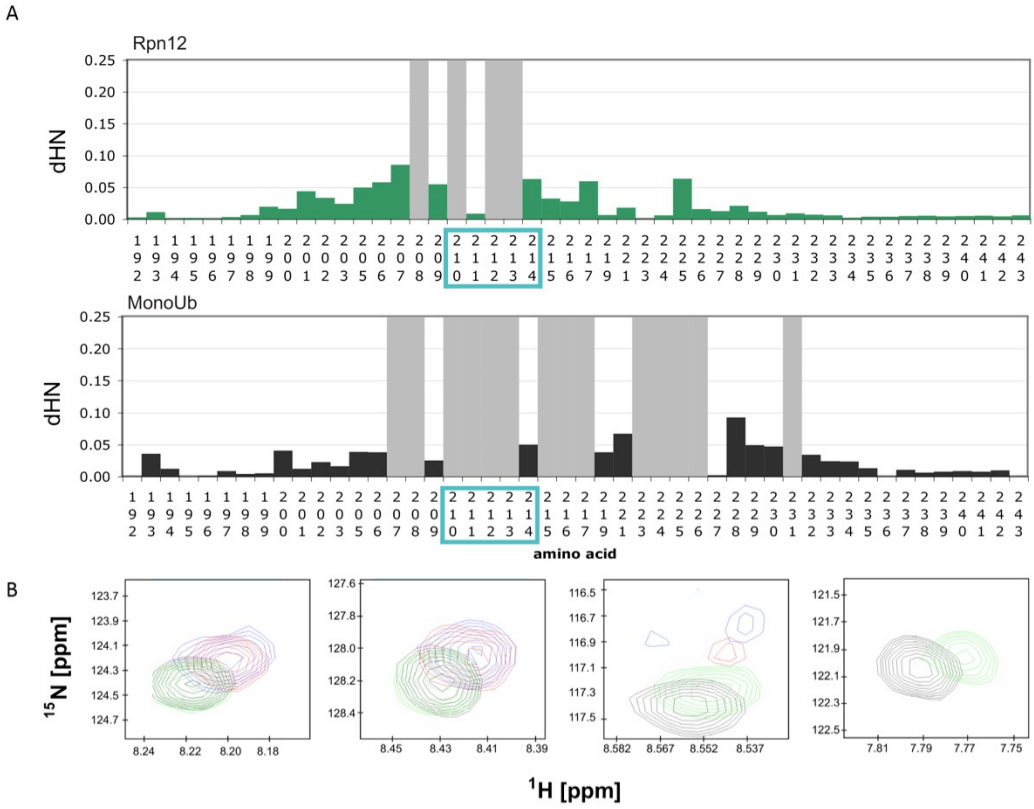


Figure 3.9. Rpn12 interacts with the Rpn10 UIM and competes for ubiquitin binding. (A) Comparison of CSP of Rpn10 residues 192-243 upon addition of Rpn12 (top panel) or monoubiquitin (lower panel). In both cases the UIM contributes to most of the interactions. **(B)** Selected residues of Rpn10 in the apo state (black), in the presence of 3-fold excess of Rpn12 (green), in the presence of equimolar K48-Ub_2 (blue) and at 1:3:1 ratios of Rpn10:Rpn12:K48-Ub_2 (red).

To quantitate our findings, Rpn12 was fluorescently labelled with the thiol-reactive dye Oregon Green 488 maleimide and binding to Rpn10-FL, Rpn10 VWA and Rpn10 UIM was measured (see Section 6.7 for experimental details). All three constructs bound to Rpn12 (Figure 3.11A). The isolated VWA domain bound most weakly, with a measured K_d of 75 ± 11 μM. The UIM and the full-length protein bound more strongly and with equivalent affinities (18 ± 9 μM and 7.3 ± 1.8 μM, respectively, Figure 3.11A, B). To calibrate this method with ITC the interaction between the Rpn10 UIM and K48-Ub_2 was also measured. Using this method, a K_d of 9 ± 2 μM was

determined for the interaction, a similar value to that measured by ITC (13.9 ± 0.6 μM), (Figure 3.11B).

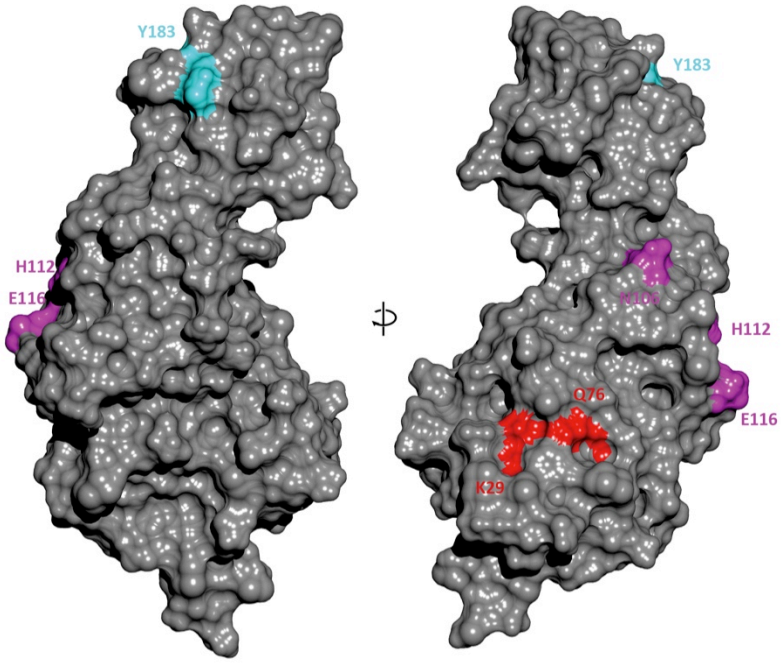


Figure 3.10. Rpn12 mutagenesis. The three surface patches mutated to alanine are highlighted in cyan (Y183), purple (N106, H112, E116) and red (K29, Q76).

In order to characterize the interactions between Rpn12 and Rpn10 in more detail, a series of Rpn12 mutants was created. Based on Rpn12 surface analysis (Figure 3.7) three patches were mutated to alanine: K29/Q76, N106/H112/E116 and Y183 (Figure 3.10). After confirming that these mutants were correctly folded by comparing their 1D ^{1}H NMR spectra to that of the wild-type protein, their effect on Rpn10 binding was studied by NMR and fluorescence anisotropy. Wild type Rpn12 and mutants were added in fivefold excess to ^{15}N-labelled Rpn10, in order to detect and compare the extent of binding. Using this approach, the Y183A and the triple mutations showed identical chemical shift changes compared to wild-type Rpn12, suggesting that binding to Rpn10 is not affected by these mutations. However, the Rpn12 29/76 mutant completely abolished binding to Rpn10, as the spectrum in the presence of the mutant was identical to that in the unbound state of Rpn10 UIM. This suggests that the 29/76 mutation is located at the main binding site mediating the interaction with Rpn10.

To further probe the interactions, fluorescence anisotropy titrations were subsequently performed in 384-well plates with a total reaction volume of 20 µl. Increasing concentrations of Rpn10 were titrated into fluorescently labelled Rpn12 (prepared as described in 6.3.4) and the fluorescence anisotropy was recorded. These experiments confirmed the essentially wild-type affinity of the Rpn12 N106A/H112A/E116A mutant for Rpn10 (Figure 3.11A). However, the measured affinity of the K29A/Q76A mutant was only reduced threefold from K_d=7.3 ± 0.8 µM to 18.3 ± 1.3 µM. This value is still higher than that of the Rpn10 VWA domain for wild-type Rpn12 (75 ± 11 µM, Figure 3.11A). It is unclear why no CSPs could be observed in the NMR titration.

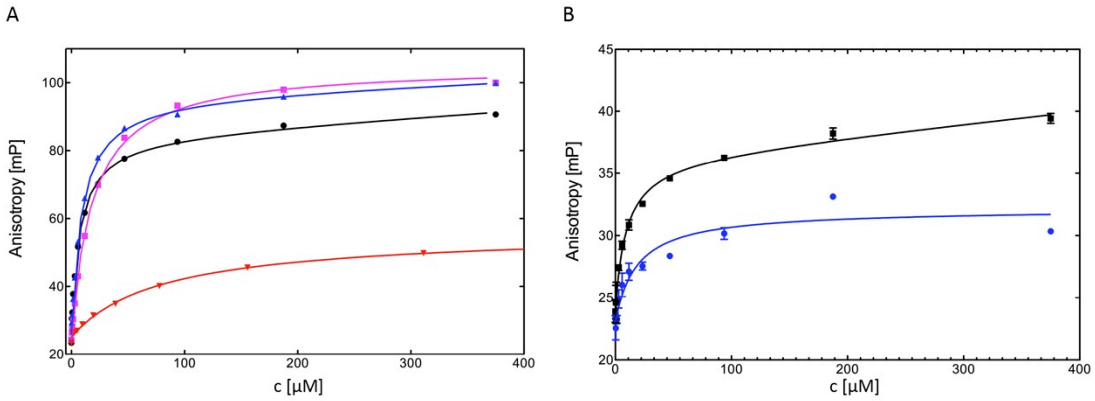


Figure 3.11. Interaction between Rpn10 and Rpn12 as measured by fluorescence anisotropy. (A). The VWA domain (red) had an affinity to wild-type Rpn12 of 75 µM while the full-length protein bound with 7.6 µM (black). The N106A/H112A/E116A mutant (blue) essentially showed wild-type behaviour (8.8 µM) but the K29A/Q76A mutation (magenta) reduced the affinity to 18.3 µM. **(B)** Both K48-Ub_2 and Rpn12 bound to the C-terminus (196-243) of Rpn10. Compared to full-length Rpn10 the affinity towards Rpn12 is reduced to 18 µM (blue) but that of K48-Ub_2 is essentially unaltered 9 µM (black).

3.6 Discussion

TPR domains are extremely versatile units involved in protein-protein interactions. This is also the reason why no conserved binding site has been identified. Indeed, interactions within protein complexes have been mapped to as diverse faces of the proteins as a groove formed by a slight offset of the repeats in some TPR domains, loops connecting TPRs or surfaces formed by the α-helices of a TPR. The $Rpn12_{1-224}$ crystal structure is an Rpn12-like atypical PCI domain and contains both a TPR and a winged helix domain. Analysis of the solvent-exposed surface identified patches that are conserved amongst Rpn12 orthologs and are thus potential protein-protein interaction sites. Interestingly all of these patches are concentrated on the convex side of the molecule with the only other notable patch being located between TPRs1 and 2 (Figure 3.7B).

VWA domains are also protein-protein interaction domains. Its unique location within the 19S RP between base and lid makes Rpn10 not only a ubiquitin receptor but also a mediator and it has been shown to interact with various other subunits of the 19S RP (Rpn1, Rpn12 and Rpn9). VWA domains are involved in a number of signalling pathways and a number of binding sites have been mapped to different parts of the VWA surface (Springer 2006). This is compatible with the proposed role of Rpn10 as a mediator between base and lid. In integrins the metal ion dependant adhesion site (MIDAS) undergoes a conformational change upon binding that is transduced to the C-terminal domain. In Rpn10 the MIDAS is not conserved and no metal ion can be coordinated. This makes signal transduction through the domain unlikely – but not impossible – because it is the change in the metal ion coordination that triggers the conformational change. Interestingly, however, the structural organisation of the loops contributing to the MIDAS is very well conserved in Rpn10 (Figure 2.7B, C). Together with a patch of conserved surface residues and a solvent-exposed tryptophan residue the top loop region of Rpn10 has the features of a potential protein-binding site. The mutations we introduced into Rpn10 (Figure 2.8C) did not have any effect *in vivo* in the context of a *Δrpn10, Δrhp23 S. pombe* strain. This assay, however, might just address the ubiquitin

binding functionality of Rpn10. It would be interesting to characterise the impact of these mutations on the interactions with proposed binding partners such as Rpn9, Rpn11 and Rpn12.

All of these interactions had been proposed to be mediated via the VWA domain so our finding that the UIM is involved in the interaction with Rpn12 was unexpected. The initial experiments that indicated this behaviour were surface-based SPR measurements in which the binding curves exhibited an unusually slow off-rate (Figure 3.8A). SPR experiments can be problematic because the immobilised binding partner can be oriented in a specific way that favours or impedes binding. Nevertheless, it is possible to retrieve information on the binding mode by measuring the dissociation rate after increasingly long association times. If the dissociation rate is significantly higher after a short analyte injection but slower after a long association, a slow event strengthening the interaction (such as a conformational change) is likely to have happened. This experiment did not unambiguously demonstrate either behaviour as the differences in the dissociation rates are very small (Figure 3.8C) but a slightly decreased dissociation rate could be observed after the longest sample injection. The two binding sites we characterised could explain this behaviour: a stronger interaction via the UIM and a much weaker interaction via the VWA binding site. But even if the VWA domain is involved in Rpn12 binding the question still remains as to why Rpn12 interacts with Rpn10 in a way that can exclude ubiquitin-binding. Most importantly, the affinities of the Rpn10 UIM for both Rpn12 and K48-Ub_2 are modest and very similar. This might allow displacement of Rpn12 by ubiquitin chains and could provide a potentially powerful way of modulating ubiquitin recognition by Rpn10. It could be possible that only ubiquitin chains of a certain kind and length effectively compete with Rpn12 resulting in a larger discriminatory effect than the one resulting from the Rpn10-ubiquitin interaction alone.

It has been demonstrated that an intact UIM is necessary to rescue an $Rpn12_{1\text{-}197}$ temperature-sensitive *S. pombe* strain by overexpression of Rpn10. Rpn12 is essential for proteasomal integrity and deletion of the gene is lethal. The truncated construct removes the WHD, highlighting the importance of this domain

for proteasome integrity. Over-expression of Rpn10 in this strain rescues the temperature-sensitive phenotype, possibly by helping to maintain the structure of the 19S RP. It is conceivable that a higher concentration of Rpn10 can help to restore the interaction with Rpn12 despite the loss of one potential binding site on the WHD. Attempts to co-crystallise Rpn10 and Rpn12 failed. However, this is perhaps not surprising given the relatively weak interaction. Unfortunately we were also not able to characterise the binding sites on both proteins sufficiently well to use this data as restraints to generate a model of the complex with the CNS software HADDOCK (Dominguez et al. 2003; Brünger et al. 1998). Another route to identify the Rpn12 UIM binding site would be soaking of $Rpn12_{1-228}$ crystals or co-crystallisation with a short Rpn10 peptide containing only the central UIM binding motif as identified by NMR. Small-angle X-ray scattering (SAXS) could also be used to calculate a surface envelope that could be used as additional restraints in creating a model based on the two crystal structures. Such a model of the Rpn10-Rpn12 complex would possibly be of sufficient size to locate unambiguously within an EM envelope. For this experiment W. Baumeister has already kindly agreed to provide us with a new *S. pombe* EM dataset that his group is currently collecting. This would be a significant step towards a better understanding of 19S regulatory particle organisation.

4 Rhp23 Binds Selectively to K48-Ub_2

4.1 Introduction

The gene encoding Rad23, the *S. cerevisiae* ortholog of Rhp23, was discovered in an attempt to decipher the reason for UV-sensitivity in a yeast strain expressing abnormally high levels of iso-2-cytochrome C (Rothstein et al. 1980). The deletion of two adjacent genes was later found to be responsible for this phenotype. One of the genes (*Rad23*) was the sole cause of increased UV-sensitivity (McKnight et al. 1981). It was later revealed that it participates in a process that detects and repairs DNA lesions that distort the DNA helix known as nucleotide excision repair (NER). NER is well conserved between yeast and humans and most research has been undertaken using either Rad23 or the human orthologs hHR23a or hHR23b. The NER pathway can be divided into four steps (reviewed in Dantuma et al. 2009): (i) recognition of the DNA lesions by Rad4/XPC, which is exclusively found in complex with Rad23/hHR23. Rad4 can only efficiently bind to distorted DNA and thus indirectly recognises DNA lesions that destabilise the DNA helix. These lesions such as cyclobutane pyrimidine dimers and 6-4 phosphoproducts are primarily caused by short wavelength UV light. (ii) These lesions result in the recruitment of more factors including the DNA helicases Rad3/XPD and Rad25/XPB that unwind the DNA. (iii) The endonucleases Rad1/XPF and Rad2/XPG then remove a 25-30 nucleotide long patch of single stranded DNA including the lesion followed by (iv) repair synthesis and ligation. Despite decades of research the specific role of Rad23/hHR23 in this process remains obscure and currently only a function stabilising Rad4/XPC is well established (Ng et al. 2003; Ortolan et al. 2004).

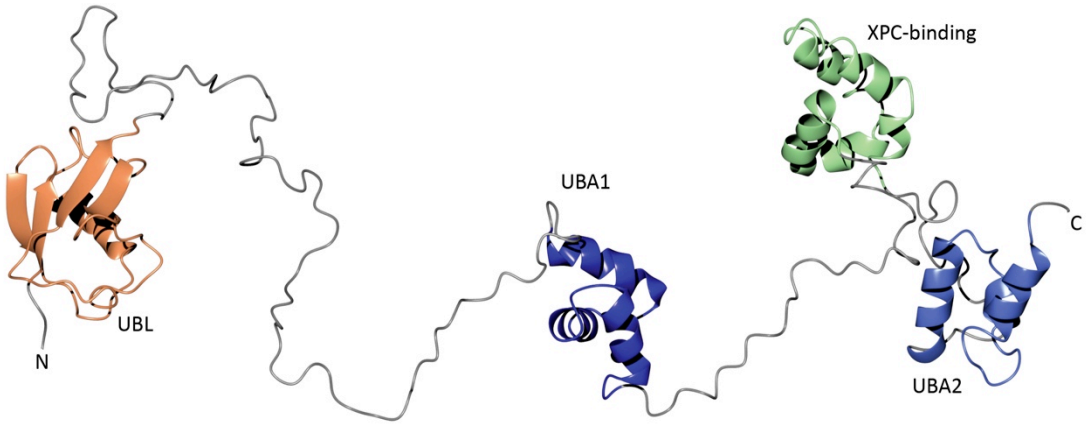


Figure 4.1. NMR structure of the human ortholog of Rad23 (PDBID 1OQY). A C-terminal UBL domain (orange) is connected via a long unstructured linker to two UBA domains (blue) that are separated by an XPC-binding domain (green).

In addition Rad23 plays an important role in ubiquitin-mediated proteasomal degradation (L. Chen et al. 2002). It contains four domains: a C-terminal ubiquitin-like domain (UBL) and two UBA domains separated by the Rad4-binding domain (Figure 4.1). UBL domains are typically located at the extreme N-terminus of a protein and adopt a ubiquitin-like β-grasp fold despite only 20-25% sequence identity. They are capable of interacting with UBDs (Ryu et al. 2003) similarly to ubiquitin and can also interact with the proteasomal subunits Rpn1 and Rpn2 that do not interact with ubiquitin directly (Schreiner et al. 2008; Elsasser et al. 2002). The Rad23 UBL domain can be exchanged for ubiquitin without impairing NER function (Lambertson et al. 2003). However, the functionality of Rad23 then becomes dependant on enzymes involved in targeting proteins with an N-terminal ubiquitin moiety (ubiquitin fusion degradation (UFD) substrates) for proteasomal degradation.

Both UBA domains of Rad23 are capable of binding ubiquitin but have been reported to do so with different specificities. UBA1 possessed a slight preference for K63-linked polyubiquitin chains whereas UBA2 selectively bound K48-Ub_4 (Raasi et al. 2005). It is also important to note that earlier studies demonstrated that both UBA domains of hHR23a form interactions with the UBL domain (Walters et al. 2003) that may alter the affinities and/or selectivity towards different ubiquitin species.

Rad23 is a member of a class of proteins that has been defined by their possession of both UBL and UBA domains. Because they possess domains that enable potentially simultaneous engagement with ubiquitin and with the proteasome but are not integral stoichiometric subunits of the proteasome, they have been described as shuttle proteins. A model has been proposed in which the proteasome-interacting UBL domain of the shuttle protein is occupied by a UBA domain unless ubiquitylated substrate is bound. The ubiquitin chain of that substrate can displace the UBL domain upon binding allowing its interaction with the proteasome. After recycling of the ubiquitin, the re-established UBL-UBA interaction finally results in the protein's dissociation from the proteasome. This is an appealing model and several studies have demonstrated that UBL-UBA proteins play a positive role in substrate degradation (Funakoshi et al. 2002; Lambertson et al. 1999; Rao et al. 2002). However, *in vitro* studies have often found an inhibitory effect of Rad23 (Raasi et al. 2003). These seemingly contradictory results are conceivable because UBA domains can principally mask ubiquitin chains and UBL domains can block binding sites at the proteasome but, at the same time, are also able to foster interactions between the proteasome and ubiquitylated substrates. Fine-tuned concentration-dependant interplay between the various ubiquitin receptors seems to be necessary for efficient degradation of all substrates. This is reflected in recent studies proposing functions such as chain-length sensing by the interplay between Rpn10 and Dsk2 (Figure 2.4, Zhang, Chen, et al. 2009).

Another interesting property of UBL-UBA proteins is the ability to interact directly via their UBL domain with ubiquitin ligases capturing the ubiquitin chain while the substrate is still bound to the ligase. For example Rad23 and Dsk2 both bind Ufd2 (Richly et al. 2005; I. Kim et al. 2004) whereas Ddi1 interacts with Ufo1 (Kaplun et al. 2005).

Rhp23 and its orthologs are the only proteasome associated UBL-UBA proteins with more than one UBA domain. The second UBA domain might promote avid interactions with longer polyubiquitin chains but *in vivo* and *in vitro* studies have shown that the interaction of UBA2 with ubiquitin has no impact on Rad23 function (Bertolaet et al. 2001; L. Chen et al. 2002). Other than binding ubiquitin, a

stabilising effect on Rad23 has more recently been attributed to UBA2 (Heessen et al. 2005). This study elegantly demonstrated that GFP-based substrates of the UPS were stabilised by the C-terminal addition of Rad23 UBA2 but not UBA1. Other C-terminal UBA domains of Dsk2, Ddi1 and Ede1 exhibited the same stabilising effect. This might be a necessary feature of proteins that have to interact with the proteasome to avoid degradation.

To assess the effect of the ubiquitin-binding properties of the Rhp23 UBA domains we designed mutations that abolish their affinities for ubiquitin. Based on NMR studies with the human ortholog hHR23B (Ryu et al. 2003) and structural alignment we designed mutations in both UBA domains to disrupt the ubiquitin binding properties (M157A, L183A within UBA1 and L303A, Y358A within UBA2). In order to de-convolute the contributions of both UBA domains each of them was mutated in the context of the full-length protein (FL-UBA1 m., FL-UBA2 m., FL-UBA1+2 m.) and within the context of the individual domains (UBA1 m., UBA2 m.). The effects of these mutations were compared to the wild-type Rhp23 (FL-wt, UBA1 wt, UBA2 wt) proteins.

4.2 Expression and Purification of Rhp23

The eight Rhp23 constructs tested in this study were most successfully expressed in *E. coli* in terrific broth medium following induction with 500 μM IPTG (Section 6.3.5). Glutathione sepharose affinity chromatography resulted in ~100 mg GST-fusion protein per litre medium for each of the individual UBA domains and ~20 mg for full-length Rhp23. Size exclusion chromatography (SEC) was used as the final purification step and resulted in >90% pure protein (Figure 4.2).

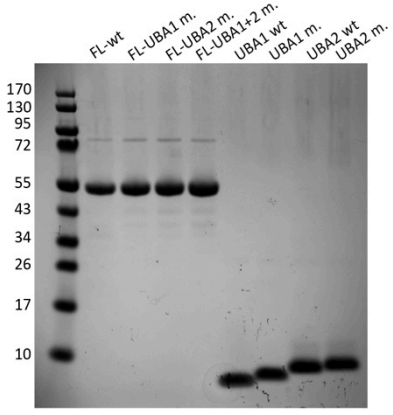


Figure 4.2. Purified Rhp23 constructs. A combination of glutathione sepharose followed by size-exclusion chromatography resulted in pure protein

4.2.1 Generation of Fluorescently Labelled Ubiquitin

As described previously, UBL-UBA proteins such as Rhp23 not only bind ubiquitin via their UBA domain(s) but also form inter- and intra-molecular complexes via the interaction of the UBA with the UBL domain. ITC is not suitable to measure the binding of ubiquitin chains to these proteins because the UBL-UBA interactions make unambiguous data analysis impossible without the introduction of a substantial number of mutations. To overcome these problems, we decided to use fluorescence anisotropy to measure binding to fluorescently labelled ubiquitin species.

The most specific method for covalently attaching fluorophores to a protein is to exploit thiol-reactive groups as found in dye derivatives such as the Oregon Green 488 maleimide used here. At the correct pH they exclusively react with reduced cysteine residues in the protein. As ubiquitin does not contain any cysteine residues the ubiquitin mutants generated for the defined synthesis of diubiquitin (Section 6.3.2) were exploited without the need for any further mutagenesis. Because the fluorescent dye coupled to C48 of ubiquitin negatively influenced its binding properties, a double mutant (K48R, K63C) was produced to generate fluorescently-labelled K48-Ub_2. All ubiquitin mutants expressed at about 30 mg of protein per litre terrific broth and were purified by perchloric acid precipitation and ion exchange chromatography (Figure 4.3A). The labelled proteins could be separated efficiently from unreacted dye by SEC and eluted as a single peak (Figure 4.3B). The labelled and purified ubiquitin was then used to generate diubiquitin as described in Section 6.3.2 and was purified by ion-exchange chromatography (Figure 4.3C). This resulted in pure mono-, K48- or K63-linked diubiquitin carrying the fluorescent label on C63 (Figure 4.3D). C63-labelled K48-Ub_2 bound to Rpn10 with an affinity very similar (9 μM, Figure 3.11D) to that determined by ITC (13.9 μM, Figure 2.10B).

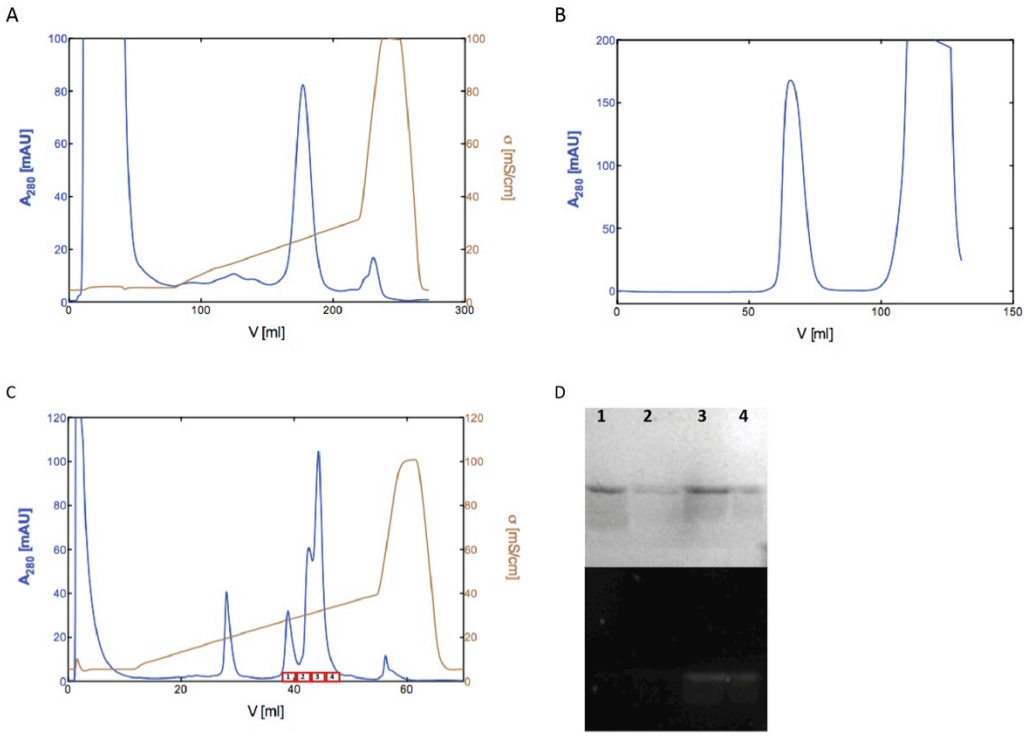


Figure 4.3. Purification of fluorescently labelled ubiquitin. (A) The expressed ubiquitin mutants were first purified by cation exchange chromatography and then labelled with the thiol-reactive dye Oregon Green 488 maleimide. **(B)** Excessive dye was removed by size-exclusion chromatography. **(C)** After the ubiquitin chain-formation reaction, pure labelled Ub_2 was obtained by cation exchange chromatography. **(D)** Peak fractions highlighted in (C) were resolved by SDS PAGE and visualised by Coomassie staining (top panel) and UV-light excitation (lower panel). The separation of labelled from unlabelled Ub_2 is clearly demonstrated.

4.3 The Interactions of Rhp23 With Ubiquitin

Rhp23 contains an N-terminal UBL and two C-terminal UBA domains separated by an XPC-binding domain. It has been speculated that the main function of UBA2 of the *S. cerevisiae* ortholog Rad23 is not as a ubiquitin receptor but rather that it serves as an internal stabilisation signal to avoid degradation by the proteasome (Heessen et al. 2005). Our collaborator C. Gordon tested full-length Rhp23 containing either one or both mutant UBA domains for their ability to rescue a *Δrpn10, Δrhp23* temperature-sensitive *S. pombe* strain. Only Rhp23 containing the mutated UBA2 domain could rescue the strain at 36 °C (**Error! Reference source not found.**).

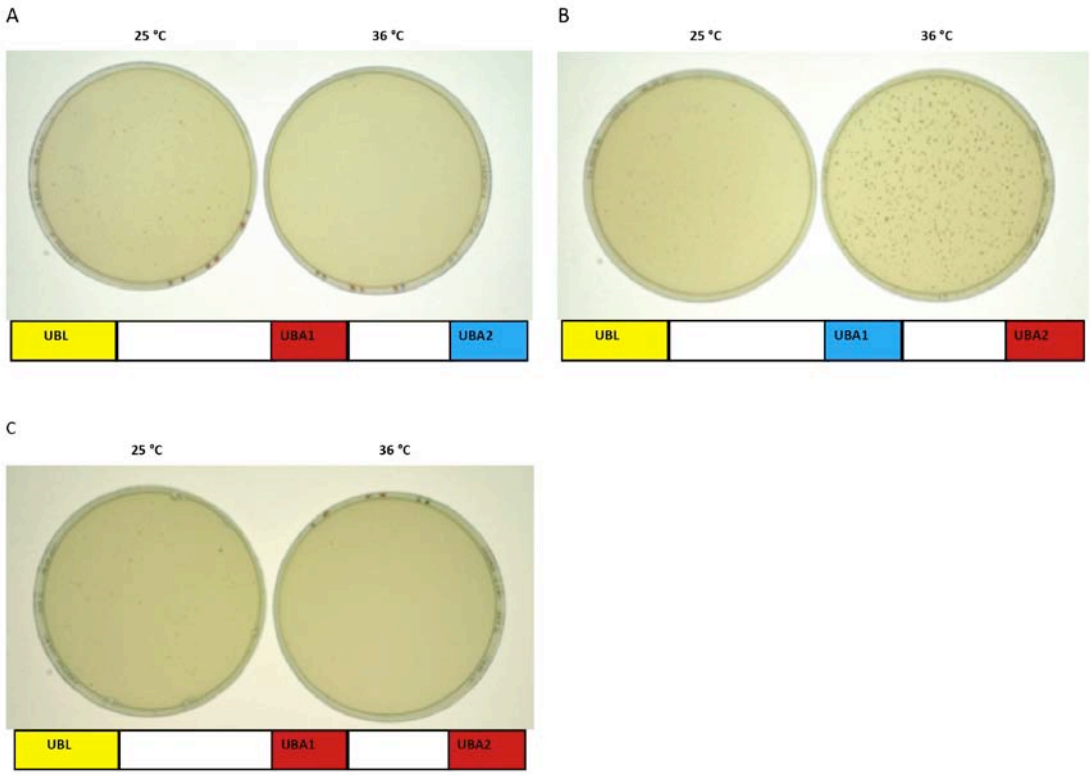


Figure 4.4. Rhp23 rescue experiments. Wild-type and mutant (unable to bind ubiquitin) UBA domains are coloured blue and red, respectively. (A) An intact UBA2 in the context of full-length *Rhp23* is not sufficient to rescue the *Δrpn10, Δrhp23 S. pombe* strain. (B) Mutation of UBA2 has little or no effect in this assay. (C) *Rhp23* in which both UBAs are mutated cannot rescue the *Δrpn10, Δrhp23 S. pombe* strain.

In order to elucidate the reason for this behaviour, we used the fluorescence anisotropy assay as described (Section 6.7) to characterise the interactions of the different Rhp23 full-length (Rhp23-FL) constructs as well as the isolated UBA domains with ubiquitin. We first set out to establish whether Rhp23 binds preferentially to monoubiquitin, K48-Ub2 or K63-Ub2. We tested the wild-type constructs of both the full-length protein and the individual UBA domains. All constructs preferentially bound K48-Ub2. The strongest (~15-fold) discriminatory effect was detected for the full-length protein where the affinity increased from 138 ± 47 μM to 103 ± 36 μM to 7.0 ± 0.8 μM for monoubiquitin, K63-Ub2 and K48-Ub2, respectively (**Figure 4.5**A). The UBA1 domain in isolation bound monoubiquitin with an affinity of 126 ± 80 μM and discriminated against K63-Ub2 about 2.5-fold (51 ± 16 μM). K48-Ub2 had an affinity of 18.8 ± 4.7 μM for the UBA1 domain, only threefold weaker than the full-length protein (**Figure 4.5**B). It was not surprising then

that the UBA2 domain generally exhibited much weaker binding to all ubiquitin chains tested. It bound K48-Ub2 with 48 ± 13 μM affinity and the interactions with both monoubiquitin and K63-Ub2 at >200 μM were too weak to determine a precise value with this method (**Figure 4.5**C).

As these experiments clearly established the K48-linked ubiquitin chain selectivity of Rhp23, we focused on K48-Ub_2 to characterise the different UBA-mutant constructs. As expected from the previous results, mutation of UBA2 had almost no impact on the affinity for K48-Ub_2 in the context of Rhp23-FL. The overall affinity decreased from 7.0 ± 0.8 μM to 11.9 ± 0.7 μM (Figure 4.5D). We suspect the residual ubiquitin-binding properties of UBA2 (>200 μM) to be the reason for this slightly stronger interaction when compared to the UBA1 domain alone (Figure 4.5E). The effect of mutating UBA1, however, was much stronger. In the context of the full-length protein the affinity dropped to 34 ± 7 μM, which is in good agreement with the value obtained for UBA2 in isolation (48 ± 13 μM). The isolated and mutated domain had no detectable binding (Figure 4.5D, E). Not surprisingly the mutation of both UBA domains in Rhp23-FL had the severest effect reducing the affinity to >200 μM. These results are in agreement with the *in vivo* experiments and underline the importance of the ubiquitin-binding properties of UBA1.

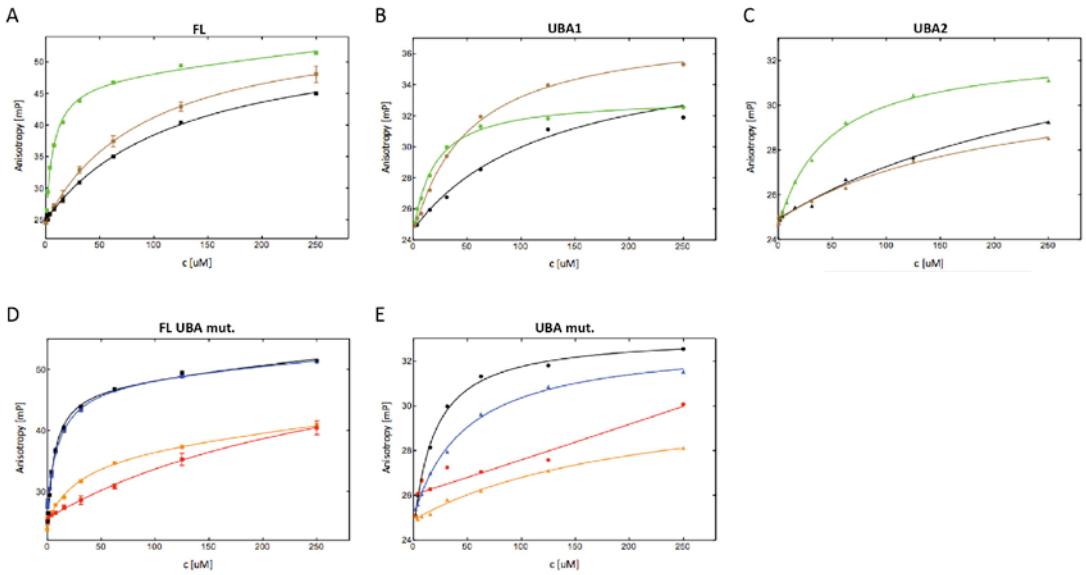


Figure 4.5. Rhp23-ubiquitin interactions as measured by fluorescence anisotropy. (A) Full-length Rhp23 bound K48-Ub_2 with 7.0 μM affinity (green) but both monoubiquitin (138 μM, black) and K63-Ub_2 (103 μM, brown) more than tenfold weaker. The individual UBA domains displayed the same behaviour but to a different extent. **(B)** UBA1 had an approximately six-fold lower affinity for monoubiquitin compared with K48-Ub_2 (19 μM vs. 126 μM) but the discrimination towards K63-Ub_2 was less (51 μM). **(C)** UBA2 had an overall much reduced affinity for ubiquitin. The affinities for both monoubiquitin and K63-Ub_2 were >200 μM. K48-Ub_2 bound more strongly (48 μM) but much weaker than UBA1. **(D-E)** The effects of mutations within the UBA domains to K48-Ub_2-binding. **(D)** In the context of the full-length protein, mutation of UBA2 (blue) had almost no effect reducing the affinity from 7.0 μM to 10.9 μM. If UBA1 was mutated, however, the affinity decreased about 4.5-fold to 31 μM (orange). Mutation of both UBAs resulted in almost complete loss of binding with an affinity of >200 μM (red). **(E)** Looking at the individual UBA domains, the UBA1 mutation (red) resulted in the loss of binding while mutation of UBA2 retained some residual affinity that was, however, >200 μM (orange).

4.4 Discussion

Previous reports have claimed that UBA1 and UBA2 of Rad23 have different chain selectivity. These SPR experiments, however, used GST-fusion proteins that artificially favour K63-linked ubiquitin chains (Sims, A. Haririnia, et al. 2009). Our fluorescence anisotropy experiments clearly demonstrated a preference for K48-Ub_2 over both monoubiquitin and K63-Ub_2 and confirmed previous findings that UBA2 has very weak affinity for all ubiquitin chains tested (Figure 4.5). It has been shown that the UBA2 but not UBA1 domain is essential for the stability of Rad23 *in vivo* by preventing the degradation of Rhp23 (Heessen et al. 2005). This functionality is probably independent of its ubiquitin-binding properties and it was speculated that

UBA2 provides a compact fold that does not allow the initiation of unfolding, one condition for proteasomal degradation (Prakash et al. 2004). *In vivo* experiments performed by our collaborator C. Gordon have shown that the ubiquitin-binding properties of UBA1 but not UBA2 are essential for the rescue of a *Δrhp23*, *Δrpn10 S. pombe* strain function (**Error! Reference source not found.**).

The role of UBA2 remains poorly understood. Its ubiquitin binding properties are not essential even in a *Δrpn10* background and Rad23 without any UBA domain can still function in DNA repair (Bertolaet et al. 2001). Its stabilising effect on Rhp23 has clearly been demonstrated (Heessen et al. 2005) but it is questionable whether this is the sole function of UBA2 or it is a property commonly found in the C-terminal domains of proteins that reversibly associate with the proteasome. This idea is supported by the findings that the single UBA domains of Ddi1 and Ede1 exhibit the same stabilising effects.

Although not essential, the Rhp23 UBA2 domain did have an effect on the affinity of the full-length protein towards the different ubiquitin species. While the affinity towards K48-Ub_2 only increased by about 1.5-fold it did, however, contribute to a circa 5.4-fold increased specificity of the full-length protein toward K48-Ub_2 when compared with UBA1 alone (from a 2.7-fold difference between K48-Ub_2 and K63-Ub_2 in UBA1 to a 14.7-fold difference in Rhp23-FL). This might reflect the beginning of an adaptation towards a higher specificity for K48-linked polyubiquitin chains by exploiting the avidity effects arising from a second UBA domain.

5 Discussion and Concluding Remarks

It is remarkable how a single protein can be utilised in such a variety of signalling pathways by simply altering its polymerisation state. When the ubiquitin proteasome system was discovered it was probably mostly considered to be an irreversible route to degradation at the proteasome. Three decades of research have elucidated many of the steps involved and drawn the picture of an extremely diverse system that is highly convoluted and most importantly still incomplete. The precise role of the different polyubiquitin chains is still not clear, more ubiquitin receptors have been identified than are necessary for the current model of ubiquitin recognition and the mechanism by which the 19S RP binds and unfolds substrates while recycling the ubiquitin moieties remains to be understood.

One of the most pressing questions is the reason for the redundancy of ubiquitin receptors. Five proteasome-associated receptors have been identified that are all (and also in combination) non-essential in *S. cerevisiae* (D. Finley, personal communication). The existence of more ubiquitin receptors is therefore highly likely but it is questionable whether one essential "master" receptor exists. More likely, several receptors have evolved with each of them having slight preferences for a subset of substrates while retaining the ability to recognise and deliver other substrates in situations of stress (such as the loss of another receptor). The evolution of ever more complex multicellular organisms required additional levels of regulation including the UPS. It is likely that this specialisation involves the development of substrate-subset-specific polyubiquitin receptors. This idea is supported by findings that this redundant functionality is reduced in higher eukaryotes that do rely on Rpn10 during embryonic development (Hamazaki et al. 2007).

Eight different types of polyubiquitin chains exist (excluding branched chains) and each encodes a different signal. It is therefore necessary for the system to be able to effectively discriminate between them. It might be supposed that this role would fall to specialised ubiquitin receptors. However, even though some receptors have preferences for binding one chain over another *in vitro,* overall their abilities to

discriminate seem rather modest. This is especially true in the presence of the very high cellular ubiquitin concentrations. Unless other ubiquitin receptors with much higher selectivity are discovered, it is more likely that a well-balanced system has evolved in which modest changes in the affinities are integrated over the whole pathway and culminate in an increased likelihood of degradation. In an extreme case, a ubiquitin-protein ligase could already temporarily be associated with the proteasome while producing K48-linked polyubiquitin. This chain then tethers the substrate efficiently to the proteasome for a time long enough to initiate unfolding and degradation. In the case of a "wrong" chain ending up at the proteasome, in addition to a weaker affinity for proteasomal ubiquitin receptors, DUBs could enhance this effect by preferentially removing those chains.

While the mechanism of proteasomal proteolysis is very well understood thanks, in part, to structures at atomic resolution of the 20S catalytic particle, the precise functions of the 19S regulatory particle are still rather obscure. We know it binds polyubiquitylated substrates, recycles the ubiquitin and unfolds the target protein but for no step is the mechanism known. It becomes even more complicated – but also more interesting – because substrate denaturation seems to be coupled to deubiquitylation in order to avoid degradation of the ubiquitin moieties.

In this thesis we hope to contribute to answering questions related to both 19S organisation and ubiquitin recognition. In the absence of any high-resolution structure of the 19S RP the two novel crystal structures presented here can serve as a starting point to the stepwise structural elucidation of this crucial cellular complex. One of them, Rpn10, serves as a ubiquitin receptor and we have analysed its ubiquitin-binding properties in detail. We have demonstrated *in vitro* that Rpn10 preferentially binds to K48-Ub_2 over both monoubiquitin and K63-Ub_2. The discriminatory effect, however, is modest and it remains to be determined whether this is sufficient to mediate Rpn10 function *in vivo*. In this respect the interaction of Rpn10 and Rpn12 may be relevant. Here it is shown for the first time that a UIM can bind to a protein unrelated to ubiquitin. The fact that this binding event excludes a simultaneous interaction with ubiquitin raises the question as to whether the masking of the UIM by Rpn12 has a regulatory function. Masking of a UIM by

monoubiquitylation of a nearby lysine residue to negatively regulate its binding properties has recently been described (Isasa et al. 2010) and intra-molecular interactions in UBL-UBA molecules might serve a similar purpose. For a detailed description of the effects of the Rpn10-Rpn12 interaction careful analysis in the context of intact proteasomes will be necessary. The interaction between the UIM and Rpn12 is weak but may not be as easily disrupted by ubiquitin if Rpn10 and Rpn12 are already bound in a tight complex within the 19S regulatory particle. Thus, any discriminatory effect might only exist when these proteins are part of the proteasome.

Analysis of the crystal structures of Rpn12 and the Rpn10 VWA domain identified several residues that when mutated could interrupt the interactions with potential binding partners. In *in vivo* rescue experiments the Rpn10 mutations exhibited wild-type like behaviour in a *ΔRpn10*, *ΔRhp23 S. pombe* strain, possibly because this rescue does not directly rely on the VWA domain. Once initial recombinant expression difficulties are overcome, we will be able to measure their interactions with Rpn12 and other 19S RP subunits such as Rpn9 *in vitro*. In Rpn12, the mutation of K29 and Q76 to alanine did negatively affect binding to Rpn10 and together with an assignment of the Rpn12 $^{1}H/^{15}N$ HSQC spectrum we are hoping to identify the Rpn10-Rpn12 binding interface. However, a complex with a size of >50 kDa is a challenging sample for NMR and we are currently aiming to determine better restraints for the calculation of a model of the complex structure that could be docked into a high-resolution cryo-EM map. This would be the first step toward a better understanding of the organisation and thus function of the 19S RP.

6 Materials and Methods

6.1 Buffers and Reagents

All experiments were performed at 20 °C, if not stated otherwise. Chemicals were obtained from Sigma. Primers were purchased from VHBio/IDT and restriction enzymes, DNA ladders, Antarctic Phosphatase, T4 DNA Ligase and Phusion DNA Polymerase were from New England Biolabs. Gel Extraction and MiniPrep kits from Qiagen and all FPLC columns and column materials from GE Healthcare. The In-Fusion cloning kit and chemically competent molecular cloning grade cells were purchased from Clontech. The plasmids for the expression of E2 enzymes were a kind gift from R. Poon (Cdc34) and D. Komander (UBC13, UEV1A). The murine E1 plasmid was a gift from K. Iwai and the protein was prepared by N. Brown. GST-tagged full-length Rpn10, Rpn10 VWA (1-195) and full-length Rpn12 constructs subcloned into pGEX-6P-1 (GE Healthcare) were a gift from Jean-François Trempe.

6.2 Molecular Biology Methods

6.2.1 Competent Cell Preparation and Transformation

Chemically competent *E. coli* cells of the expression strains B834, BL21(DE3) Star and Rosetta 2 were prepared by a method described previously (Inoue et al. 1990). 250 ml cultures were grown in 2 l flasks in Super Optimal broth with Catabolic repression (SOC) at 18 °C to an OD_{600} of 0.75. The cells were then harvested by centrifugation at 2500× g and 4 °C for ten minutes. The pellet was then resuspended in 80 ml ice-cold transformation buffer (10 mM HEPES, 55 mM $MnCl_2$, 15 mM $CaCl_2$, 250 mM KCl, pH 6.7). After another identical centrifugation step the cells were resuspended in 20 ml ice-cold transformation buffer and DMSO was slowly added to a final concentration of 7%. Subsequently, the suspension was incubated on ice for ten minutes before being aliquotted (200 µl) and flash-frozen in liquid nitrogen prior to storage at -80 °C.

A few nanograms of plasmid DNA were transformed into chemically competent *E. coli* cells. The *E. coli* DH5α strain was used for cloning, the BL21(DE3) Star strain was used for general protein expression, the Rosetta 2 strain for the expression of genes containing rare codons and the methionine auxotroph strain B834 for expression of selenomethionine derivatives. 50 µl aliquots of the cells were carefully thawed on ice and incubated for 30 minutes with no more than 5 µl plasmid DNA. After a 30 s heatshock in a 42 °C waterbath the cells were immediately put on ice for another two minutes. Subsequently, 250 µl SOC were added to the cells and incubated for one hour at 37 °C with vigorous shaking. The cells were then spun down at 500× g, resuspended in 50 µl SOC and spread onto an LB agar plate containing the appropriate antibiotics.

6.2.2 PCR Reactions

Phusion High-Fidelity DNA Polymerase was used in all high-fidelity PCRs. Typically, a 50 µl reaction contained 1 × Phusion HF Buffer, 0.2 mM dNTPs, 0.5 µM forward and reverse primers, 2 ng template DNA and 1 U Phusion DNA Polymerase. The reaction was carried out in a Bio-Rad iCycler apparatus. After an initial 30 s denaturation step at 98 °C, 25 cycles of denaturation (98 °C, 10 s), annealing (65 °C, 30 s) and extension (72 °C, 15 s/kb) were followed by five minutes final extension. Subsequently, the PCR products were checked by 1% agarose gel electrophoresis in TAE buffer (40 mM Tris, 20 mM acetate, 1 mM EDTA, pH 8.5). The products were purified using the Qiaquick Gel Extraction kit (Qiagen).

6.2.3 Restriction Enzyme Digestion

For restriction enzyme/DNA ligation-based cloning the desired coding sequences were amplified by PCR using primers containing different restriction sites compatible with the vector multiple cloning site. Typically, double-digests were performed. 1 µg of both the PCR product and vector were each incubated for three hours at 37 °C with 10 U of the restriction enzymes, 1× the appropriate reaction buffer and BSA in a total volume of 50 µl. Subsequently the products were purified by 1% agarose gel electrophoresis and gel extraction. The linearised vector was then dephosphorylated

with Antarctic Phosphatase by incubating 1 μg DNA with 10 U enzyme for one hour at 37 °C to reduce background of re-ligated empty vector. Prior to ligation the phosphatase was inactivated by incubation at 65 °C for 15 minutes. Ligation the linearised vector with the insert was carried out at 20 °C over night with 80 U T4 DNA ligase, 0.1-0.5 μM linearised vector and threefold molar excess of insert DNA in 10 μl 1× T4 DNA Ligase buffer. After inactivation by incubating the reaction for 10 minutes at 65 °C, 3 μl were transformed into chemically competent DH5α *E. coli* cells.

6.2.4 Ligation-Independent Cloning

For ligation-free cloning linearised vector DNA was produced by PCR to avoid any background from transformed empty vector. The insert was amplified by PCR using primers that at their ends contain 15 bases homologous to the vector ends. Insert and vector DNA were incubated for 15 minutes at 37 °C and at a 2:1 molar ratio in the In-Fusion (Clontech) reaction buffer containing the proprietary enzyme mix. After another incubation period of 15 minutes at 50 °C the reaction was diluted fivefold in TE buffer (10 mM Tris pH 8.0, 1 mM EDTA) and ~8 ng DNA were transformed into chemically competent FusionBlue (Clontech) cells.

6.2.5 Site-Directed Mutagenesis

All mutagenesis was performed by using mutagenic primer pairs containing the desired mutation flanked by 15 complementary bases on each side amplifying the whole plasmid by PCR (Section 6.2.2). Subsequently, the template (unmutated) DNA was digested by addition of 10 U Dpn1 and incubation at 37 °C for one hour. About 5 ng of DNA were then transformed into *E. coli* DH5α chemically competent cells.

6.3 Protein Expression and Purification

Cultures were grown in terrific broth or autoinduction medium (Studier 2005) supplemented with the appropriate antibiotics (carbenicillin at 100 μg/ml, kanamycin at 50 μg/ml and chloramphenicol at 35 μg/ml) at 37 °C with vigorous shaking. At on OD_{600} of 0.6, the temperature was reduced to 20 °C and in case of

non-autoinduction medium protein expression was induced with 500 μM IPTG. The cells were harvested after eight (non-autoinduction medium) or 24 h by centrifugation at 4000× g for 20 minutes at 4 °C, resuspended in HBS supplemented with Roche Protease Inhibitor Cocktail (EDTA free) at the recommended concentration and stored at -80 °C.

6.3.1 Expression of GST-Fusion Proteins

A number of proteins used in this study including all E2 enzymes were expressed and purified as GST-fusions and a similar purification strategy was used in each case. *E. coli* cell pellets were thawed under cold running water, sonicated and the cell lysate clarified by centrifugation at 75 000× g. The soluble fraction was applied to a glutathione-sepharose 4B column (GE Healthcare, typical column volume 10 ml) equilibrated at 4 °C in HBS. Following protein loading, the column was washed to baseline with HBS and then GST-fusion proteins were eluted with 20 mM glutathione in HBS. The protein yield was determined by measuring the A_{280} absorbance. Removal of the GST was achieved by incubation with 3C protease at a molar ratio of 1:500 at 4 °C overnight. Samples were then concentrated (if required) to a volume of 10 ml prior to loading onto a K26 Sephadex 75 (GE Healthcare) size-exclusion column equilibrated in HBS at 20 °C. This step provided an additional purification step and buffer-exchanged the sample to remove the glutathione. For those proteins that co-eluted from the column with GST, fractions containing the protein of interest were pooled and re-applied to a clean glutathione sepharose column and the flow-through collected. Proteins were either used immediately and kept at 4 °C for short-term or at -80 °C for long-term storage.

6.3.2 Ubiquitin

Ubiquitin chains were essentially prepared as described previously (Pickart et al. 2005). This involved the generation of ubiquitin constructs that have either the C-terminal G76 or the lysine responsible for the desired linkage type blocked. Addition of D77 (Ub^{77D}) and mutation of K48/K63 (Ub^{K48C}/Ub^{K63C}) to cysteine achieved this. The ubiquitin mutants were generated in a pGEX-6P-1 vector backbone with the

GST-fusion tag deleted and expressed in terrific broth media at 37 °C for 5 h after induction with 200 μM IPTG. The purification strategy takes advantage of the extraordinary stability of ubiquitin: precipitation of the lysate with perchloric acid at a final concentration of 0.5% precipitated almost all bacterial proteins. For further purification the supernatant was dialysed against 50 mM ammonium acetate pH 4.5, supplemented with 1 mM DTT, applied to a cation exchange column (SP-Sepharose) equilibrated in the same buffer and after washing finally eluted in 20 column volumes with a linear gradient of 0-0.5 M NaCl in 50 mM ammonium acetate pH 4.5. The ubiquitin-containing fractions were identified by SDS-PAGE, concentrated and buffer exchanged by alternating cycles of concentration and dilution in HBS.

For the synthesis of Ub_2, 1 mM of each Ub^{77D} and either Ub^{K48C} or Ub^{K63C} were incubated at 37 °C over night with 1 μM E1, 8 μM E2 and 2 mM ATP in 50 mM Tris pH 8.0, 0.5 mM DTT. To reduce E2-Ub conjugates the reaction was incubated for another 20 minutes with 5 mM DTT and finally brought to pH 4.5 by addition of acetic acid. Unreacted ubiquitin and the enzymes were removed by the same cation-exchange chromatography described above.

To fluorescently label the ubiquitin, 100 μM of the ubiquitin cysteine mutant were incubated in HBS for several hours at room temperature with 1 mM Oregon Green maleimide (Invitrogen). The reaction was stopped by the addition of a tenfold molar excess of β-mercaptoethanol and unreacted dye was removed with a 15 ml G-25 column equilibrated in HBS.

6.3.3 Rpn10

$SpRpn10\text{-}VWA_{1\text{-}193}$ including the stop-codon was cloned into pET-21d to generate an untagged protein and transformed into *E. coli* BL21(DE3) cells for protein expression. Cells were thawed under running cold water, and after sonication the lysate was cleared by centrifugation for 45 minutes at 75 000× g. The soluble fraction was applied to a column packed with 15 ml hydroxyapatite (Calbiochem), washed with five column volumes HB (20 mM HEPES, pH 7.5), and then bound protein was eluted with a linear gradient of 0-0.4 M Na_2HPO_4 in the same buffer. Fractions containing

significant amounts of Rpn10-VWA_{1-193} were identified by SDS-PAGE, pooled, and concentrated to circa 8 ml using a 10 000 MWCO ultrafiltration device (Vivascience). The concentrated sample was applied to a K26 Sephadex 75 size-exclusion column (GE Healthcare), equilibrated with HBS (20 mM HEPES, 150 mM NaCl, 0.5 mM DTT, 0.02% NaN_3, pH 7.5) and Rpn10-VWA_{1-193}-containing fractions were subsequently pooled and concentrated to 8 ml. After dialysis overnight against HB at 4 °C, the sample was applied to a 1.6 ml Mono Q column equilibrated in HB, and eluted with a linear gradient of 0.0-1.0 M NaCl in HB. The fractions containing Rpn10-VWA_{1-193} were then applied to a K16 Sephadex 30 column previously equilibrated in HBS. For optimal crystallisation conditions the sample was concentrated to 18 mg/ml in 10 mM HEPES, 12 mM NaCl, pH 7.5.

6.3.4 Rpn12

A proteolytically stable construct was identified by subjecting the full-length protein to proteolysis by subtilisin A at 1:1000 and 1:10 000 molar ratios for 30 minutes at 37 °C. Stable fragments were identified by N-terminal sequencing and mass spectrometry and subsequently PCR amplified and sub-cloned into the GST-fusion plasmid pGEX-6P-1 as 1-228 and 1-250 fragments. Selenomethionine derivative protein was expressed in SelenoMethionine Medium Complete (Molecular Dimensions). An overnight culture was grown in LB, harvested by centrifugation at 4000× g for ten minutes and washed once with selenomethionine media. All cells were used to inoculate 4× 0.5 l Selenomethionine media. At an OD_{600} of 0.6, $Rpn12_{1-228}$ protein expression was induced with 200 μM IPTG. Unlabelled protein was expressed as described above. The purification was identical to that of all other GST-fusion proteins.

Fluorescently labelled Rpn12 was prepared equivalently to ubiquitin (Section 6.3.2).

6.3.5 Rhp23

All Rhp23 constructs were recombinantly expressed in *E. coli* BL21(DE3) cells growing in terrific broth media at 37 °C. At an OD_{600} of 0.6 the temperature was reduced to 20 °C and protein expression was induced overnight with 500 μM IPTG. The

purification procedure was identical to that of the other GST-fusion proteins described above.

6.4 Crystallisation, Data Collection and Structure Calculation

6.4.1 Rpn10 VWA Domain

Sitting crystallisation drops were set up in 96-well plates containing commercially available sparse matrix and systematic grid screens totalling 576 conditions (Wizard I + II (Emerald Biosystems), Structure Screens 1 + 2, Pact Premier, JCSG-Plus, Stura Footprint (Molecular Dimensions), Index, PEG/Ion, Natrix, NaCl Grid Sreen, $(NH_4)_2SO_4$ Grid Screen, Crystal Screen Lite (Hampton Research). 100 nl mother-liquor and 100 nl protein solution (Rpn10-VWA_{1-193} at 18 mg/ml in 10 mM HEPES, 12 mM NaCl, pH 7.5) were set up as sitting drops using the Mosquito robot (TTP LabTech). A native dataset was collected at the ESRF beamline ID14-2 from a crystal grown in 0.1 M Tris/HCl pH 8.5, 1.5 M $(NH_4)_2SO_4$, 25 mM $MgSO_4$, cryo-protected in the same buffer containing 25% glycerol. 180 images with an oscillation range of 1° were taken to a resolution of 1.27 Å. Datasets were integrated, and scaled using MOSFLM (Leslie 1992) and SCALA (Collaborative Computational Project 1994) respectively, and finally converted to SHELX format (hkl) with MTZ2VARIOUS (Collaborative Computational Project 1994). The sulphur single-wavelength anomalous dispersion dataset was collected in-house on the same crystal using a 2.7 kW Bruker MicroStar X-ray generator with Montel optics and a Bruker SMART 6000 CCD detector to a resolution of 2.1 Å. Various phi- and chi-scans were performed to yield an anomalous multiplicity of 21.57 over the complete resolution range (30-2.1 Å). Both datasets were prepared for further analysis with SHELXC (Sheldrick 2008) to find ten sulphur atoms using 1000 iterations. The positions of ten sulphur atoms were determined with SHELXD using data to a resolution of 2.3 Å. These ten sites were consistent with eight methionine and two cysteine residues in the structure. Phasing calculations and solvent flattening were performed in SHELXE using 400 iterations. A solvent content of 25.9% gave the best results. By applying the “free lunch algorithm” implemented in SHELXE the resolution was extended to 1.0 Å. Subsequently, SHELXE analysed the output map, searched for more sites, and checked the quality of the

input sites. This resulted in a total of 13 sites, ten of which had a peak height of around 15 sigma or greater and those of the next three peaks found were much smaller. These sites might represent sulphate ions present in the solvent. The pseudo-free correlation coefficient of 84.67% indicates the very high quality of the map. The "correct hand" was identified by visual inspection of the electron density map. At this point, an electron density map with single atoms resolved allowed model building in COOT (Emsley et al. 2004) for most of the structure. Some loop regions, however, displayed high disorder and could only be built after several alternating building-refinement cycles in COOT and PHENIX.REFINE (Paul D Adams et al. 2010), respectively. Due to a strongly increasing mean (I)/σ(mean) beyond this value, the resolution was cut to 1.3 Å. The final model comprises residues 2-147 and 151-193. 264 water molecules and three sulphate ions, supported by peaks in the difference electron density map, could be modelled into the density. It possesses a crystallographic R-factor of 12.28% and an R_{free} of 16.46%. Data collection and refinement statistics can be found in Table 2.1.

6.4.2 Rpn12

Sitting crystallisation drops containing equal volumes (100 nl) of protein solution and mother liquor were set up in 96-well plates containing a mixture of commercially available sparse matrix screens (Wizard I + II (Emerald Biosystems), Structure Screens 1 + 2, Pact Premier, JCSG-Plus (Molecular Dimensions), Index Screen (Hampton Research) totalling 480 conditions using the Mosquito robot (TTP LabTech). Initial hits were obtained in Bis-Tris Propane pH 6.5 and 7.5, 20% PEG 3350 and 0.2 M of NaI, NaBr, or $NaNO_3$. These conditions were optimised by varying pH and PEG concentrations and crystals grown in 17-23% PEG 3350, 150-250 mM $NaNO_3$, 100 mM Bis-Tris Propane pH 7.5 were cryo-protected in the same buffer containing 25% glycerol. As heavy metal soaking experiments were not successful selenomethionine derivative crystals were grown in the same conditions. Native and SAD datasets were collected at ESRF beamline ID 14-4. The data were integrated and scaled using MOSFLM (Leslie 1992) and SCALA (Collaborative Computational Project 1994), respectively and the SHELX (Sheldrick 2008) suite was used to obtain

experimental phases. Both datasets were prepared with SHELXC. Because of a drop in the anomalous signal to noise ratio to below 1.5 at 2.3 Å, SHELXD only used data to 2.4 Å to identify 14 anomalous sites with a correlation coefficient greater than 0.3. This was consistent with seven methionine residues per $Rpn12_{1-228}$ molecule. SHELXE used this substructure identified by SHELXD to calculate initial phases for all reflections and refined these by solvent-flattening and continuity restraints. The initial model was built into the electron-density using COOT (Emsley et al. 2004) followed by alternating cycles of refinement in PHENIX.REFINE (Paul D Adams et al. 2010) and manual building in COOT.

6.5 Isothermal Titration Calorimetry

All samples were prepared in 25 mM sodium phosphate pH 6.5, 50 mM sodium chloride. Concentrations were determined spectrophotometrically at 280 nm using the experimentally determined extinction coefficients of 3425 M^{-1} cm^{-1} (Ub_2) and 13118 M^{-1} cm^{-1} (Rpn10). All experiments were performed on an iTC200 microcalorimeter (MicroCal) at 25 °C. The titrations were started with an initial injection of 0.8 µl to minimise losses due to protein diffusing into the cell, a second 3.2 µl injection, followed by nine 4 µl injections. For the titration of monoubiquitin, the cell contained 603 µM Rpn10 and the syringe contained 6.9 mM monoubiquitin. For the titration with K48-Ub_2, the cell contained 126 µM Rpn10 and the syringe 1.26 mM K48-Ub_2, whereas for the K63-linked Ub_2 titration the cell contained 134 µM Rpn10 and the syringe 1.45 mM K63-Ub_2. For all experiments the first injection was removed and the heat of dilution integrals were subtracted prior to analysis with the instrument software.

6.6 Sedimentation Velocity Ultracentrifugation

Samples of K48-Ub_2, K63-Ub_2, and SpRpn10 were prepared in 25 mM sodium phosphate pH 6.5, 50 mM sodium chloride. Six concentration ratios of Ub_2:Rpn10 ranging from 1:5 to 4:1 were prepared while keeping the total protein concentration constant at 300 µM. Sedimentation velocity experiments were performed in a Beckmann Optima XL-1 analytical ultracentrifuge at 20 °C and 40 000 rpm in an An-

60 Ti rotor. 280 nm absorbance measurements were taken at ten-minutes intervals for a total of 500 minutes. c(s) analysis for each sample was performed and the weight-average sedimentation coefficient was calculated using SEDFIT (Schuck 2000).

6.7 Fluorescence Anisotropy

Each reaction contained 100 nM of fluorescent probe and 0-375 μM Rpn10 or 0-250 μM Rhp23 in a total volume of 20 μl. Fluorescence polarization was recorded in black low protein-binding 384-well plates (Corning) using the PHERAstar FS platereader (BMG Labtech) equipped with a fluorescein polarisation module. The gain of the parallel and perpendicular channel was adjusted against the zero-point of each titration to result in a starting anisotropy of 23 mP. The raw intensities were recorded with 200 flashes per point. The binding curves were analysed in Prism 5 (Graphpad) using a one-site binding model with nonspecific binding (NS), starting anisotropy level (BG), maximal binding signal (B_{max}), dissociation constant (K_d) and ligand depletion by binding to the fluorescent probe (P):

Equation 1. Single-Site binding with ligand depletion.

$$Y = \frac{B_{max} \times (P + x + K_d) - \sqrt{(P + x + K_d)^2 - (4P \times x)}}{2P} + NS \times x + BG$$

6.8 Surface Plasmon Resonance

All experiments were performed on a BIAcore T100 instrument at 25 °C with a flow rate of 10 μl/min. Anti-GST antibody was coupled to a CM5 chip as described by the manufacturer (GE Healthcare). This resulted in ca. 11 000 immobilised response units (RU) on each flow cell. 1 μM GST-Rpn12 and GST as negative control were captured during 60 s injections, followed by a 100 s stabilisation period. Subsequently, the analyte (full-length or ΔUIM Rpn10) was injected for 1200 s at increasing concentrations ranging from 1 nM to 50 μM. This was followed by a 1200 s dissociation period and two 60 s regeneration steps using the glycine pH 2.2 solution supplied by the manufacturer (GE Healthcare) and 0.05% SDS. After a

stabilization period of 300 s the next cycle was started. The data was analysed using the BIAcore T100 evaluation software (GE Healthcare) and corrected for the amount of captured GST-fusion protein in Excel (Microsoft). Equilibrium binding values were plotted against the concentration and fitted to a one-site equilibrium binding equation using Prism 5 (Graphpad).

6.9 Table of constructs

protein	residues	mutations	vector	tag (N-terminal)
Rpn10	1-243	wt	pGEX-6P-1	GST
	1-243	E14A, W15A, N18A	pENTR-3C	-
	1-243	E14A, W15A, N18A	pDEST-15	GST
	1-243	D31A, H34A, D68A, H77A	pENTR-3C	-
	1-243	D31A, H34A, D68A, H77A	pDEST-15	GST
	1-243	T64A, Q92A, H100A, E102A, K104A, R107A	pENTR-3C	-
	1-243	T64A, Q92A, H100A, E102A, K104A, R107A	pDEST-15	GST
	1-243	K123A, K127A, K130A, K133A, K134A	pENTR-3C	-
	1-243	K123A, K127A, K130A, K133A, K134A	pDEST-15	GST
	1-243	D164A, H167A, D180A	pENTR-3C	-
	1-243	D164A, H167A, D180A	pDEST-15	GST
	1-193	wt	pET-21d	-
	1-195	wt	pGEX-6P-1	GST
	196-243	wt	pGEX-6P-1	GST
	196-243	E235C	pGEX-6P-1	GST
Rpn12	1-270	wt	pGEX-6P-1	GST
	1-250	wt	pGEX-6P-1	GST
	1-250	K29A, Q76A	pGEX-6P-1	GST
	1-250	N106A, H112A, E116A	pGEX-6P-1	GST
	1-250	Y183A	pGEX-6P-1	GST
	1-228	wt	pGEX-6P-1	GST
	1-176	wt	pGEX-6P-1	GST
Ubiquitin	1-76	wt	pGEX-6P-1 mod.	-
	1-76	K48C	pGEX-6P-1 mod.	-
	1-76	K63C	pGEX-6P-1 mod.	-
	1-76	K48R, K63C	pGEX-6P-1 mod.	-
	1-77	77D	pGEX-6P-1 mod.	-
Rhp23	1-368	wt	pGEX-6P-1	GST
	1-368	M157A, L183A	pGEX-6P-1	GST
	1-368	L322A, F338A	pGEX-6P-1	GST
	1-368	M157A, L183A, L322A, F338A	pGEX-6P-1	GST
	146-190	wt	pGEX-6P-1	GST
	146-190	M157A, L183A	pGEX-6P-1	GST
	316-368	wt	pGEX-6P-1	GST
	316-368	L322A, F338A	pGEX-6P-1	GST

7 References

Adams, Paul D et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. 66, 213-21.

Alam, S. L., Langelier, C., Whitby, F. G., Koirala, S., Robinson, H., Hill, Christopher P, and Sundquist, W. I. (2006). Structural basis for ubiquitin recognition by the human ESCRT-II EAP45 GLUE domain. Nat. Struct. Mol. Biol. 13, 1029-30.

Beal, R., Deveraux, Q., Xia, G., Rechsteiner, M, and Pickart, C. M. (1996). Surface hydrophobic residues of multiubiquitin chains essential for proteolytic targeting. Proc Natl Acad Sci U S A 93, 861-6.

Bergink, S., and Jentsch, S. (2009). Principles of ubiquitin and SUMO modifications in DNA repair. Nature 458, 461-7.

Bertolaet, B. L., Clarke, D. J., Wolff, M., Watson, M. H., Henze, M., Divita, G., and Reed, S. I. (2001). UBA domains of DNA damage-inducible proteins interact with ubiquitin. Nat. Struct. Biol. 8, 417-22.

Bienkowska, J., Cruz, M., Atiemo, A., Handin, R., and Liddington, R. (1997). The von willebrand factor A3 domain does not contain a metal ion-dependent adhesion site motif. J. Biol. Chem. 272, 25162-7.

Borissenko, L., and Groll, M. (2007). 20S proteasome and its inhibitors: crystallographic knowledge for drug development. Chem. Rev. 107, 687-717.

Bremm, A., Freund, S. M. V., and Komander, D. (2010). Lys11-linked ubiquitin chains adopt compact conformations and are preferentially hydrolyzed by the deubiquitinase Cezanne. Nat. Struct. Mol. Biol. 17, 939-47.

Brzovic, P. S., Lissounov, A., Christensen, D. E., Hoyt, D. W., and Klevit, R. E. (2006). A UbcH5/ubiquitin noncovalent complex is required for processive BRCA1-directed ubiquitination. Mol. Cell 21, 873-80.

Brünger, A. T. et al. (1998). Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Crystallogr. D Biol. Crystallogr. 54, 905-21.

Cascio, P., Call, M., Petre, B. M., Walz, T., and Goldberg, Alfred L (2002). Properties of the hybrid form of the 26S proteasome containing both 19S and PA28 complexes. EMBO J. 21, 2636-45.

Chastagner, P., Israël, A., and Brou, C. (2006). Itch/AIP4 mediates Deltex degradation through the formation of K29-linked polyubiquitin chains. EMBO Rep. 7, 1147-53.

Chau, V., Tobias, J. W., Bachmair, A., Marriott, D., Ecker, D. J., Gonda, D. K., and Varshavsky, A (1989). A multiubiquitin chain is confined to specific lysine in a targeted short-lived protein. Science 243, 1576-83.

Chen, L., and Madura, K. (2002). Rad23 Promotes the Targeting of Proteolytic Substrates to the Proteasome. Mol. Cell. Biol. 22, 4902-4913.

Chen, P., and Hochstrasser, M. (1996). Active Site Formation in the 20S Proteasome to Completion of Assembly. Cell 86, 961-972.

Chen, X., Lee, B.-H., Finley, D., and Walters, K. J. (2010). Structure of proteasome ubiquitin receptor hRpn13 and its activation by the scaffolding protein hRpn2. Mol. Cell 38, 404-15.

Ciechanover, A. (1994). The ubiquitin-proteasome proteolytic pathway. Cell 79, 13-21.

Ciechanover, A., Elias, S., Heller, H., Ferber, S., and Hershko, A. (1980). Characterization of the heat-stable polypeptide of the ATP-dependent proteolytic system from reticulocytes. J. Biol. Chem. 255, 7525-8.

Ciechanover, A., Heller, H., Elias, S., Haas, A. L., and Hershko, A. (1980). ATP-dependent conjugation of reticulocyte proteins with the polypeptide required for protein degradation. Proc Natl Acad Sci U S A 77, 1365-8.

Ciechanover, A., Heller, H., Katz-Etzion, R., and Hershko, A. (1981). Activation of the heat-stable polypeptide of the ATP-dependent proteolytic system. Proc Natl Acad Sci U S A 78, 761-5.

Ciechanover, A., Hod, Y., and Hershko, A. (1978). A heat-stable polypeptide component of an ATP-dependent proteolytic system from reticulocytes. Biochem. Biophys. Res. Commun. 81, 1100-1105.

Collaborative Computational Project, N. 4 (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. 50, 760-3.

Cook, W. J., Jeffrey, L C, Carson, M., Chen, Z., and Pickart, C. M. (1992). Structure of a diubiquitin conjugate and a model for interaction with ubiquitin conjugating enzyme (E2). J. Biol. Chem. 267, 16467-71.

Cook, W. J., Jeffrey, Leigh C, Kasperek, E., and Pickart, C. M. (1994). Structure of Tetraubiquitin Shows How Multiubiquitin Chains Can Be Formed. J. Mol. Biol. 236, 601-609.

Crosas, B. et al. (2006). Ubiquitin chains are remodeled at the proteasome by opposing ubiquitin ligase and deubiquitinating activities. Cell 127, 1401-13.

Dam, J., Velikovsky, C. A., Mariuzza, R. A., Urbanke, C., and Schuck, P. (2005). Sedimentation velocity analysis of heterogeneous protein-protein interactions: Lamm equation modeling and sedimentation coefficient distributions c(s). Biophys. J. 89, 619-34.

Dantuma, N. P., Heinen, C., and Hoogstraten, D. (2009). The ubiquitin receptor Rad23: at the crossroads of nucleotide excision repair and proteasomal degradation. DNA Repair 8, 449-60.

Datta, A. B., Hura, G. L., and Wolberger, C. (2009). The structure and conformation of Lys63-linked tetraubiquitin. J. Mol. Biol. 392, 1117-24.

Dessau, M., Halimi, Y., Erez, T., Chomsky-Hecht, O., Chamovitz, D. A., and Hirsch, J. A. (2008). The Arabidopsis COP9 signalosome subunit 7 is a model PCI domain protein with subdomains involved in COP9 signalosome assembly. Plant Cell 20, 2815-34.

Deveraux, Q., Nocker, S. van, Mahaffey, D., Vierstra, R. D., and Rechsteiner, Martin (1995). Inhibition of ubiquitin-mediated proteolysis by the Arabidopsis 26 S protease subunit S5a. J. Biol. Chem. 270, 29660-3.

Deveraux, Q., Ustrell, V., Pickart, C. M., and Rechsteiner, Martin (1994). A 26 S protease subunit that binds ubiquitin conjugates. J. Biol. Chem. 269, 7059-61.

Dieckmann, T., Withers-Ward, E. S., Jarosinski, M. a, Liu, C. F., Chen, I. S., and Feigon, J (1998). Structure of a human DNA repair protein UBA domain that interacts with HIV-1 Vpr. Nat. Struct. Biol. 5, 1042-7.

Dikic, I., Wakatsuki, S., and Walters, K. J. (2009). Ubiquitin-binding domains - from structures to functions. Nat. Rev. Mol. Cell Biol. 10, 659-71.

Dominguez, C., Boelens, R., and Bonvin, A. M. J. J. (2003). HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. J. Am. Chem. Soc. 125, 1731-7.

D'Andrea, L., and Regan, L. (2003). TPR proteins: the versatile helix. Trends Biochem. Sci. 28, 655-662.

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

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. 60, 2126-32.

Enchev, R. I., Schreiber, A., Beuron, F., and Morris, E. P. (2010). Structural insights into the COP9 signalosome and its common architecture with the 26S proteasome lid and eIF3. Structure 18, 518-27.

Etlinger, J. D., and Goldberg, Alfred L (1977). A soluble ATP-dependent proteolytic system responsible for the degradation of abnormal proteins in reticulocytes. Proc Natl Acad Sci U S A 74, 54-58.

Finley, D. (2009). Recognition and processing of ubiquitin-protein conjugates by the proteasome. Annu. Rev. Biochem. 78, 477-513.

Fu, H., Reis, N., Lee, Y., and Glickman, M. H. (2001). Subunit interaction maps for the regulatory particle of the 26S proteasome and the COP9 signalosome. EMBO J. 20, 7096-7107.

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

Fushman, D., and Walker, O. (2010). Exploring the linkage dependence of polyubiquitin conformations using molecular modeling. J. Mol. Biol. 395, 803-14.

Gaczynska, M, Rock, K. L., and Goldberg, A L (1993). Gamma-interferon and expression of MHC genes regulate peptide hydrolysis by proteasomes. Nature 365, 264-7.

Glickman, M. H., Rubin, D. M., Coux, O., Wefes, I., Pfeifer, G., Cjeka, Z., Baumeister, W., Fried, Victor A, and Finley, D. (1998). A subcomplex of the proteasome

regulatory particle required for ubiquitin-conjugate degradation and related to the COP9-signalosome and eIF3. Cell 94, 615-624.

Glickman, M. H., Rubin, D. M., Fried, V A, and Finley, D. (1998). The regulatory particle of the Saccharomyces cerevisiae proteasome. Mol. Cell. Biol. 18, 3149-62.

Groll, M., Bajorek, M., Köhler, A., Moroder, L., Rubin, D. M., Huber, R., Glickman, M. H., and Finley, D. (2000). A gated channel into the proteasome core particle. Nat. Struct. Biol. 7, 1062-7.

Guterman, A., and Glickman, M. H. (2004). Complementary roles for Rpn11 and Ubp6 in deubiquitination and proteolysis by the proteasome. J. Biol. Chem. 279, 1729-38.

Haglund, K (2003). Distinct monoubiquitin signals in receptor endocytosis. Trends Biochem. Sci. 28, 598-604.

Haglund, Kaisa, and Dikic, I. (2005). Ubiquitylation and cell signaling. EMBO J. 24, 3353-9.

Hamazaki, J., Sasaki, K., Kawahara, H., Hisanaga, S.-I., Tanaka, Keiji, and Murata, S. (2007). Rpn10-mediated degradation of ubiquitinated proteins is essential for mouse development. Mol. Cell. Biol. 27, 6629-38.

Haracska, L., and Udvardy, A. (1995). Cloning and sequencing a non-ATPase subunit of the regulatory complex of the Drosophila 26S protease. Eur. J. Biochem. 231, 720-5.

Haracska, L., and Udvardy, A. (1997). Mapping the ubiquitin-binding domains in the p54 regulatory complex subunit of the Drosophila 26S protease. FEBS Lett. 412, 331-6.

Haririnia, A., D'Onofrio, M., and Fushman, D. (2007). Mapping the interactions between Lys48 and Lys63-linked di-ubiquitins and a ubiquitin-interacting motif of S5a. J. Mol. Biol. 368, 753-66.

Hartmann-Petersen, R., and Gordon, C. (2004). Protein degradation: recognition of ubiquitinylated substrates. Curr. Biol. 14, R754-6.

Hayden, M. S., and Ghosh, S. (2008). Shared principles in NF-kappaB signaling. Cell 132, 344-62.

Heessen, S., Masucci, M. G., and Dantuma, N. P. (2005). The UBA2 domain functions as an intrinsic stabilization signal that protects Rad23 from proteasomal degradation. Mol. Cell 18, 225-35.

Hershko, A., Ciechanover, A., Heller, H., Haas, A. L., and Rose, I. A. (1980). Proposed Role of ATP in Protein Breakdown: Conjugation of Proteins with Multiple Chains of the Polypeptide of ATP-Dependent Proteolysis. Proc Natl Acad Sci U S A 77, 1783-1786.

Hershko, A., Ciechanover, A., and Rose, I. A. (1979). Resolution of the ATP-dependent proteolytic system from reticulocytes: a component that interacts with ATP. Proc Natl Acad Sci U S A 76, 3107-10.

Hershko, A., Heller, H., Elias, S., and Ciechanover, A. (1983). Components of ubiquitin-protein ligase system. Resolution, affinity purification, and role in protein breakdown. J. Biol. Chem. 258, 8206-14.

Hicke, L., Schubert, H. L., and Hill, Christopher P (2005). Ubiquitin-binding domains. Nat. Rev. Mol. Cell Biol. 6, 610-21.

Hirano, S., Masato, K., Hideaki, U., Kato, K., Raiborg, C., Stenmark, H., and Wakatsuki, S. (2006a). Double-sided ubiquitin binding of Hrs-UIM in endosomal protein sorting. Nat. Struct. Mol. Biol. 13, 272-277.

Hirano, S., Suzuki, N., Slagsvold, T., Kawasaki, M., Trambaiolo, D., Kato, R., Stenmark, H., and Wakatsuki, S. (2006b). Structural basis of ubiquitin recognition by mammalian Eap45 GLUE domain. Nat. Struct. Mol. Biol. 13, 1031-2.

Hough, R., Pratt, G., and Rechsteiner, Martin (1986). Ubiquitin-lysozyme conjugates. Identification and characterization of an ATP-dependent protease from rabbit reticulocyte lysates. J. Biol. Chem. 261, 2400-8.

Huizinga, E. G., Tsuji, S., Romijn, R. A. P., Schiphorst, M. E., Groot, P. G. de, Sixma, J. J., and Gros, Piet (2002). Structures of glycoprotein Ibalpha and its complex with von Willebrand factor A1 domain. Science 297, 1176-9.

Husnjak, K. et al. (2008). Proteasome subunit Rpn13 is a novel ubiquitin receptor. Nature 453, 481-488.

Ikeda, F., and Dikic, I. (2008). Atypical ubiquitin chains: new molecular signals. “Protein Modifications: Beyond the Usual Suspects” review series. EMBO Rep. 9, 536-42.

Inoue, H., Nojima, H., and Okayama, H. (1990). High efficiency transformation of Escherichia coli with plasmids. Gene 96, 23-8.

Isasa, M. et al. (2010). Monoubiquitination of RPN10 Regulates Substrate Recruitment to the Proteasome. Mol. Cell 38, 733-745.

Iwanczyk, J., Sadre-Bazzaz, K., Ferrell, K., Kondrashkina, E., Formosa, T., Hill, Christopher P, and Ortega, J. (2006). Structure of the Blm10-20 S proteasome complex by cryo-electron microscopy. Insights into the mechanism of activation of mature yeast proteasomes. J. Mol. Biol. 363, 648-59.

Janse, D. M., Crosas, B., Finley, D., and Church, G. M. (2004). Localization to the proteasome is sufficient for degradation. J. Biol. Chem. 279, 21415-20.

Jin, L., Williamson, A., Banerjee, S., Philipp, I., and Rape, M. (2008). Mechanism of ubiquitin-chain formation by the human anaphase-promoting complex. Cell 133, 653-65.

Johnson, E. S., Ma, P. C., Ota, I. M., and Varshavsky, A (1995). A proteolytic pathway that recognizes ubiquitin as a degradation signal. J. Biol. Chem. 270, 17442-56.

Kajava, A. V. (2002). What curves alpha-solenoids? Evidence for an alpha-helical toroid structure of Rpn1 and Rpn2 proteins of the 26 S proteasome. J. Biol. Chem. 277, 49791-8.

Kang, Y., Chen, X., Lary, J. W., Cole, J. L., and Walters, K. J. (2007). Defining how ubiquitin receptors hHR23a and S5a bind polyubiquitin. J. Mol. Biol. 369, 168-176.

Kaplun, L., Tzirkin, R., Bakhrat, A., Shabek, N., Ivantsiv, Y., and Raveh, D. (2005). The DNA damage-inducible UbL-UbA protein Ddi1 participates in Mec1-mediated degradation of Ho endonuclease. Mol. Cell. Biol. 25, 5355-62.

Kim, H. T., Kim, K. P., Lledias, F., Kisselev, A. F., Scaglione, K. M., Skowyra, D., Gygi, S. P., and Goldberg, Alfred L (2007). 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, 17375-86.

Kim, I., Mi, K., and Rao, H. (2004). Multiple interactions of Rad23 suggest a mechanism for ubiquitylated substrate delivery important in proteolysis. Mol. Biol. Cell 15, 3357-3365.

Kirkpatrick, D. S., Hathaway, N. A., Hanna, J., Elsasser, S., Rush, J., Finley, D., King, R. W., and Gygi, S. P. (2006). Quantitative analysis of in vitro ubiquitinated cyclin B1 reveals complex chain topology. Nat. Cell Biol. 8, 700-710.

Kleiger, G., Saha, A., Lewis, S., Kuhlman, B., and Deshaies, R. J. (2009). Rapid E2-E3 assembly and disassembly enable processive ubiquitylation of cullin-RING ubiquitin ligase substrates. Cell 139, 957-68.

Kodadek, T. (2010). No Splicing, no dicing: non-proteolytic roles of the ubiquitin-proteasome system in transcription. J. Biol. Chem. 285, 2221-6.

Komander, D., Reyes-Turcu, F. E., Licchesi, J. D. F., Odenwaelder, P., Wilkinson, K. D., and Barford, D. (2009). Molecular discrimination of structurally equivalent Lys 63-linked and linear polyubiquitin chains. EMBO Rep. 10, 466-73.

Krissinel, Evgeny, and Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. J. Biol. Chem. 372, 774-97.

Köhler, A., Cascio, P., Leggett, D. S., Woo, K. M., Goldberg, Alfred L, and Finley, D. (2001). The axial channel of the proteasome core particle is gated by the Rpt2 ATPase and controls both substrate entry and product release. Mol. Cell 7, 1143-52.

Lambert, C., Léonard, N., De Bolle, X., and Depiereux, E. (2002). ESyPred3D: Prediction of proteins 3D structures. Bioinformatics 18, 1250-6.

Lambertson, D., Chen, L., and Madura, K. (2003). Investigating the importance of proteasome-interaction for Rad23 function. Curr. Genet. 42, 199-208.

Lambertson, D., Chen, L., and Madura, K. (1999). Pleiotropic defects caused by loss of the proteasome-interacting factors Rad23 and Rpn10 of Saccharomyces cerevisiae. Genetics 153, 69-79.

Landau, M., Mayrose, I., Rosenberg, Y., Glaser, F., Martz, E., Pupko, T., and Ben-Tal, N. (2005). ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. Nucleic Acids Res. 33, W299-302.

Lapouge, K., Smith, S. J. M., Walker, P. a, Gamblin, S. J., Smerdon, S. J., and Rittinger, K. (2000). Structure of the TPR Domain of p67phox in Complex with Rac·GTP. Mol. Cell 6, 899-907.

Larkin, M. A. et al. (2007). Clustal W and Clustal X version 2.0. Bioinformatics 23, 2947-8.

Lee, S., Tsai, Y. C., Mattera, R., Smith, W. J., Kostelansky, M. S., Weissman, A. M., Bonifacino, J. S., and Hurley, J. H. (2006). Structural basis for ubiquitin recognition and autoubiquitination by Rabex-5. Nat. Struct. Mol. Biol. 13, 264-71.

Leslie, A. G. W. (1992). Recent changes to the MOSFLM package for processing film and image plate data. Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography.

Li, X., Kusmierczyk, A. R., Wong, P., Emili, A., and Hochstrasser, M. (2007). beta-Subunit appendages promote 20S proteasome assembly by overcoming an Ump1-dependent checkpoint. EMBO J. 26, 2339-49.

Löwe, J., Stock, D., Jap, B., Zwickl, P., Baumeister, W., and Huber, R. (1995). Crystal Structure of the 20S Proteasome from the Archaeon T. acidophilium at 3.4 Å Resolution. Science 268, 533-539.

Martin-Serrano, J. (2007). The role of ubiquitin in retroviral egress. Traffic 8, 1297-303.

Masson, P., Andersson, O., Petersen, U. M., and Young, P (2001). Identification and characterization of a Drosophila nuclear proteasome regulator. A homolog of human 11 S REGgamma (PA28gamma). J. Biol. Chem. 276, 1383-90.

Matiuhin, Y., Kirkpatrick, D. S., Ziv, I., Kim, W., Dakshinamurthy, A., Kleifeld, O., Gygi, S. P., Reis, N., and Glickman, M. H. (2008). Extraproteasomal Rpn10 restricts access of the polyubiquitin-binding protein Dsk2 to proteasome. Mol. Cell 32, 415-25.

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

Mueller, T. D., and Feigon, Juli (2002). Solution Structures of UBA Domains Reveal a Conserved Hydrophobic Surface for Protein–Protein Interactions. J. Mol. Biol. 319, 1243-1255.

Murata, S., Yashiroda, H., and Tanaka, Keiji (2009). Molecular mechanisms of proteasome assembly. Nat. Rev. Mol. Cell Biol. 10, 104-15.

Neely, A. N., and Mortimore, G. E. (1974). Localization of products of endogenous proteolysis in lysosomes of perfused rat liver. Biochem. Biophys. Res. Commun. 59, 680-7.

Ng, J. M. Y., Vermeulen, W., Der Horst, G. T. J. van, Bergink, S., Sugasawa, K., Vrieling, H., and Hoeijmakers, J. H. J. (2003). A novel regulation mechanism of DNA repair by damage-induced and RAD23-dependent stabilization of xeroderma pigmentosum group C protein. Genes Dev. 17, 1630-45.

Nickell, S. et al. (2009). Insights into the molecular architecture of the 26S proteasome. Proc Natl Acad Sci U S A 106, 11943-7.

Nishikawa, H., Ooka, S., Sato, K., Arima, K., Okamoto, J., Klevit, R. E., Fukuda, M., and Ohta, T. (2004). Mass spectrometric and mutational analyses reveal Lys-6-linked polyubiquitin chains catalyzed by BRCA1-BARD1 ubiquitin ligase. J. Biol. Chem. 279, 3916-24.

Nisogi, H., Kominami, K., Tanaka, K, and Toh-e, A. (1992). A new essential gene of Saccharomyces cerevisiae, a defect in it may result in instability of nucleus. Exp. Cell. Res. 200, 48-57.

Nocker, S. van, Sadis, S., Rubin, D. M., Glickman, M. H., Fu, H., Coux, O., Wefes, I., Finley, D., and Vierstra, R. D. (1996). The multiubiquitin-chain-binding protein Mcb1 is a component of the 26S proteasome in Saccharomyces cerevisiae and plays a nonessential, substrate-specific role in protein turnover. Mol. Cell. Biol. 16, 6020-8.

Ortolan, T. G., Chen, L., Tongaonkar, P., and Madura, K. (2004). Rad23 stabilizes Rad4 from degradation by the Ub/proteasome pathway. Nucleic Acids Res. 32, 6490-500.

Ozkaynak, E., Finley, D., and Varshavsky, A (1984). The yeast ubiquitin gene: head-to-tail repeats encoding a polyubiquitin precursor protein. Nature 312, 663-6.

Phillips, C. L., Thrower, J. S., Pickart, C. M., and Hill, Christopher P (2001). Structure of a new crystal form of tetraubiquitin. Acta Crystallogr. D Biol. Crystallogr. 57, 341-4.

Pick, E., Hofmann, K., and Glickman, M. H. (2009). PCI complexes: Beyond the proteasome, CSN, and eIF3 Troika. Mol. Cell 35, 260-4.

Pickart, C. M. (2001). Mechanisms underlying ubiquitination. Annua. Rev. Biochem. 70, 503-33.

Pickart, C. M., and Raasi, S. (2005). Controlled synthesis of polyubiquitin chains. Methods Enzymol. 399, 21-36.

Potterton, L. et al. (2004). Developments in the CCP4 molecular-graphics project. Acta Crystallogr. D Biol. Crystallogr. 60, 2288-94.

Prakash, S., Tian, L., Ratliff, K. S., Lehotzky, R. E., and Matouschek, A. (2004). An unstructured initiation site is required for efficient proteasome-mediated degradation. Nat. Struct. Mol. Biol. 11, 830-7.

Raasi, S., and Pickart, C. M. (2003). Rad23 Ubiquitin-associated Domains (UBA) Inhibit 26 S Proteasome-catalyzed Proteolysis by Sequestering Lysine 48-linked Polyubiquitin Chains. J. Biol. Chem. 278, 8951-8959.

Raasi, S., Varadan, R., Fushman, D., and Pickart, C. M. (2005). Diverse polyubiquitin interaction properties of ubiquitin-associated domains. Nat. Struct. Mol. Biol. 12, 708-714.

Rao, H., and Sastry, A. (2002). Recognition of specific ubiquitin conjugates is important for the proteolytic functions of the ubiquitin-associated domain proteins Dsk2 and Rad23. J. Biol. Chem. 277, 11691-11695.

Reyes-Turcu, F. E., Horton, J. R., Mullally, J. E., Heroux, A., Cheng, X., and Wilkinson, K. D. (2006). The ubiquitin binding domain ZnF UBP recognizes the C-terminal diglycine motif of unanchored ubiquitin. Cell 124, 1197-208.

Reyes-Turcu, F. E., Ventii, K. H., and Wilkinson, K. D. (2009). Regulation and cellular roles of ubiquitin-specific deubiquitinating enzymes. Annu. Rev. Biochem. 78, 363-97.

Richly, H., Rape, M., Braun, S., Rumpf, S., Hoege, C., and Jentsch, S. (2005). A series of ubiquitin binding factors connects CDC48/p97 to substrate multiubiquitylation and proteasomal targeting. Cell 120, 73-84.

Risler, J. L., Delorme, M. O., Delacroix, H., and Henaut, A. (1988). Amino acid substitutions in structurally related proteins. A pattern recognition approach. Determination of a new and efficient scoring matrix. J. Mol. Biol. 204, 1019-29.

Rosenzweig, R., Osmulski, P. A., Gaczynska, Maria, and Glickman, M. H. (2008). The central unit within the 19S regulatory particle of the proteasome. Nat. Struct. Mol. Biol. 15, 573-80.

Rothstein, R. J., and Sherman, F (1980). Genes affecting the expression of cytochrome c in yeast: genetic mapping and genetic interactions. Genetics 94, 871-89.

Ryu, K.-S., Lee, K.-J., Bae, S.-H., Kim, B.-K., Kim, K.-A., and Choi, B.-S. (2003). Binding surface mapping of intra- and interdomain interactions among hHR23B, ubiquitin, and polyubiquitin binding site 2 of S5a. J. Biol. Chem. 278, 36621-36627.

Saftig, P., and Klumperman, J. (2009). Lysosome biogenesis and lysosomal membrane proteins: trafficking meets function. Nat. Rev. Mol. Cell Biol. 10, 623-35.

Santelli, E., Bankston, L. a, Leppla, S. H., and Liddington, R. C. (2004). Crystal structure of a complex between anthrax toxin and its host cell receptor. Nature 430, 905-8.

Sato, Y., Yoshikawa, A., Mimura, H., and Fukai, S. (2009). Structural basis for specific recognition of Lys 63-linked polyubiquitin chains by tandem UIMs of RAP80. EMBO J. 28, 2461-2468.

Sato, Y., Yoshikawa, A., Yamagata, A., Mimura, H., Yamashita, M., Ookata, K., Nureki, O., Komada, M., and Fukai, S. (2008). Structural basis for specific cleavage of Lys 63-linked polyubiquitin chains. Nature 455, 358-362.

Scheel, H., and Hofmann, K. (2005). Prediction of a common structural scaffold for proteasome lid, COP9-signalosome and eIF3 complexes. BMC Bioinformatics 6, 71.

Scheufler, C., Brinker, a, Bourenkov, G., Pegoraro, S., Moroder, L., Bartunik, H., Hartl, F. U., and Moarefi, I. (2000). Structure of TPR domain-peptide complexes: critical elements in the assembly of the Hsp70-Hsp90 multichaperone machine. Cell 101, 199-210.

Schlesinger, D. H., Goldstein, G., and Niall, H. D. (1975). The complete amino acid sequence of ubiquitin, an adenylate cyclase stimulating polypeptide probably universal in living cells. Biochemistry 14, 2214-8.

Schmidt, M., Haas, W., Crosas, B., Santamaria, P. G., Gygi, S. P., Walz, T., and Finley, D. (2005). The HEAT repeat protein Blm10 regulates the yeast proteasome by capping the core particle. Nat. Struct. Mol. Biol. 12, 294-303.

Schreiner, P. et al. (2008). Ubiquitin docking at the proteasome through a novel pleckstrin-homology domain interaction. Nature 453, 548-552.

Schuck, P. (2000). Size-Distribution Analysis of Macromolecules by Sedimentation Velocity Ultracentrifugation and Lamm Equation Modeling. Biophys. J. 78, 1606-1619.

Schulman, B. a, and Harper, J. W. (2009). Ubiquitin-like protein activation by E1 enzymes: the apex for downstream signalling pathways. Nat. Rev. Mol. Cell Biol. 10, 319-31.

Seeger, M., Hartmann-Petersen, R., Wilkinson, C. R. M., Wallace, M., Samejima, I., Taylor, M. S., and Gordon, C. (2003). Interaction of the anaphase-promoting complex/cyclosome and proteasome protein complexes with multiubiquitin chain-binding proteins. J. Biol. Chem. 278, 16791-16796.

Sharon, M., Mao, H., Erba, E. B., Stephens, E., Zheng, N., and Robinson, C. V. (2009). Symmetrical Modularity of the COP9 Signalosome Complex Suggests its Multifunctionality. Structure, 31-40.

Sharon, M., Taverner, T., Ambroggio, X. I., Deshaies, R. J., and Robinson, C. V. (2006). Structural organization of the 19S proteasome lid: insights from MS of intact complexes. PLoS Biol. 4, 1314-1323.

Sheldrick, G. M. (2008). A short history of SHELX. Acta Crystallogr. A 64, 112-22.

Shibatani, T., Carlson, E. J., Larabee, F., Mccormack, A. L., and Skach, W. R. (2006). Global Organization and Function of Mammalian Cytosolic Proteasome Pools : Implications for PA28 and 19S Regulatory Complexes. Mol. Biol. Cell 17, 4962-4971.

Simpson, M. V. (1953). The release of labeled amino acids from the proteins of rat liver slices. J. Biol. Chem. 201, 143-154.

Sims, J. J., and Cohen, R. E. (2009). Linkage-specific avidity defines the lysine 63-linked polyubiquitin-binding preference of rap80. Mol. Cell 33, 775-83.

Sims, J. J., Haririnia, A., Dickinson, B. C., Fushman, D., and Cohen, R. E. (2009). Avid interactions underlie the Lys63-linked polyubiquitin binding specificities observed for UBA domains. Nat. Struct. Mol. Biol. 16, 883-9.

Springer, T. A. (2006). Complement and the multifaceted functions of VWA and integrin I domains. Structure 14, 1611-6.

Stone, M., Hartmann-Petersen, R., Seeger, M., Bech-Otschir, D., Wallace, M., and Gordon, C. (2004). Uch2/Uch37 is the major deubiquitinating enzyme associated with the 26S proteasome in fission yeast. J. Mol. Biol. 344, 697-706.

Studier, F. W. (2005). Protein production by auto-induction in high density shaking cultures. Protein Expr. Purif. 41, 207-34.

Swanson, K. a, Kang, R. S., Stamenova, S. D., Hicke, L., and Radhakrishnan, I. (2003). Solution structure of Vps27 UIM-ubiquitin complex important for endosomal sorting and receptor downregulation. EMBO J. 22, 4597-606.

Tanahashi, N., Murakami, Y., Minami, Y., Shimbara, N., Hendil, K. B., and Tanaka, K (2000). Hybrid proteasomes. Induction by interferon-gamma and contribution to ATP-dependent proteolysis. J. Biol. Chem. 275, 14336-45.

Tenno, T. et al. (2004). Structural basis for distinct roles of Lys63- and Lys48-linked polyubiquitin chains. Genes Cells 9, 865-75.

Thrower, J. S., Hoffman, L., Rechsteiner, Martin, and Pickart, C. M. (2000). Recognition of the polyubiquitin proteolytic signal. EMBO J. 19, 94-102.

Varadan, R., Assfalg, M., Haririnia, A., Raasi, S., Pickart, C. M., and Fushman, D. (2004). Solution conformation of Lys63-linked di-ubiquitin chain provides clues to functional diversity of polyubiquitin signaling. J. Biol. Chem. 8, 7055-7063.

Varadan, R., Walker, O., Pickart, C. M., and Fushman, D. (2002). Structural properties of polyubiquitin chains in solution. J. Mol. Biol 324, 637-47.

Varshavsky, Alexander (2005). Regulated protein degradation. Trends Biochem. Sci. 30, 283-6.

Verma, R., Aravind, L., Oania, R., McDonald, W. H., Yates, J. R., Koonin, E. V., and Deshaies, R. J. (2002). Role of Rpn11 metalloprotease in deubiquitination and degradation by the 26S proteasome. Science 298, 611-5.

Verma, R., Chen, S., Feldman, R., Schieltz, D., Yates, J. R., Dohmen, J., and Deshaies, R. J. (2000). Proteasomal proteomics: identification of nucleotide-sensitive proteasome-interacting proteins by mass spectrometric analysis of affinity-purified proteasomes. Mol. Biol. Cell 11, 3425-3439.

Vijay-Kumar, S., Bugg, C. E., and Cook, W. J. (1987). Structure of ubiquitin refined at 1.8 Å resolution. J. Mol. Biol. 194, 531-44.

Walters, K. J., and Chen, X. (2009). Measuring ubiquitin chain linkage: Rap80 uses a molecular ruler mechanism for ubiquitin linkage specificity. EMBO J. 28, 2307-8.

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

Wang, M., Cheng, D., Peng, J., and Pickart, C. M. (2006). Molecular determinants of polyubiquitin linkage selection by an HECT ubiquitin ligase. EMBO J. 25, 1710-9.

Wang, Q., Young, Patrick, and Walters, K. J. (2005). Structure of S5a bound to monoubiquitin provides a model for polyubiquitin recognition. J. Mol. Biol. 348, 727-739.

Wei, Z., Zhang, P., Zhou, Z., Cheng, Z., Wan, M., and Gong, W. (2004). Crystal structure of human eIF3k, the first structure of eIF3 subunits. J. Biol. Chem. 279, 34983-90.

Whitby, F. G., Masters, E. I., Kramer, L., Knowlton, J. R., Yao, Y., Wang, C. C., and Hill, C P (2000). Structural basis for the activation of 20S proteasomes by 11S regulators. Nature 408, 115-20.

Whittaker, C. A., and Hynes, R. O. (2002). Distribution and evolution of von Willebrand/integrin A domains: widely dispersed domains with roles in cell adhesion and elsewhere. Mol. Biol. Cell 13, 3369-3387.

Wilkinson, C. R. M., Ferrell, K., Penney, M., Wallace, M., Dubiel, W., and Gordon, C. (2000). Analysis of a gene encoding Rpn10 of the fission yeast proteasome reveals that the polyubiquitin-binding site of this subunit is essential when Rpn12/Mts3 activity is compromised. J. Biol. Chem. 275, 15182-15192.

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

Yao, T., and Cohen, R. E. (2002). A cryptic protease couples deubiquitination and degradation by the proteasome. Nature 419, 403-407.

Ye, Y., and Rape, M. (2009). Building ubiquitin chains: E2 enzymes at work. Nat. Rev. Mol. Cell Biol. 10, 755-64.

Zhang, D., Chen, T., Ziv, I., Rosenzweig, R., Matiuhin, Y., Bronner, V., Glickman, M. H., and Fushman, D. (2009). Together, Rpn10 and Dsk2 can serve as a polyubiquitin chain-length sensor. Mol. Cell 36, 1018-33.

Zhang, N. et al. (2009). Structure of the s5a:k48-linked diubiquitin complex and its interactions with rpn13. Mol. Cell 35, 280-90.