The proteasome and its ancillary proteins

Jonida Tafilaku

Laboratory of Molecular Biology
University of Cambridge

This dissertation is submitted for the degree of
Doctor of Philosophy

Darwin College

April 2020

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other University. This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration, except where specifically indicated in the text. This dissertation contains less than 60,000 words excluding appendices, bibliography and tables.

Jonida Tafilaku

April 2020

Abstract

The proteasome and its ancillary proteins

Jonida Tafilaku

The proteasome is an essential multi-protein complex found within all eukaryotes. It degrades unwanted proteins and is involved in regulating many cellular functions, including cell homeostasis and apoptosis. Previous research in the field mainly focused on the mechanisms by which substrate proteins are specifically tagged for proteasome-mediated degradation and the subsequent mechanisms involved in substrate processing. However, the proteasome is known to interact with a wide range of ancillary proteins, and not all of these interactions have been characterised. The main focus of this thesis was therefore to identify and investigate the human proteasome interactions at a molecular level, with a focus on three specific proteins.

Firstly, I investigated the proteasome interaction with the UBR4 E3 ligase. I optimised the preparation of 26S proteasomes and found UBR4 co-purifying. I also showed that endogenously-tagged UBR4 conversely co-purifies with proteasome complexes. My biochemical and biophysical experiments indicate a strong interaction, as it was not possible to separate UBR4 from the proteasome *in vitro*. Increasing the yield of obtainable UBR4 to perform more detailed structural and functional analysis will be one of the future challenges.

Secondly, I studied the proteasome interaction with the P97 AAA+ ATPase. Using the baculovirus mediated insect cell overexpression system, I recombinantly co-expressed the human P97 and the 20S proteolytic core of the proteasome. Although P97 was shown to co-purify with affinity tagged 20S proteasomes, the data obtained suggest that any direct interaction appears to be of transient nature or may depend on other still unidentified proteins or cofactors.

Finally, I focused on PI31, a poorly characterised protein that has been described as a proteasome inhibitor. Indeed, by adding purified PI31 to 20S proteasomes I showed that in the presence of PI31 the proteasome peptidase activities are decreased. Interestingly, cryo-EM analysis of affinity-tagged 20S proteasomes recombinantly co-expressed with PI31

showed the 20S proteasome bound to PA200 and PA28-like caps, with no evidence for the 26S proteasome 19S regulatory particle. There was also a low abundance complex with a smaller, less-resolved density, that could be assigned to PI31. Surprisingly, these experiments suggest a regulatory role of PI31 in the selective assembly of higher order proteasome complexes, rather than acting as an alternative cap directly modulating the proteasome proteolytic active sites. My results emphasise the importance of fully characterising the proteasome and its interactors, as regulation of this multi-protein complex is likely to occur on many different levels, all of which represent potential drug target sites.

Acknowledgements

First and foremost, thank you Paula for giving me this unique opportunity and providing the scientific support to investigate such interesting biological questions. I am grateful for the privilege to have learned so much and I look forward to seeing the proteasome field continue to unravel. Thank you also to Ed, who kindly offered his EM facility and expertise. My thanks extend to all the da Fonseca lab, past and present; Ana, Migle, and Pavel for the advice, support, and friendship. Special thanks especially to Migle who has been there both personally and professionally and made my time in the lab so enjoyable. Thank you to all the staff at the LMB who provided the resources and knowledge I needed to push projects forward. Special thanks to Stephen McLaughlin, Pat Edwards, Pei-chun Li, Farida Begum, Giuseppe Cannone, Toby Darling, and Jake Grimmett. Thank you to Manu Hegde and Ben Luisi for being great secondary supervisors. Manu- thank you especially for the numerous chats. To my friends within the building and in Cambridge- thank you. You made LMB/Cambridge home and I will miss you all. Special thanks to Nirupa, Chris, and Tom, the peas to my LMB pod, I will always cherish our adventures together and you'll forever hold a special place in my heart. To Fatime, Ben, Mirlinda, Ananth, Eva, Eliska, and Maryam thank you for your friendship and for making Cambridge so fun. Special thanks especially to Eva, Fatime, Migle, Andi, Cristina, Vish, Ananth, Ana, and Paula who were extremely helpful and supportive during the write up process.

To my Chiswick girls (Naima, Em, Iz, Jess, Beth, Soph), the girls I have known a lifetime, thank you for keeping me sane throughout this journey. For the laughs, giggles and silliness. To another lifetime of friendship. To Reem, I don't know where to start. Just know that I'm grateful to have had you around and you make me very proud. مع تمنياتي لك بالمزيد من المغامرات الشيقة في المستقبل القريب إن شاء الله. To Melina, thank you for our PhD chats, we may have an ocean between us but I still feel you by my side just as in Edinburgh- it will be your turn soon and you will ace it. To Mafaz, thank you for challenging every opinion I've ever had.

To my wonderful family, I hit the jackpot with you all. Glenda, thank you for being the kindest soul and the sweetest sister. For always being there, for lending me your clothes (even if I stretch them sometimes!) and your time. Eni, even though you make me feel like your older, uncool, terrible driver of a sister, you teach me so much and I'm thrilled I will always have you around now to fix my computer/bike/anything. To my favourite four-legged friend, Samush, thank you for being such a stress relief. You'll be pleased to know I'll be around for more walks now. To my grandpas, uncles and my many, many cousins: ju dua shume dhe jeni te mrrekellushme. Especially to Greta and Ada, thanks for the parties and the laughter. You are cousins in name, but sisters in spirit. To my little cousin Sonila, I am so proud of you and can't wait to see you do great things. To my aunts (Teze Bukri, Teze Lida, Teze Ari), jeni zonja fantastike dhe idhyllet to mi! Thank you for being such incredibly progressive and strong women and for teaching me that women can be ambitious, successful and kind all at once. To my parents, there are no words that could adequately thank you enough. In the words of the great John Steinbeck "perhaps it takes courage to raise children" and you have shown such courage. I am grateful and forever indebted to you both for moving countries, doing jobs you didn't particularly enjoy, and persevering in a foreign land just to give us these opportunities. You gave us the world dhe shpresoj te ju kthejme.

And finally, thank you to my nana. Who left us far too early last year, who we think of everyday and for whom this thesis is dedicated to. Per ty Nanushe.

Table of Contents

List of Figures

List of Tables

List of abbreviations

2D	Two-dimensional
3D	Three-dimensional
2XTY	Yeast extract and tryptone
AD	Alzheimer's disease
ADP	Adenosine diphosphate
ALS	Amyotrophic lateral sclerosis
AMC	7-amino-4-methylcoumarin
APC/C	Anaphase promoting complex
APF1	ATP-dependent proteolysis factor 1
ATP	Adenosine triphosphate
BSA	Bovine serum albumin
CBL	Cas8stas B-lineage lymphoma
CD2BP2	CD2 antigen cytoplasmic tail-binding protein
CD8+	Cluster of differentiation 8
Cdc4	Cell division control protein 4
Cdc48	Celle division control protein 48
Cic1	Core interacting component 1
CP	Core particle
CRISPR	clustered regularly interspaced short palindromic repeats
Cryo-EM	Cryogenic electron microscopy
Cryo-ET	Cryogenic electron tomography
CTF	Contrast transfer function
DDD	Direct electron detector
DD1 1/2	DNA-damage-inducible protein 1/2
DNAse I	Deoxyribonuclease I
DQE	Detective quantum efficiency
DTT	Dithiothreitol
DUB	Deubiquitinase
DYNLL1/2	Dynein light chain LC8-type
E1	Enzyme-1 ligase
E2	Enzyme-2 ligase
E3	Enzyme-3 ligase
Ecm29	Extracellular matrix 29 protein
EDTA	Ethylenediaminetetraacetic acid
EM	Electron microscopy
ERAD	endoplasmic reticulum-associated degradation

ET	Electron tomography
FACS	Fluorescent-activated cell sorting
FAF1	Fas-associated factor 1
Fbxo7	F-box protein 7
FSC	Fourier shell correlation
GFP	Green fluorescent protein
gRNA	Guide RNA
HbYX	Hydrophobic-Tyrosine-any amino acid
HD	Huntingdon's disease
HECT	Homologous with E6-associated protein C-terminus
HO	Homothallic endonuclease
HRD1	Hypoxia responsive domain-1
i20S	Immunoproteasome
IBR	In-between RING domain
IBs	Inclusion bodies
IEX	Ion exchange chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranaoside
ISCAT	Interformetric scattering
IVA	*In vivo* assembly
kDa	Kilodalton
KPC	Kex-2 proprotein convertase
MCP	Multicatalytic proteinase complex
MDa	Megadalton
MHC	Major histocompatibility complex
MM	Multiple myeloma
MS	Mass spectrometry
NF- κB	Nuclear factor kappa-light-chain enhance of activated B cells
NMR	Nuclear magnetic resonance
Npl4	Nuclear protein localisation protein 4
NTC	Nutcracker
OD_{600}	Optical density
PA26	Proteasome activator complex 26
PA28	Proteasome activator complex subunit 1
PA200	Proteasome activator 200
PAAF1	Proteasomal ATPase-associated factor 1
PAC	Proteasome assembly chaperone
PBS	Phosphate buffered saline
PBST	PBS with 0.1 % Tween-20
PCR	Polymerase chain reaction

PD	Parkinson’s disease
pH	Power of hydrogen
PI31	Proteasome inhibitor protein 31
PR39	Proline and arginine-rich polypeptide
POMP	Proteasome maturation protein
Rad23	UV excision repair protein
RBR	Ring-between-ring domain
RCT	Random Conical tilt
REG	11S Regulator
RHBDL4	Rhomboid-related protein 4
RING	Really interesting new gene
RP	Regulatory particle
Rpn	RP non-ATPase subunit
Rpt	RP ATpase subunit
RVP	Retroviral aspartyl protease domain
SCF	Skp1, Cullin1, F-box protein
SDS	Sodium dodecyl sulphate
SDS-PAGE	SDS polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
SEC-MALS	Size exclusion chromatography coupled with multiple angle scattering
TCEP	Tris(2-carboxyethyl)aminomethane
Ter96	Transitional endoplasmic reticulum ATPase
TNKS	ADP-ribosyl transferase tankyrase
tRNA	Transfer ribonucleic acid
TwStrep	Double strep tag
UB	Ubiquitin
UBA	Ubiquitin-associated domain
UBQLN	Ubiquillin proteins
UBL	Ubiquitin-like domain
UBR4	Ubiquitin protein ligase E3 component N-recognin 4
UBX	Ubiquitin regulatory X domain
UFD	Ubiquitin fold domain
Ufd1	ER-associated degradation protein 1
UIM	Ubiquitin-interacting domain
UN complex	Ufd1-Npl4 complex
UPL	Ubiquitin protein ligases
UPR	Unfolded protein response
UPS	Ubiquitin proteasome system
VBM	VCP binding motif

VCP	Vasolin-containing protein
VIM	VCP interacting motif

Chapter 1 Introduction

1.1 Protein degradation mechanisms

Maintaining and regulating cellular components is essential in ensuring eukaryotic cell survival. Homeostasis is the process in which cells can spatially and temporally control fundamental cellular properties, including intracellular protein content, pH levels, and ion availability (Baumeister et al., 1998). Protein homeostasis, also known as proteostasis, is the process in which proteins are specifically regulated so that biological activity is maintained according to the cellular environment and demand (Baumeister et al., 1998; Kern and Behl, 2019). Six decades of accumulated work has led to the current detailed understanding of proteostasis and its role in maintaining proper cell functionality (Ciehanover et al., 1978; de Duve et al., 1955; Etlinger and Goldberg, 1977; Goldknopf and Busch, 1977). A key part of ensuring sufficient cellular regulation is through protein degradation. In mammalian cells, degradation ensures that unnecessary or unwanted cellular content is removed accordingly and that cellular processes such as the cell cycle are aptly maintained and regulated (Collins and Goldberg, 2017). Cellular degradation occurs within different cellular compartments through two main pathways that mediate protein degradation; autophagy or the ubiquitin-proteasome system (UPS) (Nandi et al., 2006). These two distinct but complementary mechanisms are explained in detail in this chapter.

1.1.1 Autophagy

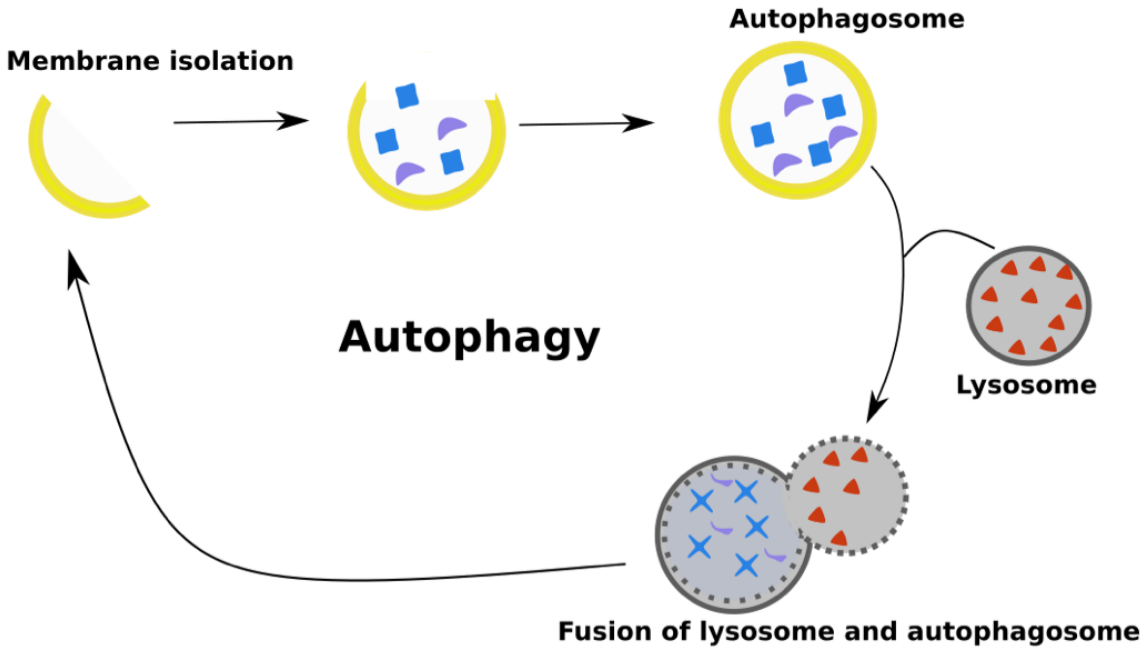


Fig 1.1: **Autophagy.** Autophagy is a multi-step process that breaks down cellular organelles and components enabling cell survival. The process begins with membrane formation, followed by identification and grouping of cellular components for degradation. Membrane sealing forms the autophagosome. The autophagosome then fuses with the lysosome to form the autolysosome. This subsequently results in the degradation of captured materials.

Autophagy is a highly conserved eukaryotic pathway that was first described by Clark and Novikoff more than six decades ago (Clark, 1957; Novikoff, 1959). These studies described mitochondria, from mouse kidneys, enwrapped within membrane-bound compartments that were later shown to consist of lysosomal enzymes (Clark, 1957; Novikoff, 1959). The discovery of the lysosome was an important turning point in the protein degradation field. Until then several experiments had shown that proteins were constantly being synthesised and degraded but experiments had not yet discovered the underlying mechanisms (Schoenheimer, 1946; Simpson, 1953). The discovery of the lysosome provided evidence of how proteolysis could occur (Fig. 1.1). This finding was later followed by the identification of membrane-bound vesicles that contained semi-digested organelles and lysosomal hydrolases (Ashford and Porter, 1962; Novikoff et al., 1962). The term ‘autophagosomes’ to describe these membrane-bound vesicles was introduced by de Duve shortly afterwards (de Duve and Wattiaux, 1966). De Duve’s work

showed that autophagosomes were related to lysosomes and made from the process of autophagy in which organelles and cytosolic matter are engulfed into membrane vesicles (Fig. 1.1) (de Duve and Wattiaux, 1966). Seminal work from the 1950s and the 1990s dissected the well-regulated processes in the lysosome. Autophagy was described as a selective process in which organelles and other exogenous proteins can be degraded (Deter and de Duve, 1967; Mortimore and Schworer, 1977; Seglen and Gordon, 1982). This work was expanded in the molecular era where pioneering studies showed the different steps in lysosomal degradation and its involvement in the digestion of many substrates and pathways, as summarised in Fig. 1.1.

Autophagy can lead to the degradation of exogenous proteins obtained from pinocytosis or receptor-mediated endocytosis, as well as intracellular cytosolic proteins and organelles by micro and macroautophagy (Goldberg, 2007; Mortimore and Pösö, 1987). Microautophagy occurs under basal metabolic conditions when parts of the cytoplasm are segregated within a membrane-bound compartment that then fuses with the lysosomes and results in digestion of protein extracts in an acidic pH environment (Fig 1.1) (Ashford and Porter, 1962). Macroautophagy, on the other hand, happens under more extreme conditions, such as starvation, when whole organelles can be engulfed (Ashford and Porter, 1962). Autophagy provided an explanation of how cellular degradation could be controlled within the confines of membrane-bound vesicles and how exogenous proteins could be targeted. Since then, autophagy has been shown to control a plethora of different cellular processes (Fig.1.2). Understanding the basic principles of autophagy mechanism answered many questions. However, several other questions remained unanswered, including how specific proteins, such as mutant or misfolded proteins and proteins with different half-lives, are targeted.

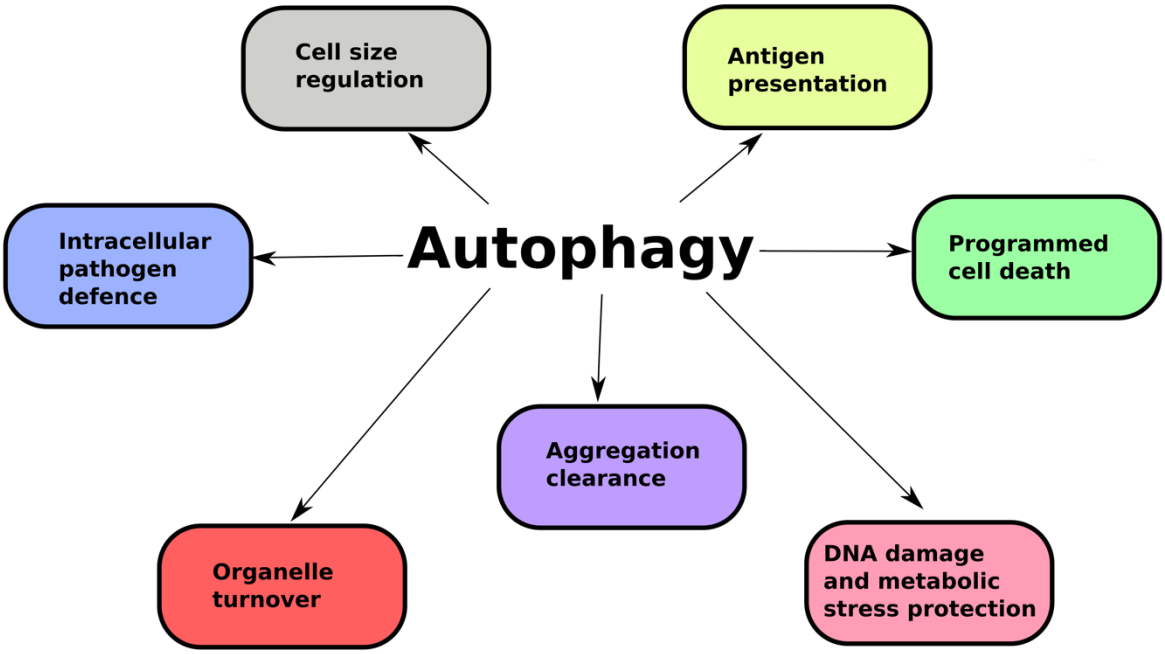


Fig. 1.2: **Autophagy roles in the cell.** Autophagy has a wide range of cellular roles, in which a summary of these functions is listed above.

Progress for discovering this other degradation system was hampered by the lack of a cell-free system that could reconstitute cellular proteolytic events. This was eventually solved by Rabinovitz and Fisher who found that abnormal haemoglobin, which contained amino-acid analogues, was efficiently degraded in lysosome-free rabbit reticulocytes (Rabinovitz and Fisher, 1964). This finding then allowed Etlinger and Goldberg to isolate the first cell-free proteolytic preparation from reticulocytes (Etlinger and Goldberg, 1977). A subsequent experiment showed that abnormal haemoglobin could be degraded at neutral pH and required ATP hydrolysis (Etlinger and Goldberg, 1977). A similar experimental set up was also replicated by Hershko and colleagues shortly afterwards (Ciehanover et al., 1978). Before these surprising findings, it had been shown by Simpson in 1953 that intracellular protein degradation requires ATP, however, no known protease requiring ATP activity had been discovered (Ciechanover, 2005; Goldberg, 2007). The finding of a ATP-dependent system led to discovering a novel proteolytic pathway, now called the ubiquitin-proteasome system (UPS), which is responsible for degrading specifically tagged proteins in the cell.

1.2 The Ubiquitin-Proteasome system

1.2.1 Ubiquitination

The first major development in the UPS system was the discovery of ubiquitin by Goldstein in the 1970s, followed by ubiquitin conjugation as a way to mark proteins for degradation (Ciehanover et al., 1978; Goldstein et al., 1975; Hershko et al., 1980). These findings used cell-free extracts from reticulocytes to purify and characterise enzymes that were involved in the UPS (Etlinger and Goldberg, 1977; Hershko et al., 1978). Using an anion-exchange resin, Ciechanover and Hershko found that fractionating reticulocyte cell extract yielded two fractions that were both needed to reconstitute the proteolytic activity in the extract (Ciechanover et al., 1978). This important finding showed that at least two components are required to make up this novel degradation system, ATP and ubiquitin. In one of the fractions, an approximately 8.5 kDa, heat-stable protein was discovered to be an activator for a protease in fraction two- this was later shown to be ubiquitin but at that stage was called ATP-dependent proteolysis factor-1 (APF1) (Ciehanover et al., 1978). Incubation of ubiquitin with fraction two was shown to require ATP (Ciechanover et al., 1980; Hershko et al., 1980). Goldstein had previously discovered ubiquitin and showed that it covalently bound histone H2A (Goldstein et al., 1975). However, experiments by Wilkinson and colleagues linked the APF1 studies with ubiquitin and showed that APF1 and ubiquitin were the same molecule, a highly conserved 76-amino acid protein present in all eukaryotes (Wilkinson et al., 1980).

In later experiments, ubiquitin's roles in protein degradation were revealed and its fundamental role in the UPS in mammalian cells was first explained by Finley and Varshavsky (1985). The mechanism of ubiquitin conjugation was shown to be similar to that of peptide-bond formation catalysed by tRNA synthetase during protein synthesis (Ciechanover et al., 1981). By immobilising ubiquitin as an affinity 'bait', the three other enzymes that are involved in the ubiquitin conjugation pathway were subsequently purified by Ciechanover, Hershko and colleagues (Ciechanover et al., 1982; Hershko et al., 1983). The three other enzymes were named enzyme-1 (E1), enzyme-2 (E2), and enzyme-3 (E3) (Ciechanover et al., 1982; Hershko et al., 1983). E1 was the ubiquitin-activating enzyme, E2 was the ubiquitin-conjugating enzyme, and E3 was the ubiquitin-protein ligase (Fig. 1.3). The discovery of E3 ligases was particularly welcomed as it provided a viable

solution to the specificity problem of degradation and the existence of proteins with varying half-lives and stabilities; different E3 ligases were involved in targeting different proteins.

1.2.1.1 Ubiquitination mechanism

Several studies provided evidence for the ubiquitin-tagging hypothesis. For instance, ubiquitin conjugation was thought to enhance the denaturation of haemoglobin (Chin et al., 1982). This was also seen with abnormal proteins that contained amino-acid analogues (Hershko et al., 1982). Several studies then led to the discovery that ubiquitination is a multi-step process (Fig. 1.3). During this process ubiquitin must first be activated by an E1 activating enzyme (E1) in an ATP-dependent manner (Ciechanover et al., 1982; Hershko et al., 1983). This process is fundamental as failure to activate ubiquitin by E1 leads to shutdown of the whole system, as seen by E1 enzyme inhibitors (Yang et al., 2007). This reaction results in the formation of a thioester between the C-terminus of ubiquitin and the E1 catalytic cysteine residue from the sulfhydryl group (Fig. 1.3) (Haas et al., 1982). The ability of E1 to transfer and activate ubiquitin may be a result of the flexibility of the ubiquitin fold. This step is then followed by ubiquitin transfer onto the E2 catalytic cysteine (Hershko et al., 1983). Upon its transfer to an E2 enzyme (E2), ubiquitin is then transferred to an E3 ligase enzyme through another ATP-dependent step (Hershko et al., 1982). As previously hypothesised in early ubiquitin studies, E3 ligases (E3) determine the specificity of the target protein and are thus key regulators in this pathway (David et al., 2011). E3s usually bind directly to the E2 and form complexes with the substrate (Hershko et al., 1983). E3 also aids in recognising the protein substrate and stabilising the conformation, however, it is the E2s that are activating the transfer through the conserved E2 cysteine residue and the C-terminus of ubiquitin (Hershko et al., 1982). Within mammalian cells, there are over 600 E3 ligases present and these will be discussed in more detail in Chapter three.

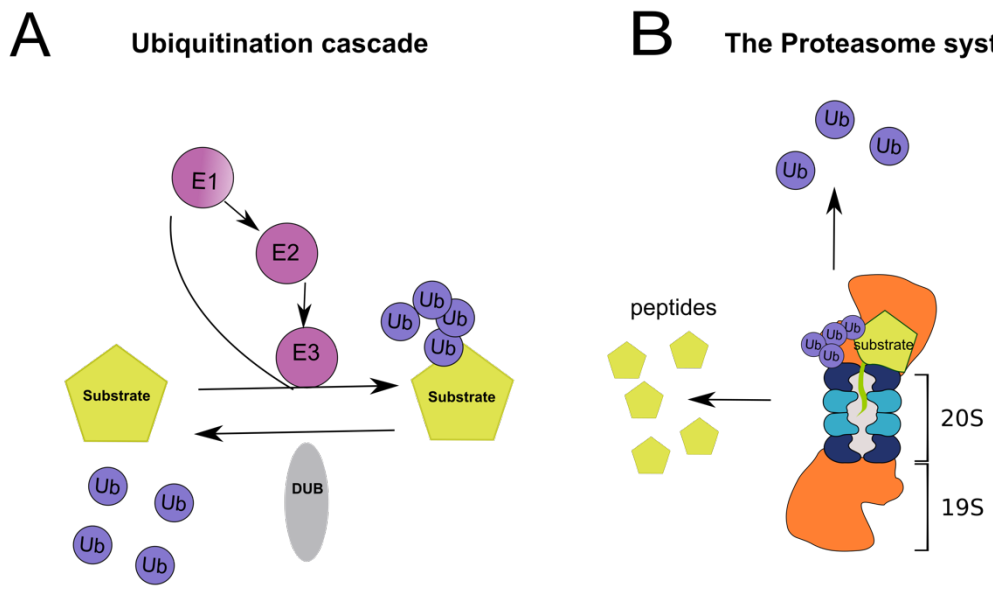


Fig. 1.3: The Ubiquitin-Proteasome System. A) Ubiquitin conjugation on a protein substrate is a result of an E1, E2, and E3 (shown in pink) enzyme cascade in which ubiquitin (purple) is activated by an E1 enzyme in an ATP-dependent step, transferred to an E2 ligase and finally to an E3 ligase where the ubiquitin is conjugated to a protein substrate (yellow pentagon). B) Once a protein substrate is tagged by ubiquitin it can then bind to the 26S proteasome that is composed of a catalytic core particle, the 20S, and an ATP-dependent regulatory particle, the 19S (shown in blue rings and orange respectively). Once bound to the proteasome, the protein substrate's ubiquitins are removed by deubiquitinases into individual ubiquitin moieties and the protein substrate is fragmented into small peptides (small pentagons) in the catalytic chamber.

Proteins can then either be tagged with a single ubiquitin (mono-ubiquitination) or several ubiquitins (poly-ubiquitination) (Fig 1.4). Once one ubiquitin is attached to the substrate, more ubiquitin moieties can be added to make a ubiquitination chain (Fig 1.4). Combinations in which ubiquitin can form chains and ubiquitinate proteins is referred to as the ubiquitin code (Grice and Nathan, 2016; Saeki, 2017). The ubiquitin code is very complex and can be made from one type of linkage or mixed linkages (Grice and Nathan, 2016; Saeki, 2017). Firstly, the protein substrate is tagged with ubiquitin through the formation of an isopeptide bond between the protein's lysine amino terminal and the carboxyl end of ubiquitin. In its sequence ubiquitin contains seven lysines (K) and any of them (K6, K11, K27, K29, K33, K48, K63) can be ubiquitinated. The lysine residues in ubiquitin are found on the surface of the molecule, allowing easy access for bond formation (Vijay-Kumar et al., 1987). In addition, ubiquitin's N-terminal methionine can also be ubiquitinated. The conformations of each ubiquitin are important in allowing

ubiquitin chains to interact with each other (Komander and Rape, 2012). All seven of the ubiquitin's lysines have been observed to undergo peptide bond formation with protein substrates, however, they have been implicated in different cellular regulatory mechanisms and determine cellular fates (Fig. 1.4). Initially, it was established that four ubiquitin chains linked via lysine 48 (K48) destined proteins for proteasomal degradation, referred to as the canonical protein degradation linkage (Thrower et al., 2000). However, more recently, other lysine linkages have also been implicated in proteasomal degradation (including K11 and K29) (Kleiger and Mayor, 2014). Other known ubiquitin linkages have contrasting functions. For example, K63 linkage is known to be implicated in autophagy, protein trafficking and innate immunity (Akutsu et al., 2016) (Fig 1.2). Contrastingly, K27 linkages are involved in nuclear translocation and the DNA damage repair response (Akutsu et al., 2016). The complexity of linkages and possible combinations enable many cellular functions to be controlled by the UPS (Kleiger and Mayor, 2014). The nature of the linkages has a dramatic effect on the overall shape of the chains and how they impart function. It also determines the specifity of the deubiquitinases and thus their regulation (Komander and Rape, 2012).

The discovery of ubiquitin clarified and propelled the field in many ways. However, during the mid-1980s, it was still unclear what the downstream protease that recognised these ubiquitinated substrates was, and this missing link needed to be solved.

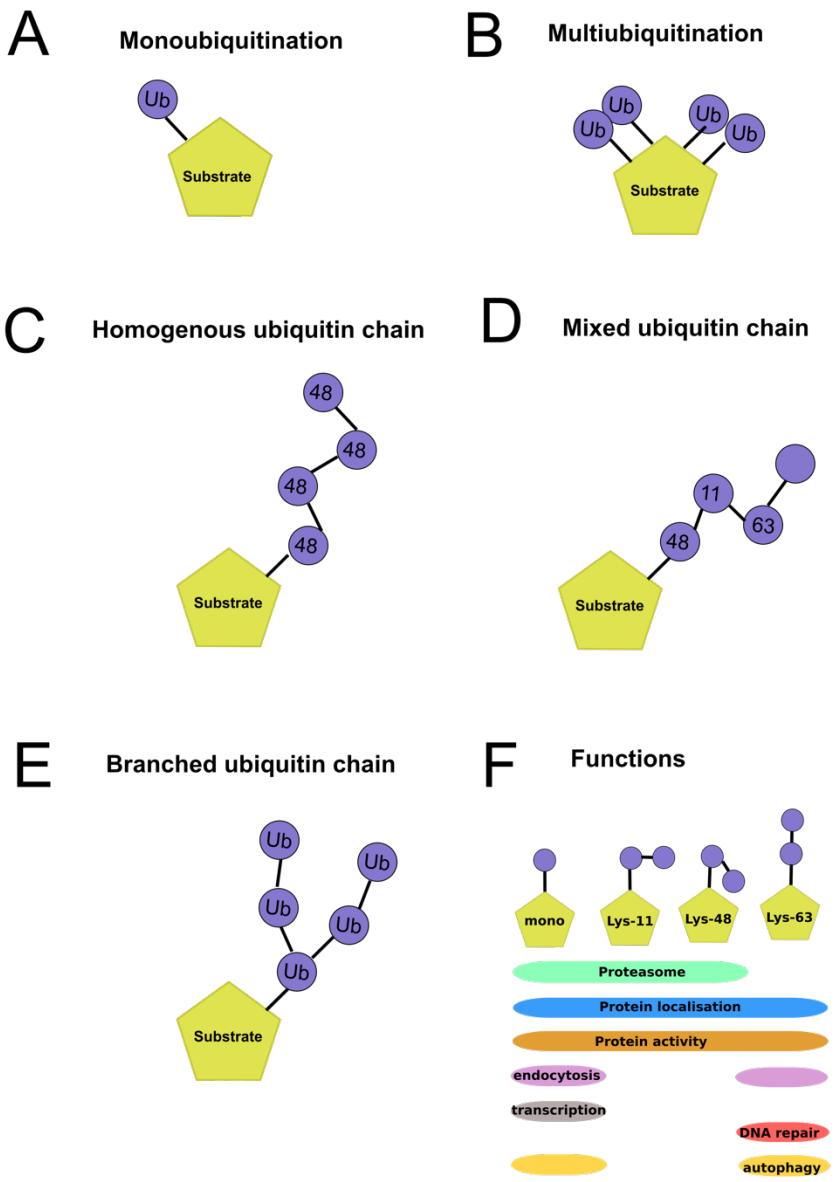


Fig 1.4: Common ubiquitin modifications and their functions. There are many different ways in which ubiquitin (purple circles) can be conjugated to protein substrates (yellow pentagons). A) Substrates can be attached to a single ubiquitin or B) to several ubiquitins. If several ubiquitins are bound, they can be linked to one another in different ways in which the N-terminal methionine or the other seven lysine residues of ubiquitin can conjugate with additional ubiquitin moieties. C) Homogenous ubiquitin chains are a result of the same lysines on ubiquitins linking, for examples Lys63 linkages. D) If different lysine linkages are used this is called mixed ubiquitin chain conjugation. E) Branched ubiquitin chain conjugation occurs when more than one lysine on a single ubiquitin moiety accepts more than one ubiquitin causing branching of ubiquitin chains. F) Ubiquitin modifications can affect the fate of the cell; the function of different ubiquitin linkages and their involvement in varying cellular processes is shown.

1.2.2 The 26S proteasome

The enigma was solved by Tanaka and colleagues who discovered another ATP-dependent step in the reticulocyte system after ubiquitin conjugation (Tanaka et al., 1983). This followed the finding by Hershko and colleagues that ATP is required for ubiquitin conjugate degradation (Hershko et al., 1984). The breakthrough came when Hough and colleagues purified and characterised a high-molecular mass protease that degraded ubiquitin-tagged lysozyme enzymes but not untagged lysozymes, and that this reaction required ATP (Hough et al., 1986). This protease was the 26S proteasome and confirmed all previous notions of it being the specific proteolytic part of the ubiquitin system. Studies by Waxman and colleagues confirmed this protease to be unusually big, around 1.5 MDa in size (Waxman et al., 1987) (Fig. 1.5). A further experiment found the 20S protease complex, together with the 26S complex, and described them as similar to the 'multicatalytic proteinase complex (MCP) identified in the bovine pituitary gland by Wilkinson and Orlowski (Fig. 1.5) (Wilkinson et al., 1980).

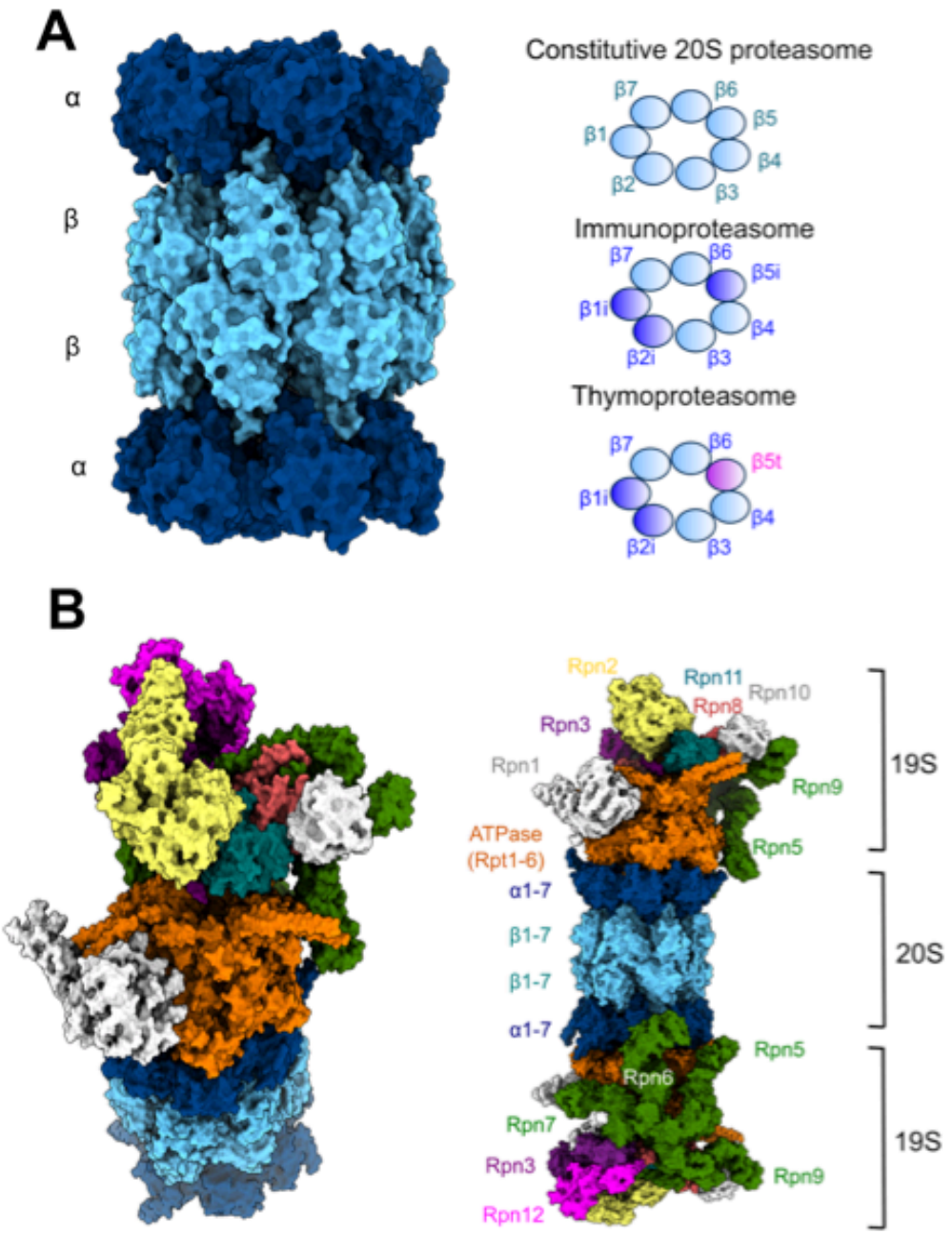


Fig. 1.5: The 20S proteasome isoforms and the 26S Proteasome. A) The 20S proteasome shown in surface representation is composed of two α rings (dark blue) and two ß rings (light blue). The constitutive 20S proteasome has seven ß subunits in which ß1, ß2, ß5 have proteolytic activity. The immunoproteasome and thymoproteasome differs from the constitutive 20S proteasome in the three catalytic ß subunits (dark blue and pink). (B) the 26S proteasome consists of the 19S regulatory particle bound to the 20S proteasome. Different protein components of the 26S proteasome are shown in different colours; the structure was generated by UCSF Chimera by Dr. Kisonaite based on the structure by Huang et al. (2016) (PDB: 5GJR).

1.2.2.1 20S proteasome

The 20S proteasome complex was shown to be an ATP-independent complex with different catalytic activities prompting Hough and colleagues to question the possibility of this complex being part of the 26S proteasome (Hough et al., 1986). This was indeed proved to be the case (Fig. 1.5) (Driscoll and Goldberg, 1990; Eytan et al., 1989). Further evidence was presented by Hoffman and colleagues showing the 20S complex combined with a further 19S subcomplex, which was predicated to have a regulatory role, to form the 26S proteasome (Hoffman et al., 1992). Since then, pioneering work conducted by many scientists have revealed many other intricacies of the 20S complex. We now know that the 20S proteasome is an incredibly complex protease that degrades hundreds of substrates and catalyses approximately 70% of protein degradation in mammals (Collins and Goldberg, 2017). The 20S proteasome's cylindrical shape consists of 14 distinct subunits in eukaryotes which form four heteroheptameric rings, giving rise to the complex known as the 20S core particle (CP) (Fig. 1.5) (Baumeister et al., 1998; Goldberg, 2007; Groll et al., 1997). Within the CP there are seven α subunits forming two α outer rings on each end of the CP and seven ß subunits forming two beta inner rings as shown in Fig.1.5. The catalytic subunits ß1, ß2 and ß5 are located in the inner chamber formed by the ß rings and they confer different proteolytic activity to the 20S proteasome: ß1 confers caspase-like activity, ß2 trypsin-like activity and ß5 chymotrypsin-like activity and cleave acidic, basic and hydrophobic residues respectively. (Dick et al., 1998; Groll et al., 1997). The α subunits, on the other hand, control entry into the central proteolytic chamber. Specifically, in the absence of a regulatory particle, the N-terminal tails of the α subunits are positioned at the gate axis occluding entry to the catalytic pore of the 20S core particle but relocate for substrate entry upon binding of a regulatory particle (Groll et al., 2000).

In mammals, besides the constitutive 20S proteasome, there are at least two other 20S proteasome isoforms, the immunoproteasome and the thymoproteasome (Fig 1.5). The immunoproteasome was discovered in 1994 as a result of studies in antigen presentation which showed it was essential in major histocompatibility complex (MHC) class I antigen processing (Rock et al., 1994). Subsequent experiments showed that 20S function changed upon interferon- γ (IFN-γ) treatment (Aki et al., 1994; Boes et al., 1994; Driscoll et al., 1993; Gaczynska et al., 1993). Later studies showed that the immunoproteasome has different catalytic ß subunit isoforms in its composition (termed ß1i, ß2i, ß5i) that replace

the canonical constitutive 20S proteasome ß subunits. These subunits are specific to the immunoproteasome and expressed in immune system-related tissues (Kloetzel, 2004; Murata et al., 2007). Immunoproteasomes play a crucial role in the immune response and are responsible for degrading inflammatory mediators within immune cells (Murata et al., 2007; Tomko and Hochstrasser, 2013). Similarly, the thymoproteasome, which was discovered in 2007 as specifically expressed in the thymus, share two of the three ß isoforms found in immunoproteasomes (ß1i, ß2i) but a distinct ß5 subunit isoform is only found in this proteasome subtype (termed ß5t). Thymoproteasomes play a role in T-cell development within the thymus (Kloetzel, 2004; Murata et al., 2007).

1.2.2.2 Assembly of the 20S proteasome

Due to its complexity and its fundamental role in protein regulation, assembly of the 14 subunits of the 20S proteasome is a highly regulated process that must occur precisely. The pathway itself relies on the 20S proteasome's intrinsic ability of subunits to self-assemble, with assembly chaperones involved to aid this process (Kunjappu and Hochstrasser, 2014). Assembly requires 20S chaperones, specifically protease assembly chaperone (PAC) 1, PAC2, PAC3, PAC4 and proteasome maturation protein (POMP) (Fig. 1.6). Like other chaperones, the proteasome chaperones aid 20S core particle assembly and do not form part of the final matured proteasome. The chaperones required are the same for all 20S isoforms described above, however the order is different. More specifically, it has been shown that there is simultaneous incorporation of ß2i and ß1i in the first step assembly of 20S isoforms, as well as ß5i incorporation that is independent of ß4 (Bai et al., 2014). The first step of 20S proteasome assembly is the formation of the α ring (Fig. 1.6) (Tomko and Hochstrasser, 2013). This requires chaperones PAC1 and PAC2, which promote the intermediate ring formation of α1, α2, α5, α6, α7. The PAC3-PAC4 hetereodimer then interacts with α5 and causes incorporation of α3 and α4 onto the ring (Fig. 1.6) (Frentzel et al., 1994). Next, the ß ring is formed by initial binding of ß2 to the already-formed α ring, and subsequent recruitment of ß3 and ß4. Incorporation of ß3 and ß4 causes dissociation of the PAC3-PAC4 dimer resulting in the intermediate 13S complex (Fig. 1.6)(Schmidt et al., 1997). POMP then binds to ß5 and recruits ß1 and ß6 to the 13S complex (Hirano et al., 2005). Finally, ß7 is incorporated and the half-proteasome is formed (Fig. 1.6) (Ramos et al., 1998; Yang et al., 1995). Half proteasomes can then

dimerise resulting in the full 20S proteasome formation. The final step results in the maturation of the 20S proteasome in which propeptides present in most ß subunits N-termini are self-cleaved, triggering POMP degradation and release of the PAC1-PAC2 complex (Hirano et al., 2005; Li et al., 2007; Ramos et al., 1998; Stadtmueller et al., 2012).

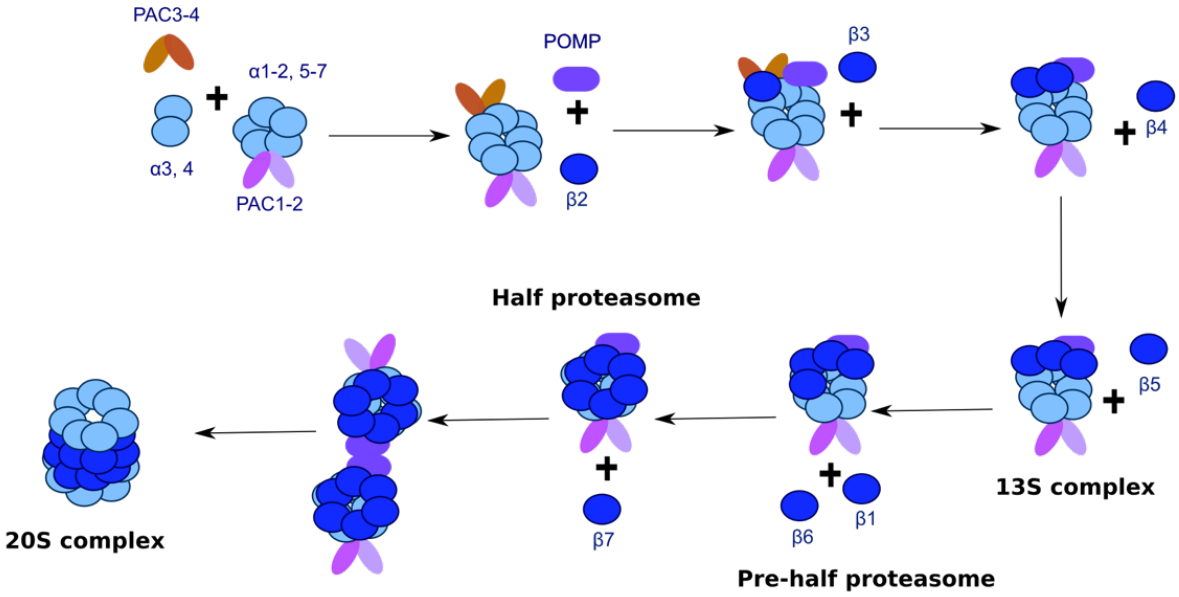


Fig. 1.6: Schematic diagram of the 20S complex assembly. The immunoproteasome and the thymoproteasome assembly mechanism consist of the same chaperones. However, there is simultaneous incorporation of ß2i and ß1i in the first step as well as ß5i incorporation that is independent of ß4 incorporation.

Assembly of the 19S RP, which is discussed later, is also a very regulated process, and requires a number of chaperones. Briefly, the lid and the base assemble independently and then associate with one another to form the 19S cap (Rosseau and Bertolotti, 2018). The base-dedicated chaperones in mammals are S5b, p27, gankyrin, and Rpn14. The chaperones assist in forming three interemediates: Nas2-Rpt4-Rpt5, Nas6-Rpt3-Rpt6-Rpn14 and Hsm3-Rpn1-Rpt1-Rpt2, which then come together and form a heterohexameric ring that binds Rpn2, Rpn10 and Rpn13 (Tomko and Hochstrasser, 2013). Conversely the lid forms two modules: Rpn5-6-Rpn11-Rpn9 and Rpn7-Rpn3-Rpn15 which are then joined and Rpn12 is incorporated with the aid of Hsp90 and Yin6 chaperones (Im and Chung, 2016). Finally, the RP and CP can associate by the HbYX ('Hb' for hydrophobic residue, 'Y' for tyrosine, 'X' for an any amino acid) motifs of the Rpt subunits and the α subunits of the 20S proteasome (Eisele et al., 2018). Various proteins have been implicated

to regulate the RP-CP associate in yeast and mammals including Ecm29 and Hsp90 (Rosseau and Bertolotti, 2018).

1.2.2.3 Structural studies of the 20S proteasome

Structural work on the 20S proteasome began in 1995 when the first x-ray crystal structure was solved for the 20S proteasome from *Thermoplasma acidophilum* (Löwe et al., 1995). This was subsequently followed by the yeast and mammalian proteasomes (Groll et al., 1999; Unno et al., 2002). These studies revealed that archaea and bacteria encode a single α subunit and a single ß subunit that assembles into a sevenfold symmetric 20S CP, whereas eukaryotic 20S proteasomes encode seven different α subunits and seven different ß subunits which assemble into a pseudo-sevenfold symmetric 20S CP (Groll et al., 1999). The yeast structure along with proteasome inhibitor initially showed the distinctive specificities of the three catalytically active, ß1, ß2, and ß5 subunits, in eukaryotic 20S proteasomes (Seemüller et al., 1995). These three catalytic subunits conduct proteolysis through a threonine-dependent nucleophilic attack reaction as shown in Fig.1.7 (Kisselev and Goldberg, 2001; Unno et al., 2002). Subsequent structural studies from bovine 20S proteasomes showed similar catalytically active ß subunits (Unno et al., 2002). Further structural work has shown that the 20S proteasome is structurally conserved in archaea, yeast and mammals (Kish-Trier and Hill, 2013). The human 20S proteasome is 150 Å long, 115 Å in dimeter, and contains an approximately 13 Å central pore. Structures of the 20S proteasome have been solved at the maximum resolution of 1.6 Å (Schrader et al., 2016).

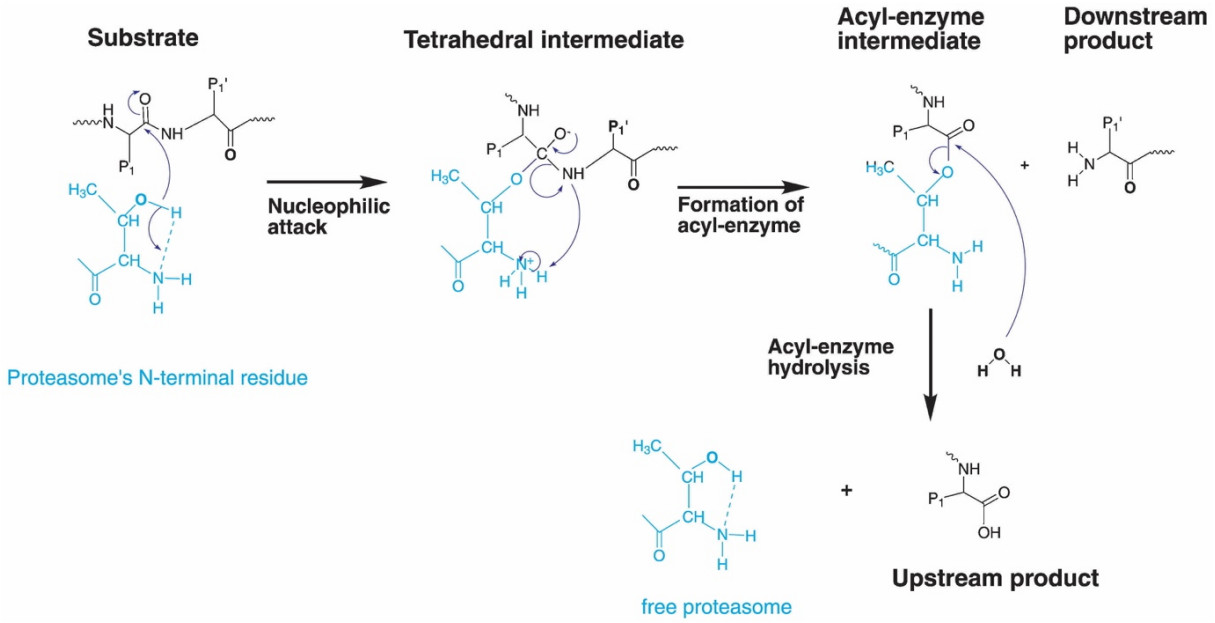


Fig. 1.7: The proteasome catalytic mechanism. Proteasomes catalyse substrate proteolysis by generating smaller peptides with sizes between three to 22 amino acids in length. The chemical mechanism is shown where the substrate is labelled in black and the proteasome in cyan. Reprinted from Kisselv and Goldberg (2001) with permission from Elsevier.

1.3 Proteasome regulators

The 20S proteasome on its own has low proteolytic activity but can be activated by 20S proteasome regulators. Proteasome activation, carried out by proteasome regulators (also known as activators), is critical for the normal function of the 26S proteasome in the UPS (Stadtmueller and Hill, 2011). Proteasome activation in this case occurs by binding of the 19S RP as described below. However, proteasome activity and substrate selection can also be regulated by other proteasome regulators (Stadtmueller and Hill, 2011). Whilst some regulators are well known and established, such as PA28 and PA200, there are others such as P97, Ecm29 and PI31 that are thought to affect proteasome activation but which require further investigation (Inobe and Matouschek, 2014; Stadtmueller and Hill, 2011). Activation of 20S proteasomes seems to involve binding and docking onto the 20S CP, which opens up the 20S gate and allows access to the ß subunits within the 20S CP that confer proteolytic activity (Smith et al., 2007). More specifically, proteasome activators have been shown to interact with the α subunits in the 20S CP, controlling the opening and the closing of the 20S gated pore (Förster et al., 2005). However, it is not yet clear the

extent to which the activators control pore opening or closing, as closed 20S proteasomes have also been found when binding the 19S RP (Unverdorben et al., 2014).

1.3.1 19S regulatory particle

The 19S regulatory particle (RP), also known as PA700, is the canonical regulator for the 20S CP (Fig. 1.5). The 19S RP is about 900 kDa and composed of 18 subunits, each comprising of different regulatory roles. The 19S includes ubiquitin receptors that can recognise ubiquitinated substrates labelled for degradation, de-ubiquitinating enzymes that remove ubiquitin, and proteins that are involved in substrate unfolding and translocation (Ehlinger and Walters, 2013). All of these subunits are listed in Table 1.1. The 19S itself can be sub-divided into the base and lid sub-complexes (see Fig 1.5 and Table 1.1).

The base consists of the six ATPase subunits (Rpt) and four non-ATPase subunits (Rpn) and is responsible for the recognition of ubiquitinated proteins, substrate unfolding and to facilitate the opening of the 20S proteasome gate (Table 1.1) (Glickman and Raveh, 2005). The Rpn subunits in the base are Rpn1, Rpn2, Rpn10 and Rpn13 of which Rpn10 and Rpn13 are known UB receptors, and with Rpn13 only found in yeast 26S proteasome purifications (Fig 1.5 and Table 1.1) (Huang et al., 2016; Schweitzer et al., 2016). Rpn1 has also recently emerged as a ubiquitin receptor but is thought to also have scaffolding roles, along with Rpn2, within the 19S RP (Lander et al., 2012; Shi et al., 2016, p. 1). The Rpt subunits, on the other hand, form a heterohexameric ring responsible for unfolding and translocation of substrates. Like other regulators, the Rpt subunits contain HbYX motifs at most of its C-terminus (Förster et al., 2005; Smith et al., 2007). The HbYX motifs interact with the 20S by inserting into grooves at the 20S α ring interface (Förster et al., 2005; Smith et al., 2007).

Table 1.1: Mammalian 19S regulatory particle subunits

Sub-particle	Subunits	Function
Base	Rpn1, Rpn10, Rpn13	Initial Ub binding, Ub receptor
	Rpt1, Rpt2, Rpt3, Rpt4, Rpt5, Rpt6	AAA+ ATPase involved in unfolding and translocation
	Rpt2, Rpt3, Rpt5	20S gate opening
	Rpn2	Substrate binding
Lid	Rpn3, Rpn5, Rpn6, Rpn7, Rpn8, Rpn9, Rpn12, Rpn15	Scaffold
	Rpn11	Removes Ub from substrate

The lid of the 19S on the other hand contains Rpn3, Rpn5-9, Rpn11, Rpn12 and Rpn15 (Livneh et al., 2016). Rpn3, Rpn5-9, Rpn 12 and Rpn15 are scaffolding subunits, whereas Rpn11 is a deubiquitinating enzyme that binds and forms a heterodimer with Rpn8 (Fig 1.5 and Table 1.1) (Bhattacharyya et al., 2014; Verma et al., 2002). The subunits found in the 19S lid are homologous to the COP9 signalosome complex, which is critical in the development of multicellular organism, and have been found in most eukaryotic organisms, including yeast and human (Cope and Deshaies, 2003; Glickman et al., 1998).

As mentioned above (and summarised in Table 1.1) the 19S subunits are involved in the ubiquitin-dependent degradation of substrates. The general mechanism of ubiquitin-dependent protein degradation requires two degradation signals (also called degrons): an unfolded initiation region of the protein that can begin the unfolding process, and a ubiquitin tag which can be recognised by ubiquitin receptors Rpn10, Rpn13 and Rpn1 (Inobe et al., 2011). Once the substrate is engaged, the proteasome can initiate degradation. The initiation receptors are not specifically known, however, the pores located within the ATPase rings of the 19S RP have been proposed as receptor sites (Huang et al. 2016). The association of the protein with the ATPase rings commits a substrate for degradation and is also important in the unfolding and translocation steps (Collins and Golberg, 2017). As mentioned previously, the Rpt subunits with their ATPase activity are also important in gate-opening and allowing the substate to progress to the 20S CP for degradation. ATPase activity, which is also required for ubiquitin chain binding, is also important to committing

a substrate for degradation (Collins and Goldberg, 2017). Translocation and unfolding steps in the degradation mechanism have proved very difficult to study, however, studies have shown that translocation causes the upstream protein to linearise and the movement is generated by the loops in the ATPase rings (Collins and Goldberg, 2017). For successful degradation to occur, the engaged substrate is also deubiuqitinated by deubiquitinases (such as Rpn11), as discussed further in Chapter three, which are proteases that reverse the ubiquitin modification on proteins (Reyes-Turcu et al., 2009).

As well as ubiquitin-dependent degradation, ubiquitin-independent degradation of some substrates such as Rpn4 have been identified, which do not involve the 19S RP (Erales and Coffino, 2014). Some further examples of ubiquitin-independent degradation are discussed in Chapter 5.

1.3.1.1 Structural studies of the 19S RP and the 26S proteasome

The first 26S proteasome structure was solved in 1993 using negative stain electron microscopy (EM) to show the double-capped structure of the yeast 26S proteasome (Peters et al., 1993). However, it was not until 2016 when a high resolution structure (3.9 Å) of the 26S revealed new structural insights of the 19S RP organization on the 20S CP (Schweitzer et al., 2016). Additional high resolution structures of the 26S proteasome were published by Luan et al. (2016), Huang et al. (2016) and Chen et al. (2016). Previously, all 32 subunits had been assigned to the 26S proteasome low/mid resolution structures using other techniques such as protein labelling, crosslinking mass spectrometry, and mutagenesis studies (Beck et al., 2012; da Fonseca et al., 2012; Lander et al., 2012; Lasker et al., 2012; Sakata et al., 2012). The high resolution structures showed a number of different conformations and new structural insights on the nucleotide occupancy on the Rpt subunits. For instance, one particular conformation showed that Rpt6 was ADP bound while the five other Rpts were ATP bound (Huang et al., 2016). The connection between Rpn1 and Rpn2 subunits, thought to be a result of Rpn1 movement, as well as C-terminus protrusion of Rpn3 in the 20S gate interface, was also revealed. More recently, Haselbach et al. (2017) determined the structure of human 26S proteasome with Oprozomib, the proteasome inhibitor, showing a preferential ‘ground conformation’ of the 19S RP where the proteasome is ready for substrate binding, suggesting that drugs can allosterically

regulate the 19S RP (Haselbach et al., 2017). In addition, further conformational changes of the 19S RP were described in substrate-bound complexes where ATP was found to specifically induce these conformational changes thus driving substrate translocation through the 20S central pore (de la Peña et al., 2018; Eisele et al., 2018). The many different conformational states now described for the proteasome highlight the complexity of the mechanism and why it has taken so long to solve high resolution structures of the 26S proteasome.

The C-terminal HbYX motifs of Rpt1, 2, 3, 5 and 6 of 19S RP interact with the α subunits in the 20S proteasome. More specifically, the adjacent α subunits form pockets which the Rpt can insert and form salt bridges with the lysines from α subunits (Rabl et al., 2008). Additional contact sites between the 19S RP and the 20S proteasome include Rpn5 and Rpn6 with the side of the 20S proteasome. The interaction between 19S RP and 20S proteasome is also thought to be regulated by other proteins, such as Ecm29 and Hsp90, however, the exact mechanisms remain unknown (Murata et al., 2009).

1.3.2 PA28/11S

The proteasome activator complex 28 (PA28), also known as 11S or REG, is an alternative regulator of approximately 28 kDa and a well-known activator of the 20S proteasome that does not require ATP hydrolysis (Fig. 1.8) (Mott et al., 1994). In higher eukaryotes, three PA28 isoforms are expressed, PA28 α, ß and γ, with PA28γ thought to be a more ancient variant as it is found in a wider range of metazoans, contrary to PA28α or PA28ß (Masson et al., 2009; Rechsteiner and Hill, 2005). PA28α and PA28ß form a heteroheptamer whereas PA28γ forms a homoheptamer (Masson et al., 2009). Other PA28 variants can also be found in nature. For example, a homoheptameric proteasome activator, PA26, was found in *Trypanosoma brucei* and although it has been shown to activate the proteasome, its sequence seems to have diverged from that of PA28 (Whitby et al., 2000). PA28αß is thought to have an important role in the immune cell response with studies showing that PA28αß-bound proteasomes are involved in the production of MHC-class I ligands. However, the exact mechanism by which this happens remains unclear (Groettrup et al., 2010). In general, it is well reported that PA28 stimulates peptide hydrolysis, rather than

protein degradation, however, PA28γ has been implicated in the degradation of natively unfolded proteins (Mao et al., 2008; Nie et al., 2010; Suzuki et al., 2009).

Structures of PA28 complexes have been determined by x-ray crystallography (Förster et al., 2003; Huber and Groll, 2017; Knowlton et al., 1997; Whitby et al., 2000). Structurally, PA28α includes elongated helical bundle subunits that assemble to form a heptameric ring with a diameter of 20-30 Å (Knowlton et al., 1997). As with the 19S RP, the carboxylate groups of the C-terminal residues within the HbYX motifs interact with proteasome lysine side chains of the α subunit (Förster et al., 2005, 2003; Whitby et al., 2000). Activation seems to be a result of the disruption of hydrogen bond packing within the 20S proteasome, triggering a widening of the pore and allowing new contacts to be made between the activators and the 20S proteasome (Förster et al., 2003). Moreover, residues Y7, D9, P17 and Y26 in the yeast and archaeal 20S proteasome α subunits seem to stabilise the open conformation and are highly conserved even amongst species that do not express PA28 (Förster et al., 2003). However, this is a highly contested issue. Further structural studies on the mammalian proteasome activator has shown that PA28αß heteroheptamers bind the proteasome, whereas PA28ß homoheptamers, that can also form at least *in vitro*, do not associate with the proteasome (Huber and Groll, 2017). Despite structural and biochemical work shedding light into how PA28 binds to the proteasome and promotes gate opening, the physiological and functional implication of this association remain elusive. Previous work has shown that PA28 plays an important role in degrading substrates, however, how these substrates bind or translocate into the 20S proteolytic chamber is not apparent (Stadtmueller and Hill, 2011). One possibility is that substrates translocate through the activators central pore to the 20S catalytic chamber without directly binding to it or bind to the PA28 directly and then subsequently translocate to the 20S catalytic chamber (Stadtmueller and Hill, 2011). However, this remains unknown.

PA28 was also found in hybrid proteasomes (proteasomes with two different regulators bound at opposite sides of the 20S CP) (Cascio and Goldberg, 2005). The biological significance of this is unknown but it has been proposed that hybrid complexes allow localisation of the 19S-activator complex to a specific cellular environment and conduct ubiquitin-linked degradation of proteins (Rechsteiner and Hill, 2005).

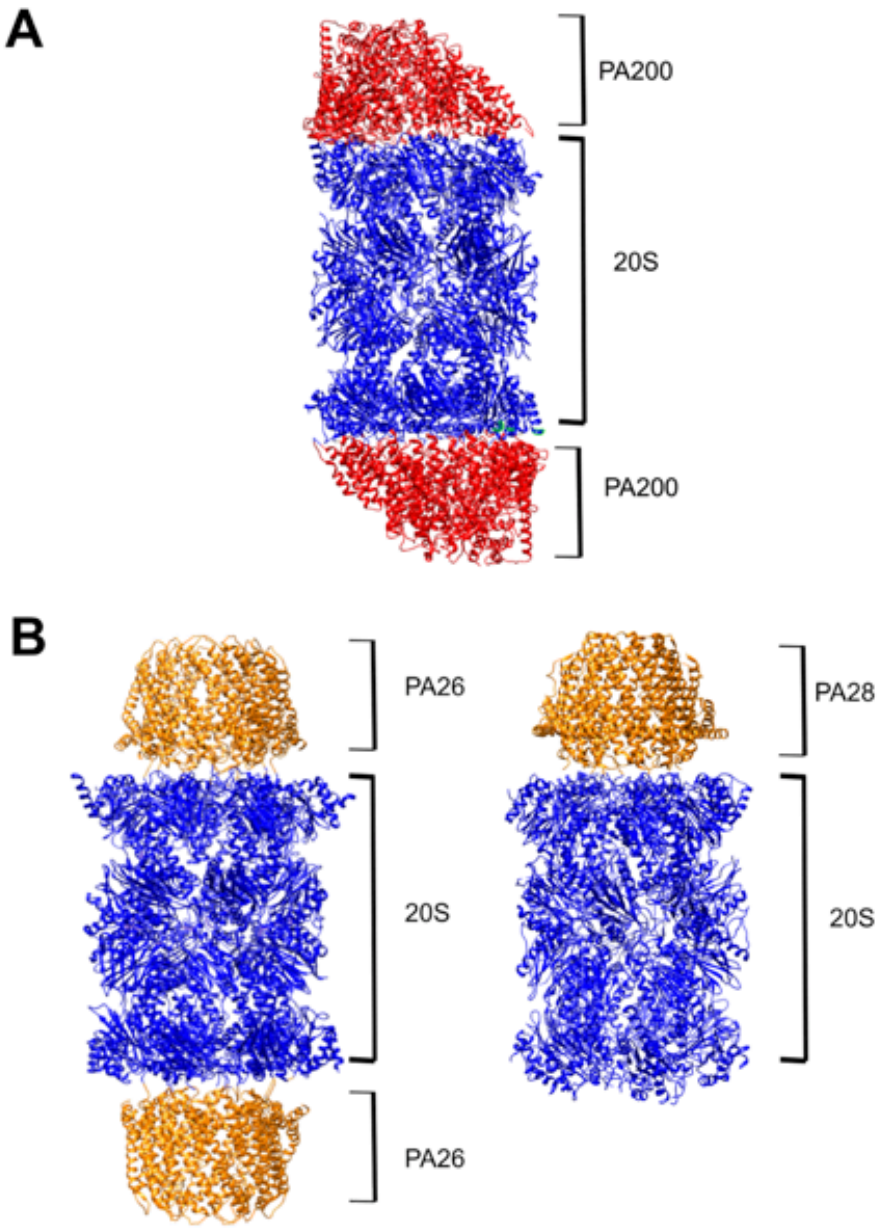


Fig. 1.8: The PA28 and PA200 proteasome activators. As well as the 19S RP, the 20S proteasome (blue) can bind to A) PA200 (red) or B) PA28 and PA26 (orange) activators, in which PA26 is a PA28 variant. Structures were produced on UCSF Chimera and are based on (PDB:6REY) for the human PA200-20S proteasome complex and PA26-20S proteasome complex from *T. brucei* (PDB: 1FNT) and *P. falciparum* PA28-20S proteasome complex (PDB: 6MUX).

1.3.3 PA200

Proteasome activator 200 (PA200), or Blm10 in yeast, is an approximately 200 kDa protein, characterised as an ATP and ubiquitin independent regulator (Fig. 1.8) (Rêgo and Fonseca, 2019). PA200 is thought to stimulate peptide, rather than protein hydrolysis and

has been implicated in a variety of cellular processes, including 20S proteasome assembly, genomic stability, DNA repair, mitochondrial regulation, and spertmatogenesis (Blickwedehl et al., 2008; Fehlker et al., 2003; Khor et al., 2006; Sadre-Bazzaz et al., 2010; Ustrell et al., 2002). A recent study on the human PA200-20S complex revealed that upon PA200 binding the proteasome showed an enhanced trypsin-like activity, usually associated with the ß2 active site, and a slight inhibition of the caspase and chymotrypsin-like activities, usually associated with ß1 and ß5, respectively (Rêgo and Fonseca, 2019). The structure by Rego and da Fonseca (2019) also revealed extensive contacts between PA200 and the surface of all the α subunits apart from the α7 subunit of the 20S proteasome. Interestingly, besides these extensive contacts, PA200 seems to bind to the 20S CP via two major anchoring points. The first is binding to the C-terminus using the typical HbYX motif and second, through an internal loop that makes novel hydrogen bonds with the 20S CP but binds in a similar alpha ring interface at the C-terminal tail. Three of the 20S alpha subunits (alpha 5, 6 and 7) N-terminal tails move away from the main pore axis into grooves on the PA200. These interactions seem to cause an unusual opening of the 20S proteasome axial channel without inducing any further significant conformational changes in the rest of the 20S proteasome complex. This finding showed the high degree of conformational plasticity of the 20S α ring and is consistent with other proteasome-activator complexes (Rêgo and Fonseca, 2019).

1.3.3.1 PI31

Proteasome inhibitor 31 (PI31) has also been shown to regulate the proteasome by inhibiting proteolytic activity (Chu-Ping et al., 1992). However, the exact physiological roles of PI31 remain elusive (Zaiss et al., 1999). PI31 is covered in more detail in Chapter five.

1.3.4 The role of proteasomes in disease

Due to its fundamental role in basic cellular processes such as differentiation, proliferation, apoptosis, gene transcription, signal transduction, metabolic regulation, and immune surveillance, the 26S proteasome has been implicated in many diseases (Asher et al., 2005; Bowerman and Kurz, 2006; Collins and Goldberg, 2017; Naujokat and Hoffmann, 2002; Strehl et al., 2005; Taylor and Jobin, 2005). These diseases include neurodegenerative

disorders, viral and bacterial infections, cancer, parasitic diseases, and inflammatory diseases (Cavo, 2006; Krause et al., 2006; McNaught and Jenner, 2001; Oh et al., 2005; Ott et al., 2003). Table 1.2 summaries these diseases and some of these links are described in more detail below.

1.3.4.1 Neurodegenerative diseases

In healthy cells, misfolded proteins are effectively degraded or refolded correctly by chaperones (Hartl and Hayer-Hartl, 2009). However, in disease-associated cells, the accumulation of misfolded proteins can give rise to diseases such as Alzheimer's disease (AD), Parkinson's disease (PD), amyotrophic lateral sclerosis (ALS), and Huntington's disease (HD) (Díaz-Hernández et al., 2004; Kabashi et al., 2004; Keck et al., 2003; McNaught and Jenner, 2001; Oh et al., 2005; Zhou et al., 2003). Several neurodegenerative diseases involve aggregation or misfolding of specific proteins into abnormal species. Under proteotoxic stress conditions, proteins can escape the cell's quality control mechanisms which ensure proper protein folding. This results in formation of aggregates that range from oligomers to high order fibrils and plaques. Once formed, higher order aggregates can be highly resistant to degradation, as in the case of amyloid aggregates found in AD patients (Finley, 2009). Proteasomes are unable to effectively degrade these aggregates as they have no partial or fully unfolded regions (Finley, 2009). Furthermore, amyloid aggregates are extremely stable as they form extensive contacts between the protein chains of the polymer (Brundin et al., 2010). Other neurodegenerative diseases, such as HD which results in aggregates of glutamine repeats, have been shown to be resistant to proteasomal degradation (Venkatraman et al., 2004). Studies have also shown that as cells' pool of misfolded proteins increase, cells experience "proteostatic collapse", where both proteasomes and chaperones become "overwhelmed" and unable to deal with them, causing the accumulation of ubiquitin inclusion bodies (IBs) (Hipp et al., 2014, 2012; Nonaka et al., 2009). IBs have also been suggested to inhibit or clog the proteasomes directly (Hipp et al., 2014). In the case of HD, ubiquitinated Huntingtin protein is found to accumulate in IBs (Myeku et al., 2016). Huntingtin has also been reported to increase degradation of the transcription co-activator CBP by proteasomes resulting in toxicity (Ross and Pickart, 2004).

As well as overwhelming the UPS system, proteins such as α-synuclein, that is genetically linked to PD, directly bind the 19S RP and inhibit the degradation of other proteins (Ross and Poirier, 2004). Further, the UB mutant UBB+1, observed in AD patients, inhibits the proteasome and results in ubiquitinated protein accumulation and cell cycle arrest (Lam et al., 2000; Lindsten et al., 2002; van Tijn et al., 2007). In general, several studies have shown clear links between the proteasome and neurodegenerative diseases. However, the paucity and difficulty of *in vivo* studies limits understanding in the field and therefore more studies are needed to shed light into specific mechanisms and therapies.

1.3.4.2 Multiple myeloma

Multiple myeloma (MM) is a type of bone marrow cancer that affects several areas of the body including the ribs, pelvis, skull and spine due to the accumulation of plasma cells (Adams et al., 1999; Chauhan et al., 2006). Plasma cells are terminally differentiated cells almost exclusively dedicated to a high rate of protein synthesis to produce secreted antibodies. In MM cells, the UPS has upregulated activity. This results in excessive degradation of protein substrates, such as the inhibitor of nuclear factor- κB (NF- κB) and the tumour suppressor p53 that usually control cell growth and inhibit tumours (Adams, 2004). More specifically, NF- κB, which plays a key role in the inflammatory response has been shown to promote the growth and survival of MM cells through the NF- κB transcription pathway (Cavo et al., 2007; Chauhan et al., 2008). Thus, erroneously degrading the inhibitor of NF- κB causes the MM cells to further spread through the body (Hideshima et al., 2001; Leestemaker et al., 2017). In general, proteasome subunit abundance and, therefore, cell growth, is regulated through controlled expression mechanisms involving the mammalian nuclear factor erythroid 2-related factor 1 (NRF1). When the proteasome is inhibited or there is proteasome overload in the cell, NRF1 is cleaved into its active form and increases transcription of proteasome subunits, thus presenting a way in which proteasome abundance can be regulated (Rosseau and Bertolotti, 2018).

Inhibiting proteasomes has been one of the most successful treatments of MM. Specifically, the proteasome inhibitor Bortezomib/PS-341 was the first successful inhibitor to enter clinical trials in cancer patients and is now a clinically approved drug for MM

(Boise et al., 2014). Bortezomib reversibly inhibits the ß5 subunit active site of the 20S proteasome and to a lesser extent the ß1 subunit active site (Crawford et al., 2006; Ruschak et al., 2011). Inhibiting of the proteasome causes protein overload in the cells, endoplasmic reticulum (ER) stress and thus cell death in MM cells (Gandolfi et al., 2017). Despite its success in MM patients, challenges such as developing more selective and less toxic inhibitors remain (Boise et al., 2014; Chauhan et al., 2005). Furthermore, observed resistance to Bortezomib is an emerging clinical impediment and one that requires further drug development (Franke et al., 2012).

1.3.4.3 Inflammatory and autoimmune diseases

The immunoproteasome has been shown to play a significant role in inflammatory diseases, with genetic studies showing that mutations on the ß5i subunit of the immunoproteasome cause inflammatory disorders (Basler et al., 2013). Mutations in ß5i can lead to accumulated poly-ubiquitinated proteins and a build-up of substrates within cells (Basler et al., 2013). Other subunits have also been shown to have a disease effect. For example, ß2i/ ß5i knock out mouse models have increased CD8+ T-cell mediated autoimmune symptoms (Basler et al., 2013). Furthermore, prior work has shown the immunoproteasome to have enhanced proteasomal proteolytic activity compared to the constitutive 20S proteasome (Gaczynska et al., 1993; Strehl et al., 2008; Voigt et al., 2010). The reason for this increased activity may be to prevent accumulation of substrates that would otherwise aggregate during inflammation (Seifert et al., 2010). However, higher levels of immunoproteasomes have also been shown in a range of inflammatory and autoimmune diseases including hepatitis and Crohn's disease (Miller et al., 2013). The immunoproteasome is thus a clear therapeutic target for a variety of immune diseases. However, the use of proteasome inhibitors to treat these diseases is not as suitable due to the adverse side effects on the immune system response to infection (Kammerl and Meiners, 2016; Miller et al., 2013).

1.3.4.4 Malaria

Malaria is caused by Plasmodium parasites with the most detrimental form caused by *Plasmodium falciparum (P. falciparum)* (Krishnan and Williamson, 2018). Plasmodium

parasites have three different proteasome species that are essential for their survival, including the typical eukaryotic 26S proteasome (Li et al., 2012). Specifically inhibiting the Plasmodium 20S proteasome is therefore a strategy for developing anti-malarial drugs, given the essential role the 26S proteasome plays in the Plasmodium life cycle (Li et al., 2016). There have been a number of 20S inhibitors tested as potential drug candidates that have shown toxicity on Plasmodium in both their infective and developmental stages (H. Li et al., 2014; Li et al., 2012). However, the concern with these drugs is that they lack specificity as they have been identified using mammalian systems and therefore may have effects on mammalian proteasomes of the affected person (Bogyo and Wang, 2002; Lee and Goldberg, 1998). Recently, the cryo-EM structure of the *P. falciparum* 20S proteasome structure in complex with parasite-specific inhibitors have been solved by Morris and da Fonseca (2017), providing a new potential approach in which specificity and toxicity can be targeted to parasitic systems without affecting mammalian hosts (Li et al., 2016; Morris and da Fonseca, 2017). Further, proteasome inhibition is being explored for other parasitic diseases, such as sleeping sickness, Chagas disease and Leishmaniasis (Bibo-Verdugo et al., 2017; Khare et al., 2016).

Table 1.2 Diseases affecting the proteasome system

Disease Group	Disease	Effect on proteasome activity	Proteasome population affected	Symptoms	References
Cardiac dysfunction	Transient ischemia	Decrease activity	26S	Apoptosis	Keller et al., 2000; Bulteau et al., 2001
	Inclusion body myositis	Decrease activity	26S	Inclusion bodies	Grune et al., 2004; Ferrer et al., 2004
Cancer	Multiple myeloma	Increase activity, depressed expression	20S, 26S, immunoproteasomes	Suppression of apoptosis, proliferation	Cavo, 2006
	Renal carcinoma				Johnson et al., 1998; Meidenbauer et al., 2004
Autoimmune	Sjogren's syndrome	Decrease expression	Subunit ß1i	Tissue destruction	Krause et al., 2006
Cataract formation	Cataract formation	Decrease activity	20S	Aggregation of proteasomes	Zetterberg et al., 2003
Viral infections	Hepatitis B	Inhibition	20S and 26S	Hepatitis	Hu et al., 1999
	HIV/adenovirus	Inhibition/decreased expression	20S, immunoproteasomes	Impaired immune response	Ott et al., 2003; Seeger et al., 1997
Neurodegenerative	Amyotrophic lateral sclerosis	Decrease activity	20S and 26s	Motor neuron loss	Kabashi et al., 2004
	Parkinson's	Decrease activity	20S and 26S	Neuronal loss, Lewy bodies	McNaught and Jenner, 2001
	Huntington's	Decrease activity	20S, 26S, induction of immunoproteasomes	Neuronal loss, poly-glutamine inclusions	Zhou et al., 2003; Diaz-Hernandez et al., 2004
	Alzheimer's	Decrease activity	20S and 26S	Tau tangles, neuronal loss, ß-amyloid plaques	Oh et al., 2005; Keck et al., 2003

1.3.5 The proteasome interaction network

As well as proteasome activators, other proteins are thought to bind to the 20S proteasome and the 19S RP, which are shown in the summary Figure 1.9. Before all 32 subunits of the 26S proteasomes were mapped, it was difficult to analyse whether a protein was a proteasome subunit or a proteasome interacting protein. To distinguish between proteasome subunits and proteasome-interacting proteins, it was generally assumed that high salt treatment would break the ionic protein-protein interactions and therefore show that the protein is an associated protein rather than an integral subunit (Schmidt et al., 2005). Using mass spectrometric methods, tagging approaches and genome-wide two hybrid analyses, many novel protein-proteasome interactions were found (Cagney et al., 2001; Davy et al., 2001; Gavin et al., 2002; Ho et al., 2002). These novel interactions include proteins with ubiquitin-like domains/ubiquitin-associated domains (UBL/UBA): ubiquitin ligases core interacting component (Cic1), and extra cellular matrix protein 29 (Ecm29) as well as the already reported activators described above. Ubiquitin ligases such as the Ubiquitin protein ligase E3 component N-recognin 4 (UBR4), P97, and proteasome inhibitors will be covered in more detail in Chapters 3, 4 and 5 respectively, whereas this section will highlight the breadth of other associated proteins acting with the proteasome.

In order to effectively degrade substrates, 26S proteasomes must recognise polyubiquitinated substrates and this is accomplished by the several ubiquitin receptors within the 19S RP, as discussed in Section 1.3.1. Proteins that bind to these receptors and bring ubiquitinated substrates are referred to as shuttling factors. Examples of shuttling factors include UV excision repair protein (Rad23A/B), Ubiquillin proteins (UBQLN1-4) and DNA-damage-inducible protein 1/2 (DD1/2). These shuttling factors contain one UBL domain and one or two UBA domains (Finley, 2009). Normally, these proteins exist in an autoinhibitory state in which UBL interacts intramolecularly with the UBA domains, however, upon ubiquitination the ubiquitinated substrates bind to the UBA domain, freeing up the UBL domain to bind to the proteasome (Saeki, 2017). In this way, ubiquitinated factors can be taken to the proteasomes. Shuttling factors can exhibit specificity for the factors in which they shuttle, for example, Rad23 is specific for K48-linked chains and is shown to bind to Rpn1 and Rpn4 (Chen et al., 2016; Nathan et al., 2013; Raasi et al., 2005). For UBQLN proteins, however, the single UBA domain binds to any ubiquitin chain linkage suggesting that some shuttling factors are more specific than others (Hjerpe

et al., 2009; Raasi et al., 2005). UBQLNs have been shown to bind to newly synthesised transmembrane proteins within the cytosol, showing their importance in facilitating membrane targeting and preventing aggregation (Itakura et al., 2016; Suzuki and Kawahara, 2016). DDI1/2, on the other hand, is unique to other shuttling factors as it contains the retroviral aspartyl protease (RVP) domain found in HIV. DDI1 and DDI2 also do not possess a UBA domain and instead binds ubiquitin through their UIM domain (Nowicka et al., 2015). Recently, DDI2 has emerged as a transcriptional regulator of proteasome genes as it cleaves the transcriptional factor Nrf1 that controls proteasome gene expression, when proteasome activity is compromised (Koizumi et al., 2016; Lehrbach and Ruvkun, 2016; Radhakrishnan et al., 2010). Several other UBL-UBA proteins exist within mammals with proteasomal functions that are not entirely understood (Saeki, 2017).

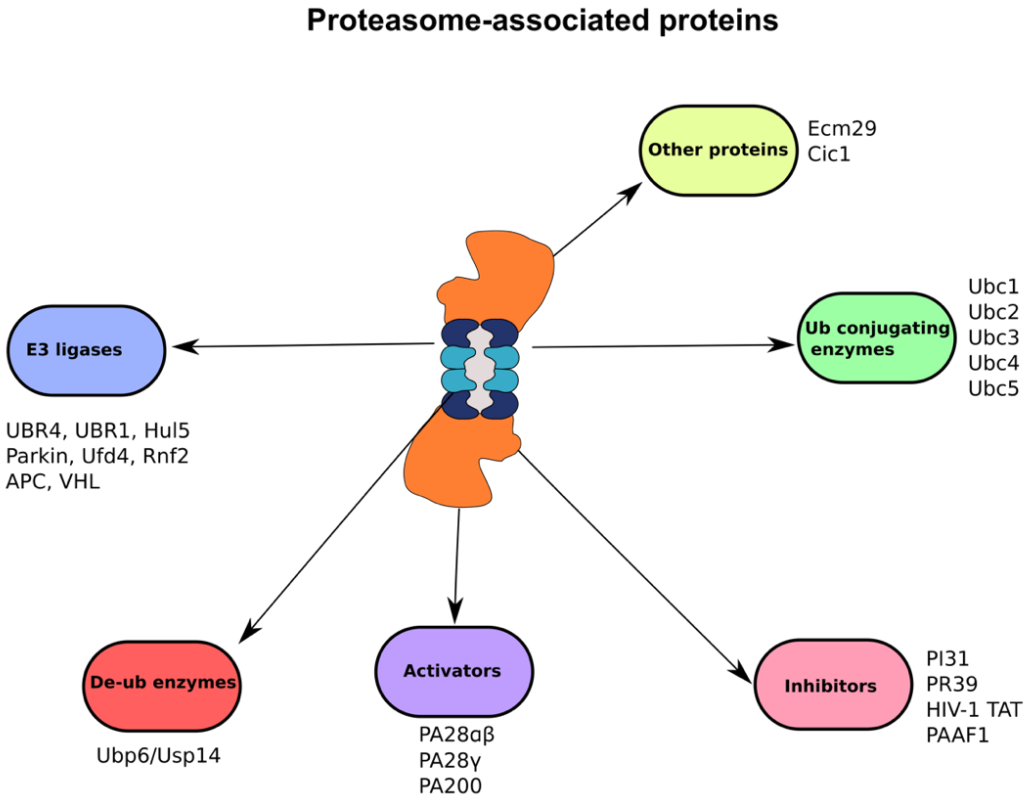


Fig. 1.9: Proteasome associated proteins. The proteasome has been shown to associate with a wide range of proteins. These vary in function and size and include proteasome inhibitors and activators (which bind to the 20S proteasome) as well as E3 ligases, and de-ubiquitinating enzymes which bind to the 19S RP.

As mentioned previously Ecm29 has also been found interacting with the proteasomes in approximately stoichiometric quantities (Leggett et al., 2002). In the study by Leggett et

al. (2002), Ecm29 was found to interact with both the RP and the CP complexes, and proteasome complexes that were purified from strains that lacked ECM29 genes were structurally less stable (Leggett et al., 2002). Within mammalian cells, Ecm29 was shown in complex with proteasomes but also binding to endosomes and ER membranes (Gorbea et al., 2004). This has prompted some to suggest that Ecm29 could be acting as an adaptor to link the proteasome to secretory membranes (Gorbea et al., 2004).

In addition, the 42.5 kDa protein, Cic1, has also been identified in complex with the 26S proteasome (Jäger et al., 2001). This study showed that Cic1 is required for the degradation of specific proteasomal substrates such as cell division control protein 4 (Cdc4) (Jäger et al., 2001). In general, there have been several proposed proteasome-interacting proteins, highlighting the complexity of the proteasome interaction network and its far-reaching effects. Several of these studies provide clues into novel pathways that involve the proteasome, but they still require further investigation.

1.4 Aim and objectives

The proteasome interacts with many transient and ancillary proteins but knowledge on the nature of these interactions and their effects on proteasome function is limited. Investigating proteasome interactors is important to better understand the proteasome system, which affects a variety of cellular processes and is a sought after drug target for diseases such as cancer.
UBR4, P97, and PI31 are commonly reported as co-eluting with proteasomes, but data has not necessarily been consistent or robust to undoubtedly prove or characterise a direct interaction. UBR4's, P97's, and PI31's role as an E3 ligase, a AAA+ ATPase, and as a proteasome inhibitor respectively, suggest regulatory roles with the proteasome.

The main focus of this thesis is to therefore deepen our understanding of the interaction of UBR4, P97 and PI31 to the human proteasome. In order to investigate and characterise these interactions directly, I optimised the 26S proteasome purification protocol and used biochemical, biophysical and cryo-EM techniques to investigate the relationship between UBR4, P97, PI31 and the human proteasome.

These were the following objectives of the projects:

- Develop and optimise a reproducible and reliable protocol for the preparation of human 26S proteasomes from HEK293F cells for biochemical and structural studies.

- Investigate whether UBR4 directly interacts with the proteasome through biochemical and biophysical methods and address the functional relevance of this interaction. Optimise and develop a novel purification method for endogenous UBR4. Depending on biochemical results, use structural biology approaches to visualise UBR4 with the proteasome.

- Investigate whether P97 binds to the human proteasome by developing an *in vivo* purification method of purifying P97 with and without the proteasome. Investigate the interaction through biochemical and biophysical tools to understand if there is a direct link of P97 to the human proteasome.

- In order to better understand the PI31-proteasome interaction, develop a reliable and reproducible purification method of purifying PI31 in isolation or in complex with the human proteasome by co-expression in insect cells. Investigate PI31 with the proteasome biochemically and biophysically and deepen our knowledge of PI31's regulatory role on proteasome activity.

- Using structural biology methods (cryo-EM), obtain a structure of the PI31-20S proteasome.

Chapter 2 Materials and Methods

2.1 26S proteasome sample preparation

2.1.1 Rad23b-UBL purification

Rad23b-UBL was used as an affinity bait to capture endogenous human 26S proteasomes from lysates (described below in Section 2.1.2). The human Rad23b-UBL construct used in the human 26S proteasome purifications was recombinantly expressed and contained an N-terminal hexahistidine tag (His-tag) and a double streptavidin-tag (Twin-Strep tag) that were separated by a Tobacco Etch Virus (TEV) protease cleavage site. This construct was already available in the lab when I joined. Rad23b-UBL was transformed into BL21 Gold *E. coli* expression strain. A streak of the resultant colonies from the transformation were used for a starter culture the night before expression using 100 ml of 2X Typtone and yeast extract (2XTY) containing 100 µg/ml of ampicillin antibiotic at 37 °C. The next day, 5 ml of starter culture was added to each 1 L of 2XTY media (total of 6 L culture) with 100 µg/ml of ampicillin antibiotic and grown to 0.6 OD_{600} at 37 °C. At 0.6 OD_{600}, cells, were induced with 1 mM Isopropyl-ß-D-1-thiogalactopyranoside (IPTG) for 3 hours at 37 °C with shaking at 180 rpm. Cells were then collected by centrifuging for 20 minutes at 6000g. Pellets of cells expressing Rad23b-UBL were then resuspended in lysis buffer containing 50 mM Tris pH 7.4, 500 mM NaCl, 1 mM dithiothreitol (DTT) and a protease inhibitor cocktail (Roche). Cells were then lysed by sonication for a total of 4.5 minutes (2 seconds on, 8 seconds off at 60% amplitude). The lysate was then centrifuged for 20 minutes at 48000g and subsequently filtered with a 5 µm filter.

After centrifugation, Rad23b-UBL was purified by loading the supernatant (lysate) in a HisTrap HP 5 ml column affinity column (GE Healthcare) that had been previously equilibrated with 50 mM Tris pH 7.4, 500 mM NaCl, 1 mM DTT, 20 mM Imidazole. Bound protein was eluted using a linear gradient from 20 to 500 mM Imidazole. Fractions containing Rad23-UBL were pooled and the his-tag cleaved by TEV protease (50:1 ratio of Rad23b-UBL: mg of TEV) and dialysed overnight at 4 °C into buffer containing 50 mM Tris pH 7.4, 100 mM NaCl, 1 mM DTT, 20 mM Imidazole. After cleavage, the protein was loaded in a pre-equilibrated His-Trap HP 5 ml column and Rad23b-UBL was

collected in the flow through. All C-terminally tagged His-TEV, uncleaved protein and other non-specific resin binding contaminants were captured in the column. The purified Rad23b-UBL collected in the flow through still contained an N-terminal Tw-Strep tag and was dialysed into 50 mM Tris pH 7.4, 50 mM NaCl, 5 mM $MgCl_2$, 10% glycerol, 0.25 mM Tris(2-carboxyethyl) phosphine (TCEP). Aliquots were subsequently flash frozen in liquid nitrogen and stored at -80 °C.

2.1.2 Purification of 26S proteasomes from HEK293F cells

Human 26S proteasomes were purified from human embryonic kidney cells (HEK293F) that had been adapted to suspension culture. FreeStyle 293 Expression Medium (Thermo Fisher Scientific) was used to grow the cells in suspension and harvested at a density of 3 million cells/ml. After harvesting, cells were centrifuged for 10 minutes at 4000g and pellets were flash frozen in liquid nitrogen and kept at -80 °C. In a typical human 26S proteasome purification, a 25 g of pellet (resulting from approximately 2 L of HEK293F suspension cells) was resuspended in 75 ml standard buffer containing 50 mM Tris pH 7.4, 150 mM NaCl, 5 mM $MgCl_2$, 10% glycerol, 5 mM ATP, 0.25 mM TCEP. If more cells were used, the same 1:3 pellet: volume of lysis buffer ratio was kept. Cells were gently broken using a dounce homogeniser (20 complete dounces), followed by a short centrifugation step of 10 minutes at 12000g. The centrifugation step allowed the cellular debris, soluble proteins, and lipids to separate as pellet, middle layer and the top layer of the supernatant respectively. The HEK293F lysate was then mixed with Rad23b-UBL at a 100:1 ratio (mg of protein in the lysate: mg of Rad23b-UBL) and left to incubate for 1 hour at 4 °C with rotation to ensure efficient mixing of lysate and Rad23b-UBL protein.

After incubation, avidin was added at a 4000:1 ratio (mg of absolute protein in lysate: mg of avidin) to reduce non-specific biotinylated proteins binding to the streptavidin resin. The first chromatography step was a streptavidin affinity chromatography using a Strep-tactin column (Qiagen) equilibrated with 50 mM Tris pH 7.4, 150 mM NaCl, 5 mM $MgCl_2$, 10% glycerol, 5 mM ATP, 0.25 mM TCEP. Proteins binding to the column through the Rad23b-UBL Twin-Strep tag were eluted with 2.5 mM d-desthiobiotin in equilibration buffer.

The second step of the purification involved glycerol gradient ultracentrifugation. The same equilibration buffer used in the first affinity chromatography step also was used in this step with the addition of glycerol. After optimisation of the gradients, the optimal

density that saw the best separation was a 15-30% glycerol gradient. Fractions containing proteasomes from the affinity chromatography (total volume of 3 ml) were loaded onto the glycerol gradient and separated on a 35 ml gradient using a Beckman SW 32 Ti rotor for 20 hours at 12,4000g at 4 °C. Fractions of 500 µl were then collected from the glycerol gradient after centrifugation and a 4-12% SDS-PAGE (Invitrogen) was run for protein evaluation. Glycerol gradient fractions that contained the 26S proteasomes were pooled and concentrated to 100 µl for a total time of approximately 45 minutes in multiple 15 minute spins at 1500g using 100 kDa Amicon Ultra (Merck) centrifugal filters.

2.2 UBR4-proteasome sample preparation

Since UBR4 was pooled down in significant amounts in human 26S proteasome preparations as described in Section 2.1.2 (approximately 0.5 mg/ml), the first steps of endogenous untagged UBR4 purification involved capturing UBR4 bound to human 26S proteasome. The purification followed the same protocol but tested buffers with and without ATP. The second step of the purification procedure was investigated further using one of the following strategies described below or with glycerol gradient centrifugation as described in Section 2.1.2.

For CRISPR-tagged UBR4 purifications, cell lines were grown in suspension as described in section 2.4.5 and the resulting cells were purified using the same purification methods described in Section 2.1.2 but without addition of Rad23b-UBL.

2.2.1 Ion-Exchange Chromatography

After the first affinity chromatography step using Strep-Tactin resin, fractions containing UBR4 as identified by SDS-PAGE were pooled together and further purified by ion exchange chromatography (IEX). A HiTrap DEAE FF 1 ml column (GE Healthcare) was equilibrated and washed in standard buffer described in Section 2.1.2. UBR4 and proteasomes were eluted with step elution. Step elution involved increasing the NaCl concentration in distinct steps (with UBR4 and proteasome eluting at a concentration of 200 mM), to a maximum NaCl concentration of 500 mM. The fractions contacting UBR4 and the proteasome were then pooled and used for further analysis and characterisation.

2.2.2 Calmodulin affinity chromatography

Calmodulin affinity chromatography was conducted as a second step, after the streptavidin affinity chromatography step described in Section 2.1.1. Fractions containing UBR4 from the first purification step were pooled and 2 mM $CaCl_2$ was added to the sample. Calmodulin beads had been equilibrated with standard calmodulin buffer containing 50 mM Tris pH 7.4, 50 mM NaCl, 2 mM $CaCl_2$, 5 mM $MgCl_2$, 1 mM ß- 2-Mercaptoethanol, 5 mM ATP. The calmodulin beads (Calmodulin-Sepharose beads from GE Healthcare) were subsequently incubated with the sample for 18 hours at 4 °C. Beads were washed with the standard calmodulin buffer and protein was eluted using the elution buffer containing the same components of standard calmodulin buffer either 3 mM or 30 mM EGTA. Fractions from the elution peak were applied on an SDS-PAGE gel and subsequently stained for detection using Pierce Silver Stain Kit (Thermo Fischer Scientific).

2.3 20S proteasome purification

Cell pellets harvested from insect cell expression of 20S proteasome as described in Section 2.6.2 were thawed and resuspended in three times volume of buffer W which consisted of 50 mM Tris pH 7.5, 150 mM NaCl, 5% (v/v) glycerol, 1 mM DTT, and 1 mM EDTA. Cells were then lysed by sonication for a total of 4.5 minutes (2 seconds on, 8 seconds off at 60% amplitude). Samples were then centrifuged at 48400 g for 30 minutes. Lysate was filtered as described in Section 2.2 and loaded onto a tandem Strep-Tactin Superflow Plus column (QIAGEN) that was equilibrated in buffer W. Protein was eluted in buffer W containing 2.5 mM d-desthiobiotin. Fractions containing proteasomes were pooled and the TwinStrep-TEV-tag was cleaved by overnight dialysis as described in Section 2.1.1. The sample was loaded onto another Strep-Tactin Superflow Plus columns as described in Section 2.2 and the flow-through was collected. Protein was concentrated with a 30 kDa cut-off Vivaspin 20 concentrator (Sartorius) and loaded onto size exclusion chromatography column (Superose 6 Increase 10/300 (GE Healthcare)) that had been equilibrated with buffer W.

2.3.1 P97-20S and P97 sample preparation

Expression and purification of P97 was conducted by co-infection of Sf9 cells with the Twin-Strep tag-TEV-p97 baculovirus, with untagged P97 baculovirus, or with P97-20S chaperones baculovirus as described in Section 2.6.1 and with 20S proteasome baculovirus. The purification protocol was the same as described for recombinant 20S proteasomes, except that for the size-exclusion chromatography steps the size exclusion chromatography column used was a Superdex 200 Increase 10/300 (GE Healthcare) instead of a Superose 6 Increase 10/300 (GE Healthcare).

2.3.2 PI31-20S and PI31 sample preparation

Expression and purification of PI31 was conducted by co-infection of Sf9 cells with the Twin-Strep tag-TEV-PI31 baculovirus, with untagged PI31 baculovirus, or with PI31-20S chaperones baculovirus as described in Section 2.6.1 and with 20S proteasome baculovirus. The purification protocol was the same described for recombinant 20S proteasomes, except that further chromatography steps were conducted (three consecutive size-exclusion chromatography in total in which the early peak (fractions 15-17) was pooled and re-injected) followed by a final analytical size-exclusion chromatography step using a Superose 6 PC 3.2/30 column (GE Healthcare) equilibrated with 50 mM Tris pH 7.4, 50 mM NaCl, 1 mM DTT, 1 mM EDTA. For individual PI31 purifications a Superdex 200 Increase 10/300 (GE Healthcare) was used instead.

2.4 CRISPR-tagging of endogenous UBR4

2.4.1 CRISPR cloning

Cas9 commercial plasmid PX458 (Addgene) was used, which contained a green fluorescent protein (GFP) marker with BbsI restriction sites. WGE CRISPR design tool (Wellcome Sanger Institute) and GPP sgRNA Designer tool (Broad Institute) packages were then used to aid in guide RNA (gRNA) design. gRNAs that had the best specificity ranking in the software packages and identified by both software programs were picked.

BbsI restriction sites and a guanine at positon -1 (shown to improve Cas9 activity) were then added to the gRNAs. gRNAs were then annealed and ligated to BbsI-digested Cas9 PX458 plasmid using 1 µl of grNA1_1, 1 µl of gRNA1_2 as shown in Table 2.1, 8 µl of H_20, heated to 95 °C and then incubated at 25 °C for 1 hour. The ligation contained 100 ng of digested vector, 2 µl of annealed oligos, 2 µl of 10X T4 ligase buffer (NEB), 1 µl of T4 ligase and H_20 to a final reaction volume of 20 µl.

For donor plasmid design, constructs were assembled using In Vivo Assembly (IVA) cloning (García-Nafría et al., 2016) with primers listed in Table 2.1 in an empty vector backbone containing a neomycin cassette (see appendix for final vector map and sequence). The plasmid was designed so that the C-terminus of UBR4 was tagged with Twin-Strep tags. PCR reactions were performed using Phusion DNA polymerase (NEB) according to the manufacture's protocol (initial denaturation step at 95 °C for 30 seconds, 30 cycles at 65 °C, and a final extension at 65 °C for 5 minutes). After PCR amplification, PCR products were treated with 1 µl FastDigest DpnI enzyme (Thermo Fisher Scientific) for 15 minutes at 37 °C. The treated PCR products were subsequently transformed in to XL10-Gold cells (Agilent Technologies) by mixing 1 µl of DNA into 50 µl competent cells, incubating at 4 °C for 30 minutes, heat shocking at 42 °C for 45 seconds, incubating at 4 °C for 2 minutes and plating onto LB agar plates. Colonies were screened for successful incorporation of the DNA fragments by restriction digestion and successful plasmids were then assessed by DNA sequencing for which the sequencing primers are provided in Table 2.2 (GATC Biotech).

Table 2.1: Primers for CRISPR cloning

Name	Primer Sequence	Details
gRNA1_1	CACCGACGATGATGGTTA CCCTTC A	Primer for gRNA oligo 1
gRNA1_2	AAACTG AAG GGTAAC CAT CATCGTC	Primer for gRNA oligo 2
PX458_F	AAACGCTTCGTCTTCGCC GCCGCAC	Forward primer for PX458 CRISPR commercial vector
PX458_R	AAACATCCAGAGAGCTTC CTGAAG C	Reverse primer for PX458 CRISPR commercial vector
Donor_plasmid_LHA_F	CCGCGGTCTAGATCCCTG TTAGGAGAACGCCTCTGC TAACATTTT	Forward primer for left homology arm for donor plasmid (CRISPR)
Donor_plasmid_LHA_R	GGGGACTGAGTTCAA CAGGTCCTT C	Reverse primer for left homology arm for donor plasmid (CRISPR)
Donor_plasmid_3XFLAG_TwStr_F	CTGTTGAACTCAGTCCCC CTGGAAGTTCTGTTCCAG GGGCC	Forward primer for cloning the tags for donor plasmid (CRISPR)
Donor_plasmid_3XFLAG_TwStr_R	UCATTATTTCTCGAATTG CGGATGGCTCCAAC	Reverse primer for cloning the tags for donor plasmid (CRISPR) Forward primer for cloning the tags for donor plasmid (CRISPR)
Donor_plasmid_RHA_F	ATCCGCAATTCGAGAAAT AATGACCACCACACAGCA GCTGCG G	Forward primer for right homology arm for donor plasmid (CRISPR)
Donor_plasmid_RHA_R	CGGGTTCCTTCCGGTAGA GAGGCCCTCGGGAGT CCT	Reverse primer for right homology arm for donor plasmid (CRISPR)
Donor_plasmid_vector_neomycin_F	TACCGGAAGGAACCCGCG CTATGA	Forward primer for the vector backbone that includes the neomycin cassette
Donor_plasmid_vector_neomycin_R	GGATCTAGACCGCGGGTA CGA C	Reverse primer for the vector backbone that includes the neomycin cassette

Table 2.2: Sequencing primers

Primer	Sequence	Use
CRISPR_UBR4_seq_F	CAGGAAACAGCTATGACC	Forward sequencing primer for tagged UBR4 CRISPR clones
CRISPR_UBR4_seq_R	ATTGCCGTCATAGCG	Reverse sequencing primer for tagged UBR4 CRISPR clones
P97_seq_F	ATGGCAAGCGGTGCAGATAG	Forward sequencing primer for tagged P97 clones.
P97_seq_R	TGGCGAATCGGACGATATGCT TCC	Reverse sequencing primer for tagged P97 clones.
PI31_seq_F	TTTCACCGCACCTATAAAAACA	Forward sequencing primer for tagged PI31 clones.
PI31_seq_F	ACGGCTGGTATGCGGATGAT	Reverse sequencing primer for tagged PI31 clones.
UBR4_screening_F	GAAGGACCTGTTGAACTCAGT CCCCCTGGAAGTTCTGTT CCA GGGGCC	Forward primer for screening tagged-CRISPR clones
UBR4_screening_R	CCGCAGCTGCTGTGTGGTGGT TATTTCTCGAATTGCGGA TGG CTCCAAC	Reverse primer for screening tagged-CRISPR clones

2.4.2 HEK293F cells: adherent and suspension cell adaption

HEK293F cells were used in the CRISPR tagging method. In order to transfect HEK293F cells with the plasmids generated as described above, HEK23F adherent cells were seeded for transfection by adding 5 ml of trypsin to cells, followed by a 5-minute incubation at 37 °C. Trypsin treatment ensured cells sufficiently detached from the wells. 15 ml of 5% FBS FreeStyle 293 Expression Medium (Thermo Fisher Scientific) was added to resuspend the cells. Cells were counted and seeded in 6 well plates at a density of 0.5 million cells/ml. At a density of 1 million cells/ml, after approximately 24 hours, 3 µl GeneJuice Transfection agent (Merck Millipore) and 1 µg of Donor plasmid and Cas9-gRNA plasmid were incubated at room temperature for 15 minutes and added to the cells drop-wise. 24 hours after transfection, 80 µg/ml of G418 antibiotic (Sigma) (calculated using a killing curve) was added to wells.

2.4.3 Cell sorting of transfected CRISPR cells

Cells that had sufficient cell growth from G418-treated wells, as described in Section 2.4.2, were sorted by fluorescence-activated cell sorting (FACS) based on incorporation of the GFP-tagged PX459 plasmid. Before FACS, each sample was prepared so that 10 million cells/ml was used after trypsin treatment as described in Section 2.4.2. Cells were subsequently filtered with 50 µm filters prior to cell sorting. Individual cells were then individually sorted into a 96 well plate. After one week, single cells that had formed colonies (cells over 50% confluent) were scaled up in larger wells and treated with G458-treated FreeStyle media as before.

Frozen aliquots of cell lines, grown after cell sorting were made by adding 10% dimethylsulfoxide (DMSO) to FreeStyle media (called freezing media) and then trypsin treatment was performed as described before in Section 2.4.1. Trypsin-treated cells were then centrifuged at 335g for 5 minutes and resuspended in 3 ml of freezing media. 1 ml aliquots of the resuspended cells were frozen at -80 °C for 48 hours then transferred in a -180 °C liquid nitrogen container for long-term storage.

2.4.4 Screening of CRISPR colonies

Successful adherent cells that had grown into large colonies after FACS were harvested by trypsin treatment and screened using PCR reactions with Phusion DNA polymerase (NEB) as described in section 2.4.1 The screening primers, shown in Table 2.2, were used to amplify the tag and part of the UBR4 insert sequence as shown in Table 2.1.

2.4.5 HEK293F suspension cell adaption

Adherent HEK293F cells were adapted to suspension to allow for large scale growth of CRISPR-tagged cell lines. Cells were initially placed in 5% FBS and FreeStyle 293 Expression Medium (Thermo Fisher Scientific). After 48 hours, cells were passaged at 0.5 million cells/ml and placed in 2% FBS Free Style 293 Expression Medium (Thermo Fisher Scientific). Cells were continuously split every 48 hours and all FBS was then removed in

the third split. Frozen aliquots of the suspended cell lines were prepared as previously described before in Section 2.4.3.

2.5 Proteasome Activity assays of samples- fluorescence spectroscopy

All samples obtained by the methods described above were tested for 20S proteasome proteolytic activity using fluorescence spectroscopy. This activity was tested using 7-amino-4-methylcoumarin (AMC) tagged fluorogenic peptides with sequences specific for each one of the three proteasome proteolytic activities: Z-L-L-E-AMC (caspase-like activity), Suc-L-L-V-Y-AMC (chymotrypsin-like activity), and Boc-L-R-R-AMC (trypsin-like activity) (Boston Biochem). In this reaction, 5 µl of samples were measured in buffer containing 50 mM Tris pH 7.4, 5 mM $MgCl_2$ and 1 mM DTT was incubated with 50 µM AMC-peptide. Once substrates and samples had been mixed, the proteolytic reaction was incubated for 30 minutes at room temperature followed by subsequent fluorescence measurement using a PHERAstar FS plate reader (BMG Labtech), at an excitation wavelength of 350 nm and an emission wavelength of 450 nm. Three concentration of chymotrypsin (Sigma) (10 ng/ml, 100 ng/ml, and 1 µg/ml) were used as positive controls, and compared with the sample to be measured. For that, chymotrypsin dilutions were incubated with 50 µM SUC-L-L-V-Y-AMC in buffer containing 50 mM Tris pH 7.4, 5 mM $MgCl_2$ and 1 mM DTT. Gain was corrected with the 1 µg/ml chymotrypsin dilution and fluorescence subtracted with a sample that only contained buffer. Activity assays were plotted using Prism v.8.2.0 program (GraphPad Software).

2.5.1 Titration assays of PI31 and 20S proteasome samples

PI31 titration assays were conducted by mixing 30 nM of recombinant 20S proteasomes with increasing concentrations of recombinant PI31 at concentrations of (0 nM, 30 nM, 50 nM, 70 nM, 90 nM, 150 nM, 300 nM) and a total volume of 55 µl. After mixing, samples were incubated for 30 minutes at 25 °C. Following initial incubations, samples were diluted and further incubated with 50 µM of each of the three proteasome fluorogenic substrates to a final volume of 100 µl as described above in Section 2.5 (resulting in a final concentration that correspond to a 20-fold dilution). Florescence was measured as

described in Section 2.5. Activity assays were plotted using Prism v.8.2.0 program (GraphPad Software).

2.6 Insect cell recombinant protein expression

The baculovirus expression system exploits the large family of viruses that infect arthropods, mainly insects. Baculoviruses have relatively large, circular double-stranded DNA genomes. These viruses have the ability to produce particles, termed “polyhedra” or “occlusion bodies”, at the end of their viral cycle and can accumulate within insect cells (Jarvis, 2009). Polyhedra are expressed within nuclei of the host organism and are contained within a paracrystalline array of polyhedrin. Baculoviruses produce large amounts of polyhedrin allowing high quantities of polyhedra packing into host cell nuclei (Jarvis, 2009). The ability of baculoviruses to produce excess polyhedrin is extremely useful and it is this system that is manipulated in many molecular biology techniques. Since baculovirus replication in insect cells is not dependent on the production of polyhedron, the polyhedron viral DNA sequence can be replaced with a foreign DNA sequence encoding the protein of interest (Fig. 2.1).

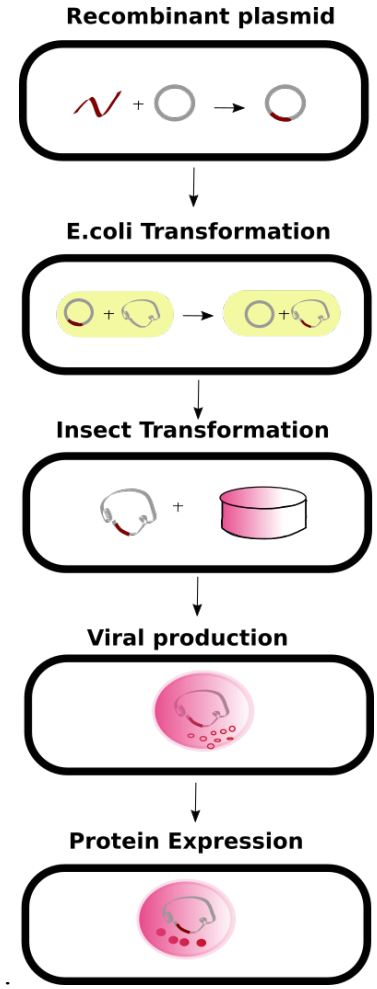


Fig. 2.1: Insect cell expression strategy. A recombinant plasmid is cloned with the gene of interest (red line). The next step involves transforming the recombinant plasmid into *E. coli* competent cells that contain baculoviral DNA (yellow) which then leads to the production of a recombinant bacmid. Bacmids are then transfected into Sf9 insect cells (pink) to make baculoviruses which are subsequently used to infect fresh Sf9 cells and allow for viral multiplication and further protein expression.

Using this technique, Dr. Ana Toste Rêgo from the da Fonseca lab had developed a system, based on the multi-baculovirus expression system, in which all 20S proteasome subunits can be over-expressed in insect cells. This involves co-infecting insect cells with two baculovirus, one expressing all 20S subunits and the other one expressing the 20S chaperones, PAC1, PAC2, PAC3, PAC4 and POMP (Hirano et al., 2006). Because of the need to express multiple genes within the same vector, the MultiBac system was used (Bieniossek et al., 2012). The MultiBac system is an advanced platform tailored for eukaryotic multiprotein complex production. Once the vectors were prepared, these were used to make baculoviruses, that have been engineered to increase protein production in

insect cells. The viruses, which are generated by bacmids, are then used to infect insect cells, which in this case were SF9 cells. Cells were then harvested and proteins were purified (Fig. 2.1).

2.6.1 Cloning and generation of baculoviruses

The 20S baculovirus contained a Twin-Strep tag at the C-terminus of ß7 so mature, homogenous proteasome complexes could be purified by pulling down on this subunit (Rêgo and Fonseca, 2019). PI31 and P97 gene sequences were synthetically synthesised using codon-optimised sequences (for P97 initial P97 constructs were made using the cDNA plasmid 113510 from Addgene) and expressed in *E. coli* (Epoch) cells. PI31 and P97 genes were then subsequently cloned into Multibac Turbo vectors pACEBac1, pIDS and/or pIDC using the IVA cloning method (García-Nafría et al., 2016). P97 was cloned with a C-terminal Twin-Strep tag followed by a TEV protease cleavage site. PI31 was cloned with a C-terminal Twin-Step tag followed by a TEV protease cleavage site. Untagged versions of PI31 and P97 were also constructed as well as untagged P97-20S chaperones and untagged P97-20S chaperones constructs that had been cre-loxed to include in the single vector pACEBac vector (Fitzgerald et al., 2006). In order to produce PI31-20S chaperones and P97-20S chaperones constructs, the vector containing all 20S chaperones and the PI31 of P97 vectors were combined recombinantly via their LoxP sites, making use of the Cre recombinase enzyme that recognises these sequences. Cre-lox recombination was conducted by mixing 350 ng/µl of 20S chaperones in the pACEBac1 vector, 100 ng of the P97 or PI31 vector in pIDC vector, 1 µl Cre-recombinase enzyme (New England BioLabs) and 2 µl of Cre-buffer (New England BioLabs) in a total 20 µl reaction. The reaction was incubated for 1 hour at 37 °C. Untagged constructs were generated in order to be co-expressed with the 20S proteasome. 10 ng/µl of plasmids were then transformed into 100 µl DH10 EmBacY cells (Geneva Biotech), left on ice for 30 minutes at 4 °C, heat shocked at 42 °C for 45 seconds, incubated at 37 °C for 4 hours with 400 µl of SOB media and plated on selective agar plates containing chloramphenicol and gentamicin. Bacmid DNA was purified using the published protocol by O'Reilly et al (1994) and the QIAGENprep Spin Miniprep kit (Qiagen). Bacmids were transfected into Sf9 cells at 1.5 x 10^6 cells/ml in InsectXPRESS media (Lonza) (P1 virus) by adding a mixture of 6 µl Fugene reagent (ThermoFisher) and 2 µg of Bacmid in a total reaction volume of 200 µl, adding it to a well in a 6 well plate and incubating for 72 hours until

viability had decreased to around 80%. The viability was measured using a Countess Automated cell counter (ThermoFischer). This made the P2 virus. The established P2 virus was spun at 1000g, filtered and the supernatant was placed in 2% Hi-FBS serum for long term storage at 4 °C and covered in foil to prevent light-induced damage. All primers used in this experimental design are shown in Table 2.3.

2.6.2 Expression of PI31, P97, PI31-20S-20S chaperones, and P97-20S-20S chaperones

Each P2 baculovirus was individually amplified by infecting Sf9 cells at a density of 1.5 x 10^6 cells/ml (at a ratio of 1: 100 of the P2 virus to cell volume in P3 virus) for 72 hours until viability had dropped to 80%. This P3 virus was filtered and used to infect Sf9 cells at approximately 2 x 10^6 cells/ml density (12 ml of each P3 virus was used per 500 ml of Sf9 cells). For co-expression, more than one P3 virus was infected per 500 ml of diluted Sf9 cells (12 ml of each virus used). Initial experiments involved triple infections with 20S proteasome baculovirus, 20S chaperones baculovirus and P97 or PI31 baculovirus. For double infections, 20S proteasome baculovirus and either PI31-20S chaperones or P97-20S chaperones baculoviruses were used. After infection, cells were incubated at 27 °C for 48 hours and harvested by centrifuging at 3000g for 20 minutes at a temperature of 4 °C. After centrifuging, cell pellets were washed with 35 ml of cold phosphate-buffered saline (PBS) solution, flash frozen in liquid nitrogen and stored at -80 °C.

Table 2.3: Primers for baculovirus cloning

	Sequence	Use
P97_pACEBac_F	GGATCCCGGTCCGAA ATG GCT TCT GGAGCC GATTCAAAAGGTG	Forward primer for cloning of P97 in pACEBac (no tag).
P97_pACEBac_R	ACTGCAGGCTCTAGATTC TTA GCC ATACAGGTCATCATCATTGTCTTCTG	Reverse primer for cloning P97 in pACEBac (no tag).
TwStrP97_pACEBac_F	G ATC CCG GTC CGA A GAA AAC CTG TATTTTCAGTCAATGGCTAGCGCA TGGAGTCATCCTC	Forward primer for cloning of P97 in pACEBac with Twin-strep tag.
TwStrP97_pACEBac_F	AC CTT TTG AAT CGG CTC CAG AAG CCATATGGGATTTTCGAACTGCGG GTG GC	Reverse primer for cloning of P97 in pACEBac with Twin-Strep tag.
PI31_paceidc_F	GGATCC CGG TCC GAA ATG GCA GGTCTGGAAGTTCTGTTTGC	Forward primer for cloning of P97 in pACEBac and pIDC.
PI31_paceidc_R	ACTGCAGGCTCTAGATTCTTACAG ATACATATCATCATAACCTGGCGGAG	Reverse primer for cloning of P97 in pACEBac and pIDC (no tag).
NTevTwStr_F	ATGTCGTACTACTGGAGCCATCCG	Forward primer for cloning the Twin-Strep tag in pACEBac vector.
NTevTwStr_R	GGCGCCCTGAAAATACAGGTTTTCCTT	Reverse primer for cloning the Twin-Strep tag in pACEBac vector.
TwStrPI31_paceidc_F	CTGTATTTTCAGGGCGCC ATG GCA GGT CTG GAA GTT CTG TTT GC	Forward primer for cloning of PI31 in pACEBac and pIDC with N-terminal Twin Strep- tag.
TwStrPI31_paceidc_R	GCTCCAGTAGTACGACAT TTC GGA CCGGGATCCGCGC	Reverse primer for cloning of PI31 in pACEBac and pIDC with N-terminal Twin-Strep tag.

2.7 Western blot analyses

2.7.1 Western blots to validate ubiquitination

Samples containing UBR4-19S and tagged UBR4 as shown in Chapter three (and purified as described in Section 2.2 and 2.4) were checked for ubiquitination states using western blotting (WB) (Burnette, 1981). 25 µl of samples at approximately 0.2 mg/ml were first analysed by SDS-PAGE and transferred using an XCell II Bolt Module (Novex) to a PVDF membrane (Sigma Aldrich). The PVDF was then soaked in methanol for 30 seconds and placed on the transfer module. The SDS-PAGE was immediately placed on the membrane whilst it was still wet and filled with transfer buffer consisting of 25 mM Tris pH 7.4, 190 mM glycine, and 10% methanol. The traditional sandwich model set up of the XCell was used and proteins from the gel were transferred to the membrane at 30V for 1 hour. After transfer, the membrane was blocked in 5% (w/v) powdered milk in PBS and 0.1% Tween-20 (PBST) for 1 hour at room temperature (RT) and incubated with primary anti-UB antibody UBI-1 (Abcam) at a 1:2000 dilution in 5% (w/v) milk in PBST and for 1 hour at room temperature. Following incubation, the membrane was washed three times with PBST for 10 minutes and incubated at a 1:10000 dilution for 1 hour at room temperature with secondary antibody- horseradish peroxidase (HRP)-conjugated anti-mouse IgG antibody (GE Healthcare) in 5% (w/v) milk in PBST buffer. The final wash was conducted with 3 rounds of PBST buffer. The Amersham ECL Western Blotting Detection Kit (GE Healthcare) was used to detect the HRP. The manufacture's protocol was used for detection. Briefly, a 5-minute incubation of a stable peroxide solution was mixed with an enhanced luminol solution on the membrane. After this, film exposed on the membrane (of approximately 2 minutes) was used to visualise the UBR4-19S samples and the level of ubiquitination present within each.

2.7.2 UBI-CAPTURE experiment

UBI-capture experiment using the UBI-QAPTURE-Q kit (Enzo Life Sciences) was conducted as described in Benvegnu et al. (2017) (Fig. 2.2). Control sample (Control ubiquinated-protein lysate from Enzo Life Sciences) provided by the kit was used to

validate the method. Other samples used were UBR4-19S and tagged-UBR4 proteins as mentioned in Chapter three. Samples were diluted 20 times with PBS to a final concentration of 25 µg of total protein and incubated with UBI-QAPTURE-Q matrix (Enzo Life Sciences) for 4 hours at 4 °C to ensure efficient matrix binding. The matrix was washed with the standard buffer described in Section 2.2. Samples were then centrifuged for 30 seconds at 5000g to collect the UBI-QAPTURE-Q matrix while the supernatant was collected as 'unbound fractions' UBI-QAPTURE-Q matrix. Three PBS washes were then performed and the sample was subsequently eluted with 100 µl PBS. An aliquot of the sample was added to 25 µl of 5x SDS-PAGE gel loading buffer while the rest was saved for analysis by western blotting (Fig. 2.2). Western blot analysis was conducted in the same way as described in Section 2.7.1 with mono-and poly-ubiquitinated conjugated monoclonal antibody (FK2) (HRP conjugate) used as the primary antibody at a dilution of 1:1000.

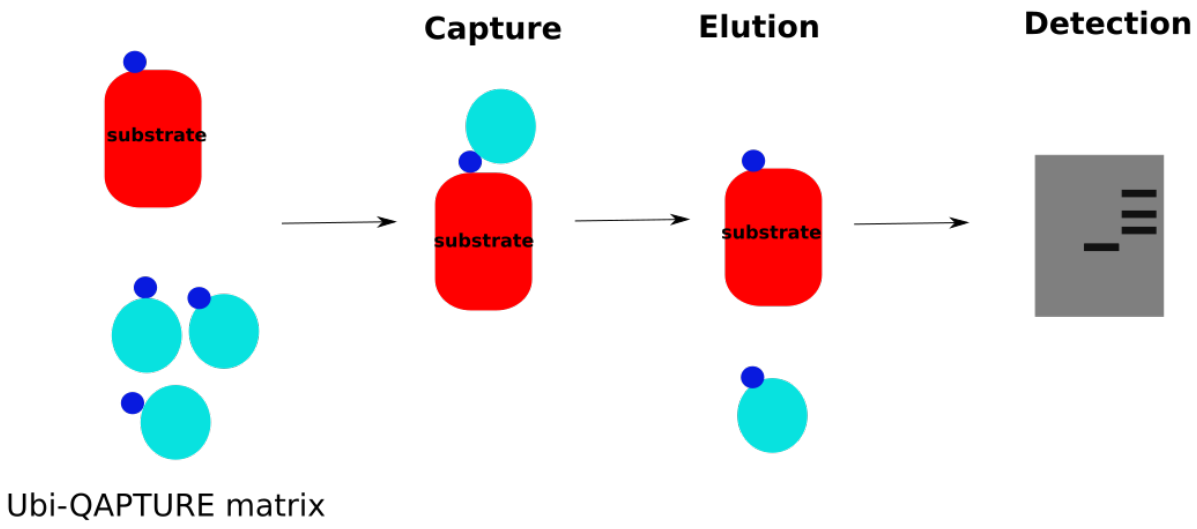


Fig. 2.2 The UBI-capture experiment. The UBI-capture tool offers a way in which ubiquitinated proteins can be detected in either a lysate or as purified sample. Substrates that are ubquitinated (blue) are captured by the high affinity matrix that can detect ubiqutinated proteins. Proteins are then washed, centrifuged and eluted and can be detected by either SDS-PAGE or western blotting.

2.8 Biophysical analyses

2.8.1 ISCAT

PI31-20S samples were purified as described in Section 2.3.2 and a 1: 20 dilution of the sample (resulting in a final concentration of 180 nM) were analysed by interformetric scattering (iSCAT) in the Refeyn One (Refeyn Ltd) as described by Young *et al.* (2018) (Young et al., 2018). 10-18 µL of each sample was applied to silicone gaskets mounted on borosilicate microscope coverslips, previously cleaned in Milli-Q water, isopropanol and Milli-Q water again, followed by drying under a stream of argon. The image was focused onto the glass surface and data collected for between 30-60s at a frame rate close to 1 kHz. For image processing, the raw image was processed by 4 x 4 pixel binning and time averaged 10-fold to give an effective frame rate of 100 Hz. To remove the static scattering background, images were processed in a ratiometric analysis to reveal changes between different frame batches. Protein landings on the surface were identified automatically as a dark point spread function that increases and decreases through successive frames. The peak contrast was used to determine the mass of the particle from a linear fit of contrast against mass of a set of protein standards. Histograms of single particle landings were generated with a binning of 5 kDa.

2.8.2 SEC-MALS

Size exclusion chromatography coupled with multiple angle scattering (SEC-MALS) was conducted using a Wyatt MALS system (Wyatt Technologies). 100 µL of samples (UBR4-19S, P97 and PI31-20S, PI31 and P97) were injected at a concentration of 0.1 mg/ml into a Superdex 200 Increase HR 10/300 (GE Healthcare) at 0.5 ml/min in standard buffer as described in Section 2.2 and 2.3. The protein concentration and light scattering at each point across the chromatographic peaks was used to determine the absolute molecular mass from the intercept of the Debye plot using Zimm's model as implemented in the ASTRA v5.3.4.20 software (Wyatt Technologies). To determine interdetector delay volumes, band-broadening constants and detector intensity normalization constants for the instrument, Bovine Serum Albumin (BSA) was used as a standard prior-to sample measurement. The data collected were plotted with the Prism v.8.2.0 program (GraphPad Software).

2.9 Binding experiments

2.9.1 PI31 and 20S binding experiment

PI31 and 20S binding experiments were conducted by individually purifying 20S proteasome or PI31 from insect cells (as described in Section 2.3.2 and 2.6) followed by mixing 1.5 μM of PI31 with 0.15 μM of 20S proteasome to ensure sufficient saturation of 20S with PI31. The samples were left at room temperature for 1 hour before concentrating the mix using 30 KDa cut-off concentrators as described previously to a final concentration of 1.2 mg/ml. The concentrated sample was then injected into the size exclusion chromatography column Superose 6 Increase 10/300 (GE Healthcare). Results were analysed using proteasome activity assays as explained in Section 2.5 and analysed using Prism v.8.2.0 program (GraphPad Software).

2.9.2 Flag-PI31 and 20S purifications

Samples were co-expressed in baculoviruses as described in Section 2.6. Proteins purified from streptavidin affinity chromatography as described in Section 2.2 were directly incubated with Anti-Flag M2 agarose beads (Sigma-Aldrich). The resin was equilibrated with 5 column volumes (CV) of standard buffer but with no glycerol at room temperature. This was followed by ten subsequent washes with 2 CVs of buffer. Samples were eluted with 1 CV of 200 μg/ml Flag peptide (Sigma Aldrich) in standard buffer. Resin was regenerated with 3 CV of 0.1 M glycine-HCl pH 3.5, and washed and stored in 50% glycine-HCl pH 3.5, TBS, 0.02% Sodium Azide.

2.10 Mass spectrometry

Mass spectrometry was conducted by the Mass Spectrometry Facility at the MRC-LMB. Briefly, gel samples were destained with 50 mM ammonium bicarbonate and 50% v/v acetonitrile, reduced with 10 mM DTT, and alkylated with 55 mM iodoacetamide. Samples were digested with 6 ng/μl trypsin (Promega, UK) overnight at 37 °C. Peptides were then extracted in 2% v/v acetronitrile, 2% v/v formic acid and analysed by nano-scale capillary LC-MS/MS (Ultimate U3000 HPLC, Thermo Fischer Scientific Dionex) at

approximately 300 nL/min flow rate. The UniProt KB was used to search for the LC-MS/MS data while the the Scaffold programme (Proteome Software Inc) was used to validate it.

2.11 Negative stain electron microscopy grid preparation

Quantifoil R1.2/1.3 400 mesh holey carbon copper (for UBR4 and P97 samples) or gold grids (for PI31 samples) (Quantifoil Micro Tools), coated with an additional thin carbon layer, were used for negative stain EM. 2 µl of each sample was applied on a grid that had been previously glow discharged at 30 mA for 30 seconds. The sample was left on the grid for one minute and then washed by placing it on the surface of two water droplets. The grid was then floated on drops of 2% uranyl acetate solution and incubated for 30 seconds before blotting with Whatmann filter paper. The grids were left to dry and subsequently viewed using a FEI T12 transmission electron microscope (FEI).

2.12 Cryo-electron microscopy sample preparation and data collection

Samples were applied on R1.2/2.3 400-mesh holey copper or gold grids (Quantifoil) previously coated with a thin layer of carbon prepared in house. 2 µl of samples were applied to the grids subsequently plunge-frozen into liquid ethane using a FEI Vitrobot (Thermo Fischer Scientific), at 22 °C and 100% humidity, blot time 2 seconds, wait time 2 seconds, drain time 1 second, blot force, 10, blot time, 1 second. The cryo-grids were stored in liquid nitrogen.

Cryo-EM data for the PI31- 20S proteasome sample was collected in house at the MRC-LMB on Titan Krios electron microscope (FEI) operating at 300 kV and with a FalconIII detector operating in counting mode. Electron counting modes allows single electron events to be detected in the electron, which improves the signal to noise ratio of the images (Mendez et al., 2019). Images were acquired at 96000 X magnification, pixel size of 0.81, with a total electron dose of 40 $e^-/Å^2$ and a set of defocus range from -1.8 to -3.0 µm. Movies were fractioned into 75 frames and automatic data acquisition was collected with EPU software (FEI, Thermo Fischer Scientific). Data was collected in three 48 hour

separate sessions, with the resultant micrographs that were used for structure presentation in this thesis totalling 5517.

Chapter 3 UBR4 and the 26S Proteasome

3.1 General features of E3 ligases

E3 ligases play a vital role within the UPS. As previously mentioned in Chapter one, E3 ligases are involved in the final step of the ubiquitination cascade where they catalyse the transfer of a ubiquitin molecule from an E2 enzyme to form a covalent bond with a substrate's lysine (Berndsen and Wolberger, 2014). E3 ligases confer specificity by selectivity binding to substrates and E2 enzymes (Berndsen and Wolberger, 2014). There are three distinct classes of E3 ligases: 'Really Interesting New Gene' (RING) E3 ligases, Ring-between-ring (RBR) E3 ligases, and homologous with E6-associated protein C-terminus (HECT) E3 ligases (Fig. 3.1) (Metzger et al., 2014; Rotin and Kumar, 2009; Sluimer and Distel, 2018). The different E3 ligases either directly or indirectly transfer ubiquitin (Ub) and ubiquitin-like (UBL) proteins. Their mechanisms have been elucidated to highlight how a range of E3 ligases (E3s) can bind to different substrates and E2 enzymes.

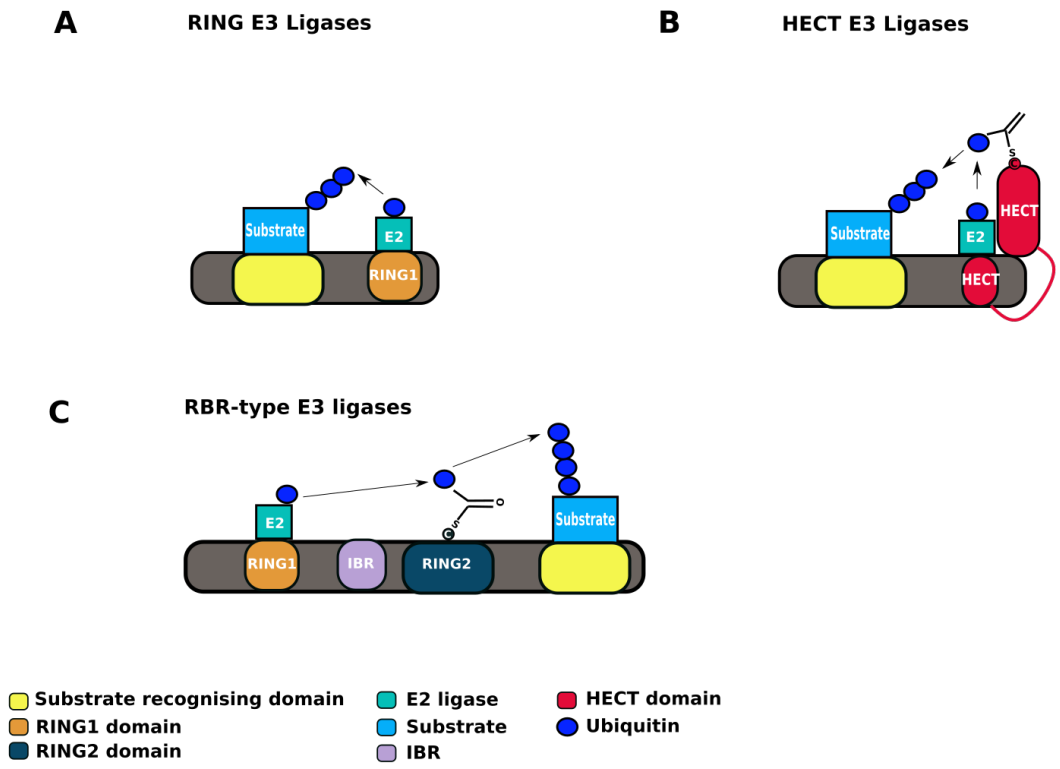


Fig 3.1: RING- HECT- and RBR-type E3 ligases and a schematic representation of their mechanism of action. The three classes of E3 ubiquitin ligases are RING-, HECT-, and RBR-type E3 ubiquitin ligases. A) RING E3 ligases bind both the substrate that will be ubiquitinated and the E2 ligase. This allows the E2 enzyme and the substrate to be in close proximity so that the ubiquitin can be transferred from the E2 enzyme to the substrate. B) HECT E3 ligases initially transfer ubiquitin from a thioester then to the substrate that will be ubiquitinated. The HECT N-terminal domain binds to an E2-ubiquitin thioester (red circle), whereas the C-terminal consists of the catalytic cysteine that is involved in ubiquitin transfer (shown as red ellipse). C) RBR-type E3 ligases consist of two RING domains that are separated by an in-between-RING domain (IBR). Ubiquitin transfer begins from the E2 bound to the first RING domain and are transferred to the second RING domain. RING2 then transfers ubiquitin through a thioester bond to the substrate. The substrate recognising domain (yellow), RING1 domain (orange), RING2 domain (grey-blue), E2 ligase (green), substrate (turquoise), IBR domain (lilac), HECT domain (red), and ubiquitin (blue circles) are displayed.

3.1.1 RING E3 ligases

RING E3 ligases use a direct mechanism to bind both the donor E2-Ub thioester and transfer the ubiquitin from the E2 enzyme to a lysine of the bound substrate (Fig. 3.1). The RING domain was first described in 1991 and comprises one of the largest enzyme families in humans with over 600 known members (Freemont et al., 1991). At the time of

discovery, 27 other proteins had been identified with the RING motives and it was initially thought that the domain may mediate DNA binding (Freemont et al., 1991). However, experiments showing that the RING protein, Rad18 which promotes ubiquitination of histones, provided a link between the RING domain and ubiquitination (Bailly et al., 1997). This finding was shortly followed by the discovery of Hypoxia responsive domain-1 (Hrd1) RING E3 ligase, required for endoplasmic reticulum-associated degradation (ERAD), the UBR1 proteins in mice and yeast which mediate the N-end rule degradation pathway, the anaphase-promoting complex Apc11 subunit, and Prt1protein in *Arabidopsis thaliana* (Bordallo et al., 1998; Kwon et al., 1998; Potuschak et al., 1998; Zachariae et al., 1998). The N-end rule pathway is a mechanism in which destabilising N-terminal residues, such as asparagine and glutamine are recognised by proteins, ubiquitinated by E3 ligases and degraded; this pathway exists in all eukaryotes (Bartel et al., 1990; Varshavsky, 2011). The ubiquitin ligase N-recognin (UBR) family of E3 ligases are discussed further in this chapter.

Despite the link between the RING domain and ubiquitination, it was not until 1999 that it was unambiguously shown that the RING domain binds ubiquitin-conjugating enzymes and is involved in the direct transfer of ubiquitin to the substrate (Kamura et al., 1999; Ohta et al., 1999; Seol et al., 1999; Skowyra et al., 1997; Tan et al., 1999). In summary, the ground-breaking 1999 papers showed the following: that RING E3s coordinate zinc and this is essential for its ubiquitin ligase activity both *in vitro* and *in vivo*; that the ubiquitination does not involve E3-linked ubiquitin intermediates, as seen in HECT E3s; and that E3 ligase activity and E2 binding are intrinsic to the RING domain (Lorick et al., 1999). Later work showed that UBR1 is essential for degrading N-end rule substrates, further discussed in Section 3.3 (Xie and Varshavsky, 1999). Several other studies linking the RING domain with ubiquitination and thus expanding the knowledge and cellular use of E3 ligases later followed (Xie and Varshavsky, 1999).

Of the 600 human RING proteins, at least half have been shown to have ubiquitin ligase function, with many others not yet been investigated (Berndsen and Wolberger, 2014; Deshaies and Joazeiro, 2009; Li et al., 2008). However, some proteins containing RING domains do not possess E3 activity. Examples include MdmX and Bmil RING proteins that are shown to function in p53 and histone H2A ubiquitination, respectively, but need other RING proteins (Mdm2 and Ring1b respectively to compensate E3 activity) (Linares

et al., 2003; Wang et al., 2004). The general features of RING domains are the common, canonical basic sequence Cys-X_2-Cys-X_{9-39}-Cys-X_{1-3}-His-X_{2-3}-Cys/His-X_2-Cys-X_{4-48}-Cys-X_2-Cys where X denotes any amino acid (Freemont et al., 1991). Within this sequence, the conserved cysteine and histidine residues are buried within the domain's core and which maintain the overall structure of the RING domain and coordinate two zinc atoms (Fig. 3.2A) (Barlow et al., 1994; Borden et al., 1995). Other RING variants have been discovered in which some cysteines and histidines are replaced by other residues that can carry out zinc coordination, such as aspartic acid in the Roc1 protein in the Skp1, Cullin1, F-box protein (SCF) ubiquitin ligase complex (Zheng et al., 2002).

The key mechanistic aspect of the RING E3 ligases is their ability to promote ubiquitin ligase activity by binding to both ubiquitin-conjugating enzymes and substrates. This was first exemplified by the Casitas B-lineage Lymphoma (Cbl) RING domain bound to the ubiquitin conjugating enzyme UbcH7 (Zheng et al., 2000). Further structural studies, including X-ray data and NMR analyses, have shown that there are key RING and E2 residues that play a crucial role in the RING domain-E2 interaction (Brzovic et al., 2003; Dominguez et al., 2004). RING domains bind to the N-terminal helix of E2, common to all E2s, and this interaction is transient if the E2 is not ubiquitin bound (Berndsen and Wolberger, 2014). Once both the substrate and E2 ligase have bound the E3 (in which the order does not matter), ubiquitin is transferred from the E2 to the substrate resulting in free E2 enzyme and ubiquitinated protein substrate (Deshaies and Joazeiro, 2009). Sequential addition of ubiquitin requires another E2-ubiquitin conjugated enzyme to bind (Fig. 3.2B). The close proximity of the E2 and the substrate is what drives E3 activity. However, more active roles for E3s have also been proposed, for example, inducing conformational changes in the E2-ubqiutination enzyme that could promote its transfer and subsequent ubiquitin transfer to the substrate (Deshaies and Joazeiro, 2009; Seol et al., 1999; Skowyra et al., 1999).

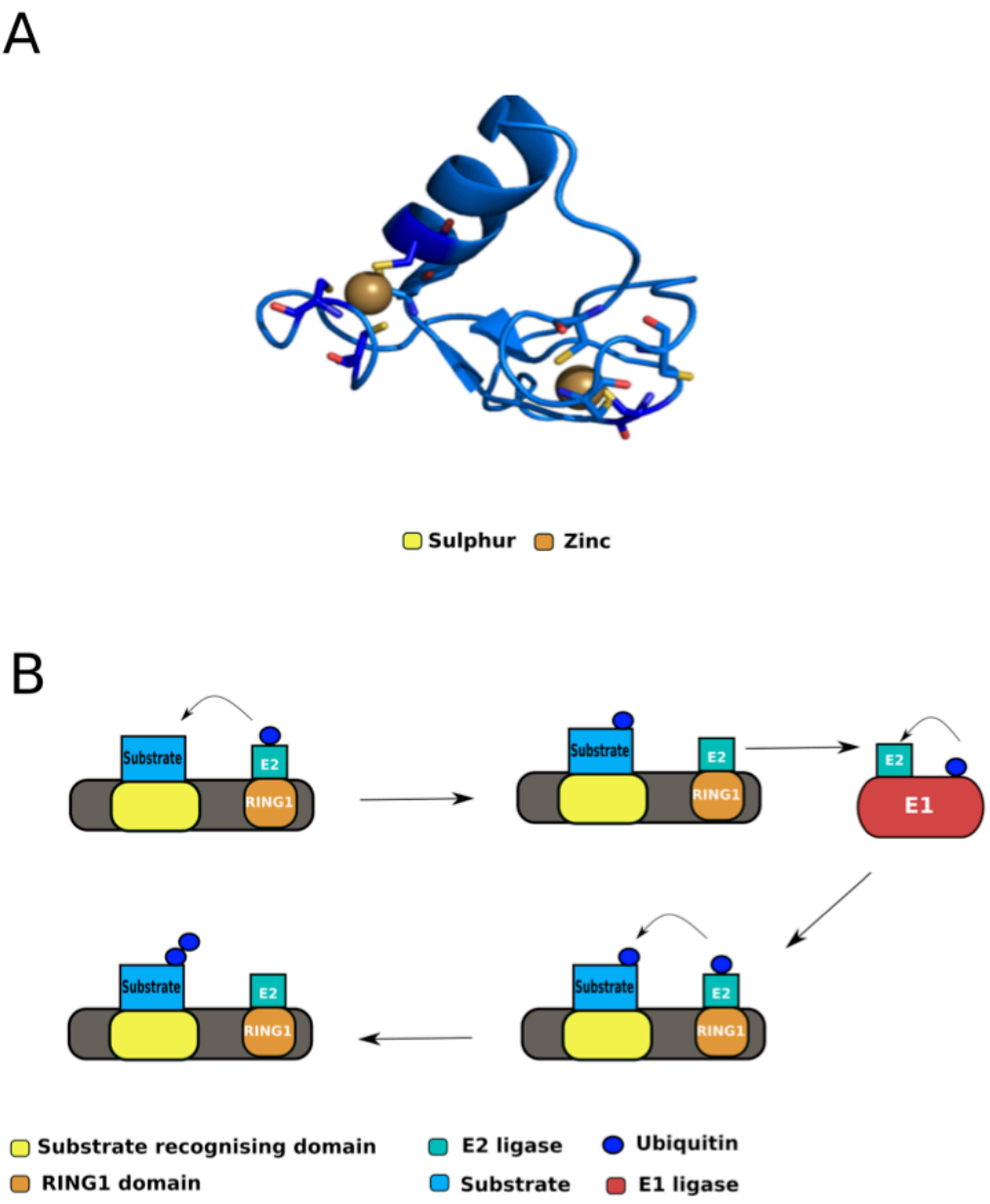


Fig 3.2: The RING E3 ligase domain and its mechanism for sequential ubiquitination. A) The RING E3 ligase domain consists of conserved cysteine residues that coordinate zinc atoms and help to maintain the overall structure of the RING domain (PDB: 1FBV). B) Substrates can be sequentially ubiquitinated by RING E3 ligases. Ubiquitin-charged E2 ligases are activated by E3 ligases and transfer the ubiquitin to the substrate protein (cyan). The E2 ligase (green) then disassociates and recharges with another ubiquitin from E1 ligases (red). The ubiquitin-charged E2 ligase can then bind to the E3 ligase and transfer an additional ubiquitin to the substrate. A key is provided showing the different proteins and domains.

3.1.2 HECT E3 ligases

HECT E3s were first discovered in 1995 and were the first E3 ligase family reported (Huibregtse et al., 1995). In humans, 28 HECT E3s have been identified (Scheffner and Staub, 2007). HECT E3s consist of a HECT domain at the C-terminus composed of approximately 350 amino acids. The N-terminal lobe of the HECT domain binds to an E2-ubiquitin thioester, whereas the C-terminal lobe contains the active site cysteine and accepts the ubiquitin from the E2 (Fig. 3.1B) (Huibregtse et al., 1995). The substrate specificity is thus determined by the N-terminal lobe of the HECT domain, whereas the C-terminal lobe of the HECT-domain possesses the catalytic activity (Fig. 3.1B). There are three subgroups of HECT E3s: Nedd4/Nedd4-like E3s, HERC family, and other HECTS in which all members are listed in Table 3.1.

Table 3.1 HECT E3 ligases family

HECT sub-family	Family member examples	Number of family members
NEDD4	NEDD4L, NEDD4L, ITCH, SMURF1, SMURF2, WWP1, WWP2, NEDL1, NEDDL2	Nine
HERC family	Small and large HERCS	Six
Other HECTs	HUWE1, HACE1, E6AP, TRIP12, UBR5	Thirteen

The exact mechanism of action of HECT E3s involves two steps (Fig. 3.1B). First, the transfer of the ubiquitin from the E2 to a cysteine in the E3 HECT domain via a transthioesterification reaction (Huibregtse et al., 1995). Second, a substrate lysine attacks the HECT-Ub thioester resulting in transfer of the ubiquitin to the protein substrate (Huibregtse et al., 1995). The C-lobe of the HECT domain has the ability to dramatically extend so the thioester can be brought in close proximity to the substrate lysine for the final ubiquitin transfer step (Kamadurai et al., 2013).

3.1.3 RBR E3 ligases

RBR E3s are mechanistically distinct and share both RING and HECT E3 ligase features. They were first identified in 1999 by the presence of two RING domains as well as an in-between RING domain (IBR domain), forming another subclass of E3 ligases (Aguilera et al., 2000; Reijden et al., 1999). Like other RING domains, the RBR domains can be found anywhere within proteins and show typical conserved cysteine and histidine residues as described earlier for RING E3 ligases (Eisenhaber et al., 2007). There are currently thirteen known RBR E3s in humans, however, they are the least understood subclass of E3 ligases with poorly defined E2 substrates and partners (Eisenhaber et al., 2007; Wenzel and Klevit, 2012). The best studied example of RBR E3 is Parkin. Structures of Parkin and other RBR E3s such as human homolog of Ariadne (HHARI) have greatly expanded our knowledge of the known mechanism of RBR E3s (Duda et al., 2013; Riley et al., 2013; Shimura et al., 2000; Trempe et al., 2013). The mechanism of substrate ubiquitination by RBR E3s is unique to RBR ligases and is composed of two steps (Fig. 3.1C). The RING1 recruits the ubiquitin-bound E2 enzyme and the ubiquitin is subsequently transferred to the RING2 domain, which then transfers the ubiquitin to the substrate (Wenzel et al., 2011, p. 7). It has been shown that this mechanism is the same for most, if not all, RBR E3s (Smit and Sixma, 2014).

3.2 E3 ligases and the 26S Proteasome

Despite the breadth of knowledge known about E3 ligases and their roles in the UPS, it remains unclear how a targeted substrate is brought to the proteasome once a protein is ubiquitinated and what role, if any, E3 ligases play in this process. For example, some proteins, such as Rad23, containing ubiquitin-like (UBL) and ubiquitin-associated (UBA) domains have been shown to associate directly with ubiquitin receptors on the 26S proteasome and act as chaperones for ubiquitinated substrates (Kim et al., 2004). Rad23 closely associates with its respective E3 ligase, Ufd2, but this interaction competes with proteasome binding (Kim et al., 2004). Other Rad23-like factors, such as Kex-2 proprotein convertase (KPC) which interacts with p27 ligase, are thought to cooperate with E3 ligases (Kamura et al., 2004), whereas proteins such as P97 have been linked to pathways delivering substrates to the proteasome (Richly et al., 2005). In addition, the UBL-UBA

protein, Ddi1, in yeast, binds to its E3 ligase Ufo1 and this is needed for the degradation of its target substrate, Homothallic (HO) endonuclease (Kaplun et al., 2005). In *Saccharomyces cerevisiae* (*S. cerevisiae)*, UBR1 and Ufd4 E3 ligases have been shown to bind to the 26S proteasome regulatory subunits Rpn2 and Rpt6 respectively (Xie and Varshavsky, 2000).

E3 ligases can therefore play a role in regulating substrate delivery to the proteasome. This is in part due to the essential role E3 ligases play in the last ubiquitination step before substrate degradation (Xie and Varshavsky, 2000). However, there have been additional studies suggesting existence of more direct links in which E3 ligases bind to the proteasome itself. For example, some E3 ligases have directly been shown to bind to the lid of the 19S, such as the RING E3 ligase UBR1 binding to the proteasome subunit Rpn2, whereas other E3 ligases have been shown to bind to the 26S proteasome base, such as the HECT E3 ligase UFD4 binding to Rpt6 (Xie and Varshavsky, 2000). These findings indicate a potential substrate delivery pathway to the proteasome where E3 ligases are involved in the formation of ubiquitin chains, as well as the delivery of ubiquitinated substrates to the proteasomes for degradation (Xie and Varshavsky, 2000).

Other E3 ligases of the HECT family have also been shown to associate with the 26S proteasome. These include the ubiquitin protein ligases (UPL) UPL1, UPL3 and UPL5 (Furniss et al., 2018). Further, the RBR E3 parkin binds and is recruited by Rpn13 through its UBL domain (Aguileta et al., 2015). Previous groups had also shown that parkin binds to the 19S Rpn10 via its UBL domain and to the 20S proteasome α4 subunit via its IBR-RING domain (Dächsel et al., 2005; Sakata et al., 2003). Parkin, which regulates various cellular processes, has been linked to a familial form of Parkinson's disease (Aguileta et al., 2015). The parkin-proteasome interaction could provide insight into how parkin is regulated as, for example, parkin activity increases upon Rpn13 binding. This increase in activity is thought to be due to parkin's relief of UBL-mediated autoinhibition (through self-ubiquitination) as its UBL is bound to the 26S proteasome instead (Aguileta et al., 2015). Parkin may also be binding to the proteasome to oppose the deubiquitinase (DUB) enzyme, Uch37, which is bound to and activated by Rpn13 (Aguileta et al., 2015; Hamazaki et al., 2006; Yao et al., 2006). Counteracting DUB activity by binding of E3s to the proteasome have been shown in other instances. For example, Ubp6 (Usp14 in mammals) binds via its UBL domain to Rpn2 in the proteasome. Ubp6 activity is

counteracted by Hul5, another E3 ligase (Crosas et al., 2006). The dynamic interplay between DUBs and E3s ensures proper recognition and degradation of substrates by the proteasome and provides a regulatory mechanism how E3 ligases binding to the proteasome can affect proteostasis (Aguileta et al., 2015). Other E3 ligases thought to directly interact with the proteasome co-purify with the 26S, including Ube3a/E6AP, Ube3c/Hul5, Rnf181, Huwe1, and UBR4 but their significance remains elusive (Besche et al., 2014).

3.3 The UBR family of E3 ligases

As mentioned previously, UBR E3 ligases are part of the family that recognise N-degrons, the N-terminal residues of a protein, for proteasomal degradation and are therefore referred to as N-recognins (Kwon et al., 2003, 2001, 1998). This pathway is called the N-end rule pathway and is one of the conserved mechanisms that determine protein half-life in cells and the rate of ubiquitin dependent proteasomal degradation. There are seven members of the family, UBR1-7 which are characterised by sharing a zinc-finger-like domain, the UBR box motif, and by recognising and ubiquitinating proteins with N-degrons for proteasome degradation following the N-end-rule pathway (Fig. 3.3 and Fig. 3.4) (Tasaki et al., 2009). N-degrons are recognized through the UBR box and they can either be type 1 (positively charged amino acid residues) or type 2 (hydrophobic and bulky amino acid residues) (Fig. 3.4) (Kaur and Subramanian, 2015; Tasaki et al., 2009). The UBR family of E3s are generally heterogeneous in size but contain domains specific to E3s, with the exception of UBR4 which has no E3 ligase activity (Tasaki et al., 2005). UBR1, UBR2 and UBR3 contain the RING ligase domain as discussed earlier, UBR5 contains the HECT domain, UBR6 the F-box domain and UBR7 contains the plant homeodomain domain (Fig. 3.3) (Callaghan et al., 1998; Tasaki et al., 2005; Tasaki and Kwon, 2007). The biochemical properties of other members of the UBR family, such as UBR3 and URB7, remain to be investigated.

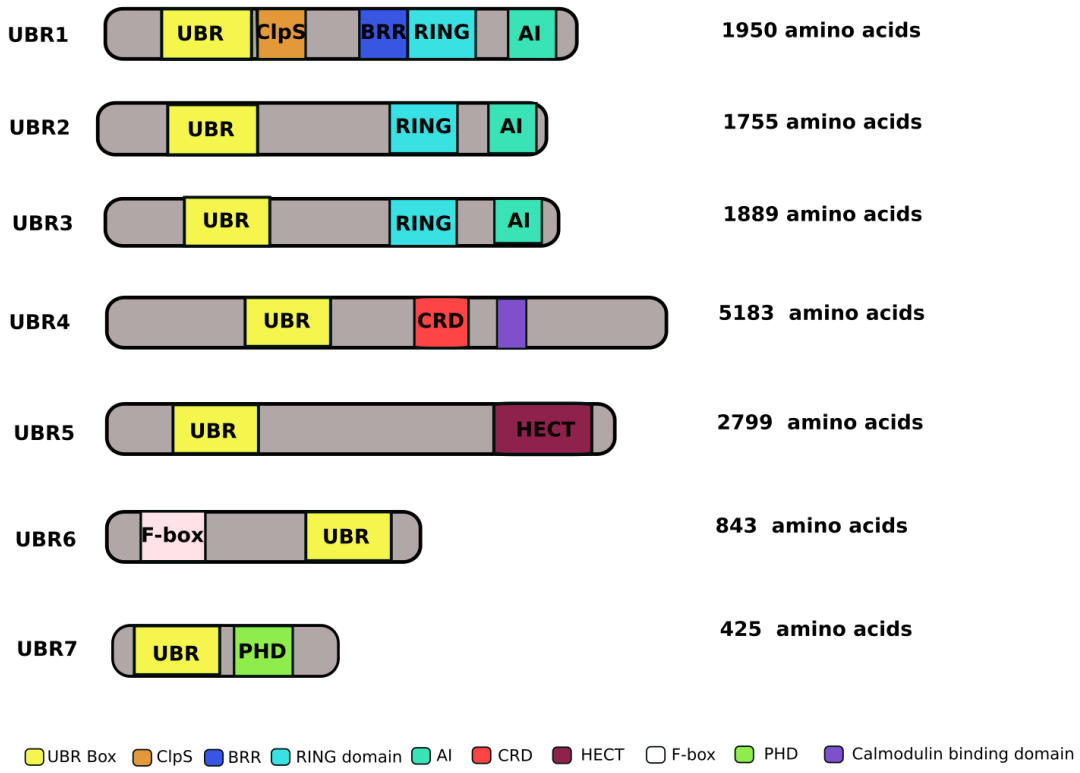


Fig. 3.3. Domain architecture of the UBR family of E3 ligases. There are seven UBR proteins that share the UBR box domain. The UBR box recognises type 1 and type 2 N-end rule substrates. The rest of the domains are labelled and a key is provided. ClpS denotes the *C. reinhardtii*-like domain; BRR is the basic residue-rich region; AI is the auto-inhibitory C-terminal domain; CRD is the cysteine-rich domain; PHD is the plant homeodomain.

3.3.1 UBR4

UBR4 is an atypical member of the E3 ligase family "ubiquitin protein ligase E3 component N-recognin" (UBR) and is a poorly structurally and biochemically-characterised protein. Initial studies of the protein identified UBR4 as an atypical UBR as it does not have typical E3 ligase activity and no known ubiquitin ligase domain (Tasaki et al., 2005). Also known as p600, due to its 600 kDa molecular weight, UBR4 was first reported as a large putative protein during the construction and characterisation of human brain cDNA libraries, corresponding to an open reading frame (ORF) present in clones named KIAA0462 and KIAA1307 (Ohara et al., 1997). Further studies showed UBR4 as a multi-functional protein interacting with the retinoblastoma protein and calmodulin, present both in the cytoplasm and the nucleus, and with roles in integrin-mediated

signalling and apoptosis (Nakatani et al., 2005). However, the significance of these interactions remains unknown.

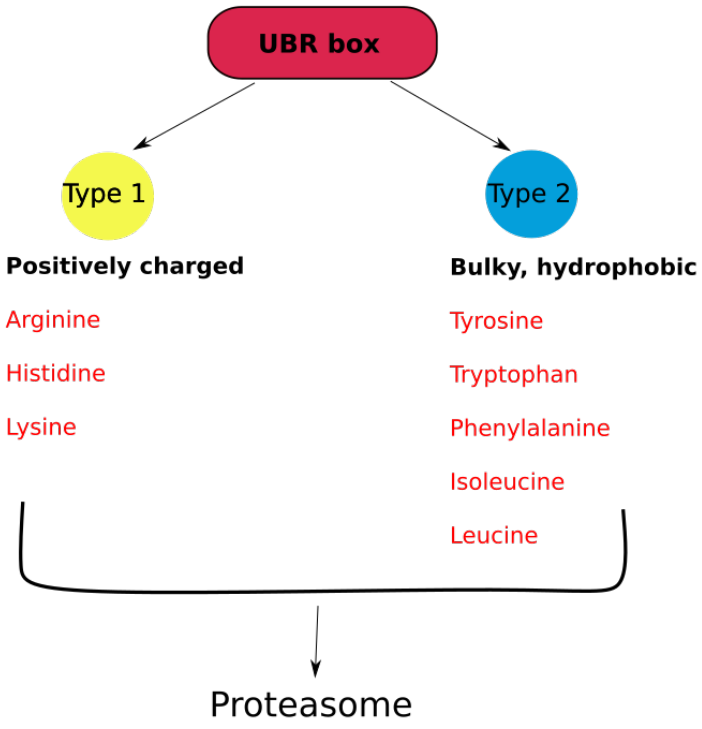


Fig. 3.4: The N-end rule pathway. The UBR box recognises type 1 and type 2 destabilising residues which are listed in the figure above. The primary destabilising residues are the product of chemical modifications to the amino acids (e.g. deamination and arginylation) resulting in N-terminal primary residues that are recognised by the UBR box in N-recognins which mediate ubiquitination and degradation by the proteasome.

UBR4 was also identified as a target of viral transforming E7 proteins of multiple papillomavirus (DeMasi et al., 2005; Huh et al., 2005). Further studies have showed that UBR4 is a critical protein in mammalian brain neurogenesis, and neuronal signalling and migration (Nakatani et al., 2005). UBR4 is abundantly expressed in all cells, and particularly enriched in the central nervous system. Ablation of UBR4 in mice leads to embryonic lethality and depletion in neuronal progenitors resulting in decreased production of neurons as well as faster neuronal differentiation (Belzil et al., 2014). Time-of-the-day dependent UBR4 expression in the suprachiasmatic nucleus suggests a role in the mammalian circadian clock regulation (Ling et al., 2014). Finally, a role of UBR4 in autophagy has also been proposed (Tasaki et al., 2013). Despite its abundance in cells and its suggested key roles in several cellular functions, UBR4 is still poorly characterised. Interestingly, it has been observed that UBR4 can co-purify with proteasomes that have

been purified by low stringency affinity chromatography (Besche et al., 2009; Jacobson et al., 2014; Scanlon et al., 2009). While these studies suggest an interaction between the proteasome and UBR4, the nature and significance of such an interaction is still unknown. Filling the knowledge gap of UBR4 research and addressing these aforementioned questions will aid in not only determining the role of this putative E3 ligase, but also in understanding how its interaction with the 26S proteasome may regulate its function. To this end, I used protein biochemistry to both reconstitute a stable complex *in vitro,* and electron microscopy and biophysical techniques to further investigate UBR4 *in vitro*.

3.4 Results

3.4.1 Preparation of human 26S proteasomes

Structural and functional analysis of proteasome ancillary proteins requires pure, homogenous, and biochemically functional 26S proteasome preparation. Several different proteasome purification methods have been described in the past. These methods have involved purifying proteasome from rabbit reticulocyte lysates, as shown by Hough et al. (1986), as well as purification from rat skeletal muscle by Dalhmann et al. (1995), mouse spleen and muscle by Besche et al. (2009) and brain and liver by Besche and Goldberg (2012).

I joined the lab at a time when the 26S proteasome purification from HEK293F cells was undergoing optimisation by Dr. Migle Kisonaite. The protocol for the purification of endogenous 26S proteasomes from HEK293F human cells, adapted to suspension culture, was modified from the previously reported pull-down method using the Rad23b UBL domain (from here on in referred to as UBL) (Besche et al., 2009; Chen et al., 2001). Rad23b, as mentioned in Chapter one, is a well-known shuttling factor containing both a UBA and UBL domain that delivers ubiquitinated proteins to the 26S proteasome via its UBL domain (Bertolaet et al., 2001). Under normal cellular conditions, the UBA and the UBL domains interact within themselves in the Rad23b complex, and this competes with binding to the 26S proteasome (Bertolaet et al., 2001). However, by truncating Rad23b and using the UBL domain alone, its affinity for the 26S proteasome can be exploited for complex purification. The UBL domain binds with high affinity to 26S proteasome via one

of the proteasome's ubiquitin receptors Rpn1 (Liang et al., 2014). The purification method involved cloning the UBL domain of Rad23b with a Twin-Strep tag at its N-terminus and using it to pull-down the 26S proteasomes (Besche et al., 2009; Chen et al., 2001). HEK293F cells were lysed using a manual dounce homogeniser and subsequently centrifuged. The resulting clarified lysate was incubated for one hour at 4 °C with Twin-Strep tagged UBL and loaded onto a streptavidin affinity chromatography column. In order to ensure maximum capture of 26S proteasomes, excess UBL bait (i.e 10-fold more than the proteasome concentration) was used. The UBL bound 26S proteasome was then eluted using 2.5 mM d-desthiobiotin. Fractions 5-7 (see Fig. 3.5A and B) where analysed by mass spectrometry (MS) and used in subsequent purification steps described below, in which proteasomes could be purified and separated from UBL.

After the initial streptavidin affinity purification, UBR4 co-eluted with the 26S proteasome (Fig. 3.5) and its identity was confirmed by MS. Interestingly, other E3 ligases could also be detected in the MS analysis such as the E3 ligase Huwe1. Fractions 5-7, which contained 26S proteasome complexes, but importantly did not strictly co-elute with the main peak UBL fractions (fractions 7-12 from Fig. 3.5), were used for further purification. However, UBL removal through a second purification step was necessary as was the separation of different proteasome species (20S, single-capped proteasomes, and double-capped proteasomes). In order to remove UBL from the sample while maintaining proteasome purity and integrity, several experimental conditions were tested by Dr. Kisonaite. I joined at a time when the second purification step by glycerol gradient ultracentrifugation was being optimised by Dr. Kisonaite. Other methods, such as size exclusion chromatography (SEC) had been tested, however, this technique did not allow for efficient separation of double and single capped proteasomes due to the elongated shape of the proteasome, which results in anomalous elution and diffusion in SEC columns. The resolution of the column did also not allow sufficient separation of double and single capped proteasomes. The method of glycerol gradient centrifugation specifically uses rate-zonal centrifugation in which the sample is added on top of a density gradient and had previously been successfully used for proteasome preparation (Hough et al., 1986, Kopp et al., 2001). Rate- zonal centrifugation separates particles based on their size and mass, and thus 20S and 26S proteasomes, and proteasomes bound to other factors can be separated. Together with Dr. Kisonaite, we found that using 30-15% glycerol

gradient centrifugation step, double and single-capped proteasomes could be separated, with other sample components, such as UBR4 and 19S, sedimenting at different glycerol concentrations (Fig. 3.6). Glycerol gradient centrifugation also allowed for efficient removal of UBL protein and no UBL protein were seen on glycerol gradient as indicated by western blot.

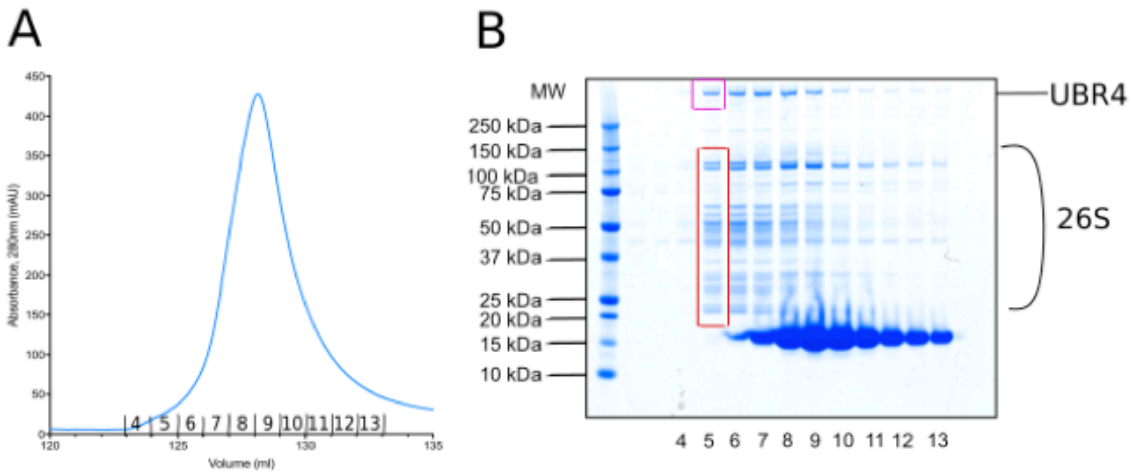


Fig 3.5: 26S proteasome purification with ATP. The 26S proteasome was purified indirectly by pulling on the streptavidin- tagged UBL domain of Rad23. A) Chromatogram of the strep-affinity chromatography purification. B) SDS-PAGE of a representative 26S purification in which 26S proteasome (red box) elutes with UBL elution. The top band (pink box) is UBR4 which seems to co-elute with the 26S proteasome. The elution fractions are indicated on the X axis.

In order to assess the functionality of the purified proteasome, fractions containing double-capped proteasomes were pooled, concentrated and their peptidase activity was measured, using highly fluorescent 7-amino-4-methyl coumarin (AMC)-tagged peptides: Suc-L-L-V-Y-AMC (chymotrypsin-like activity), Boc-L-R-R-AMC (trypsin-like activity) and Z-L-L-E-L-AMC (caspase-like activity). When these fluorogenic peptides are cleaved they release the highly fluorescent AMC from the peptide (Kisselve and Goldberg, 2005; Liggett et al., 2010) (Fig. 3.6B). The activity of the purified proteasomes was comparable to that reported for endogenous proteasome activity levels, with ß5 being the most active, followed by ß1 and then ß2 being the least active (Fig. 3.6B) (Kisselve and Goldberg, 2005; Liggett et al., 2010). Double-capped proteasomes were also evaluated by negative stain EM to check that they were correctly assembled and had retained their structural integrity throughout the purification procedure (Fig. 3.6C).

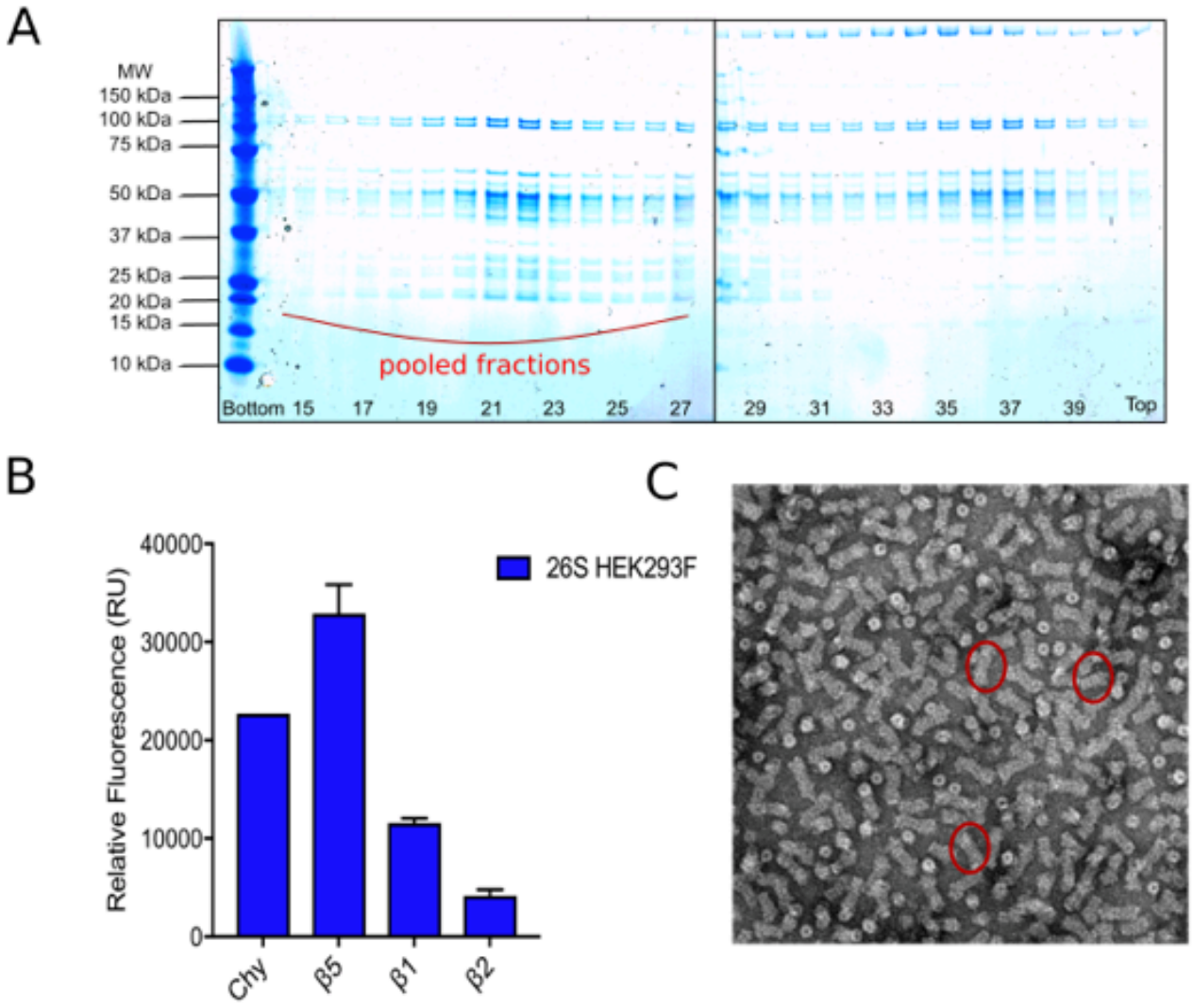


Fig 3.6: Optimisation of the 26S proteasome purification with ATP. A) After an initial streptavidin affinity purification fractions (5-7 from Fig. 3.5) were concentrated and separated by glycerol gradient centrifugation, as indicated on the SDS-PAGE gel B). Fractions (labelled pooled fractions, fractions 15-27) were pooled, concentrated and checked for proteolytic activity and C) for structural integrity using negative stain EM with a representative micrograph is shown. Error bars represent the standard error of the mean (SEM).

3.4.2 The 26S proteasome co-purification with UBR4

Early studies demonstrated that ATP is required for the *in vitro* purification of stable 26S proteasomes and before I joined the lab, 26S proteasome purifications had been conducted

in the presence of ATP and $MgCl_2$ (Eytan et al., 1989; Ganoth et al., 1988; Liu et al., 2006). During optimisation of the 26S proteasome purification we used ATP (Fig. 3.5 and Fig. 3.6). However, I decided to test whether the inclusion of ATP in purification buffers affected the amount of UBR4 observed in 26S purifications and whether this would affect the proteasome species I purified. I performed the streptavidin affinity purifications in the absence of ATP and $MgCl_2$ (Fig. 3.7). As illustrated in Fig. 3.7, removal of ATP in purification buffers did not have any effect on the co-purification of UBR4 and I thus decided to conduct further experiments without ATP present. These results also prompted Dr. Kisonaite to investigate the 26S proteasome purification without ATP. She confirmed that efficient purification of intact and stable 26S proteasomes retained proteolytic activity and structural integrity under ATP free purification conditions (Kisonaite et al., in press). Furthermore, Dr. Kisonaite observed that $MgCl_2$ exclusion from purification buffers by using EDTA was necessary in order to obtain stable proteasome complexes in the absence of ATP. The rationale behind this is that if ion chelators are not used, large reservoirs of intracellular ions released upon cell lysis may destabilise the proteasomes when there is no nucleotide presence to counteract this (Kisonaite et al., in press). Given that our results showed that removal of ATP did not have any effect on the presence of UBR4, I decided to conduct further experiments without ATP present.

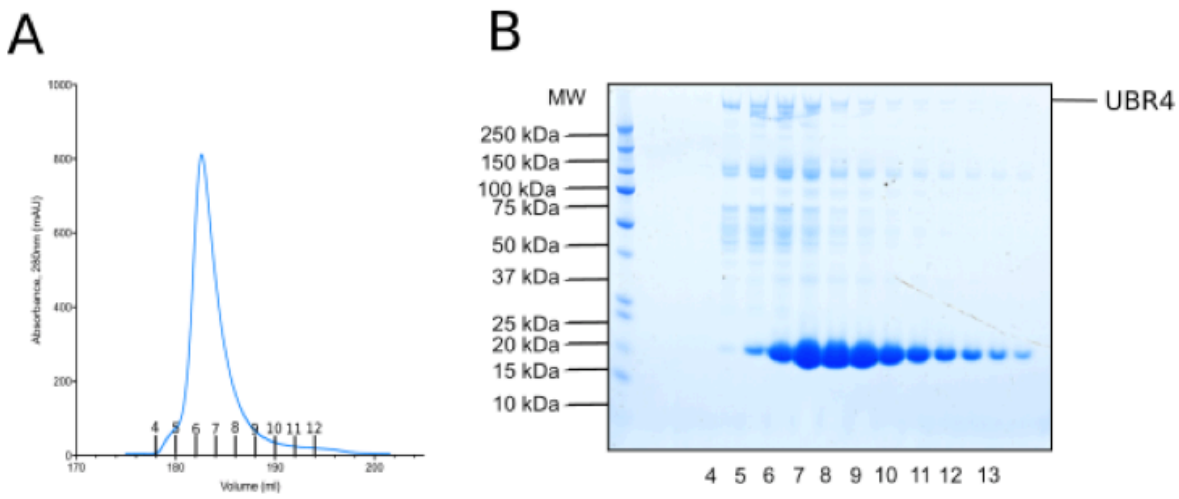


Fig. 3.7: 26S purification without ATP. The 26S purification was performed without ATP in the purification buffers in which A) Representative chromatogram and B) the elution fractions analysed by SDS-PAGE electrophoresis are shown. UBR4 (top band) is co-eluting with 26S proteasomes despite the lack of ATP in the purification buffers.

Since different proteasome species are not separated by the first streptavidin affinity step, a second step involving a variety of biochemical separation methods was adopted to further separate UBR4 from non-interacting proteasome subunits. Whilst optimising the 26S purification, I explored different glycerol gradients. The glycerol gradient was conducted after a streptavidin affinity purification step in which samples containing UBR4 and 19S were pooled and concentrated. Although I could achieve some separation of 26S proteasome from UBR4 and other unwanted proteins, the resolution needed to be optimised to ensure further separation of proteasome species. This was achieved by using 15% (top)-30% (bottom) glycerol gradient (Fig. 3.8). This further purification step allowed the clear identification of three different peaks: the initial 26S proteasome peak (fractions 8-24 highlighted in pink), a UBR4 containing peak together with the 19S (28-42 highlighted in red), and a second UBR4 containing peak with other proteins (46-53 highlighted in orange).

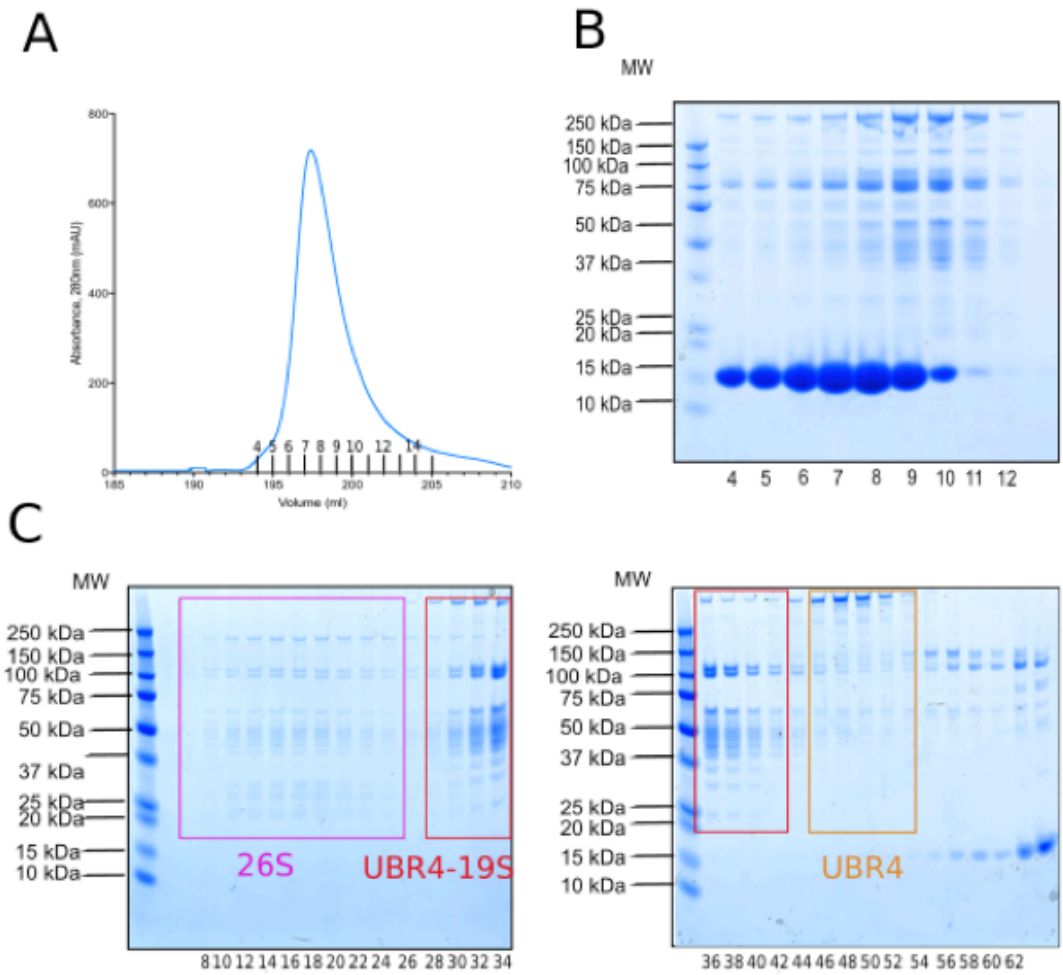


Fig. 3.8: UBR4-19S purification. UBR4 was purified using a standard 26S purification protocol described in Section 1.4.1 in which A) a representative chromatogram and B) SDS-PAGE gel are shown. Fractions 8-12 (shown in panel B) containing UBR4 were pooled after streptavidin affinity purification, concentrated and separated using glycerol gradient centrifugation where (representative SDS-PAGE is shown in C). C) Gels are from a 15-30% glycerol gradient in which the heaviest fractions are found in the early fractions (26S proteasomes), lighter fractions are separated with increasing fraction numbers. UBL separated in fractions 54-62 whereas UBR4 is seen co-eluting with 19S caps in fractions 28-42.

In order to further improve the purity of the UBR4 sample preparation, i.e. to remove other non-interacting proteins, I decided to explore additional chromatography techniques, such as ion exchange chromatography, using a variety of conditions and resins. Ion exchange chromatography separates proteins on the basis of their charge by using charged resins and proteins are eluted by increasing salt concentrations. By using positively charged DEAE

resin, I was able to separate UBR4 from non-interacting proteins. However, even under high salt concentrations, I was unable to separate UBR4 from the proteasome fractions, thus showing that the UBR4-proteasome complex was not easily separated from the proteasome (Fig. 3.9). UBR4 was seen to preferentially co-elute particularly with the 19S particle, consistent with the known effect of high salt on the interaction of 19S and 20S particles, also seen in glycerol gradients (see Fig. 3.8). These results suggest a direct interaction between UBR4 and proteasome subunits, rather than a co-elution of the two.

Ion exchange chromatography did not effectively separate many unwanted proteins after the first step of ion exchange chromatography, as additional protein bands besides the proteasome subunits appeared on the gel. I therefore wanted to explore other ways in which UBR4 could be purified with any other interacting proteins from HEK293F cells (Fig. 3.10). I decided to conduct calmodulin affinity chromatography experiments as UBR4 has been shown to associate with calmodulin (Kim et al., 2018). After streptavidin affinity chromatography co-purification of UBR4 with proteasomes and UBL protein (pooled fractions 5-8), I performed an affinity chromatography step using calmodulin beads (Fig. 3.10). The buffer used for the equilibration of calmodulin beads, washes and binding buffer consisted of 50 mM Tris pH 7.4, 50 mM NaCl, 2 mM $CaCl_2$, 5 mM $MgCl_2$, 1 mM ß- 2- Mercaptoethanol, 5 mM ATP. I conducted two parallel experiments, one in which 3 mM EGTA was added during the elution step and another in which 30 mM EGTA was added, to ensure complete protein elution from the beads. The results showed that despite retention of UBR4 with proteasome subunits, there was insufficient separation of UBL (Fig. 3.10). Furthermore, UBR4 was not free from interaction with other unwanted proteins as effectively as elicited by the glycerol gradient centrifugation.

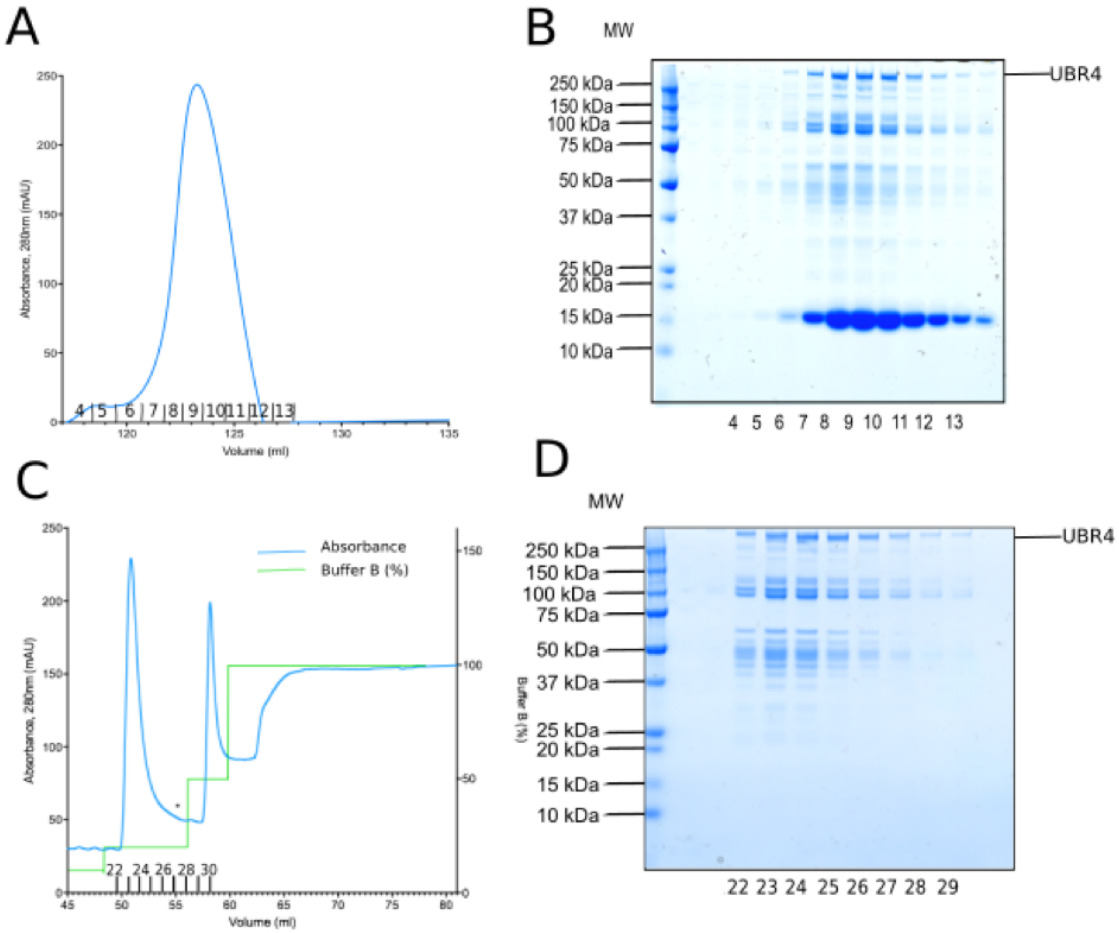


Fig. 3.9: UBR4 separation using Ion Exchange Chromatography. UBR4 was purified as before using UBL with A) representative chromatogram B) and SDS-PAGE gel shown. This was followed by ion exchange chromatography with C) a representative chromatogram and D) SDS-PAGE gel (D). UBR4 continued to co-elute with proteasome fractions as seen in panel D.

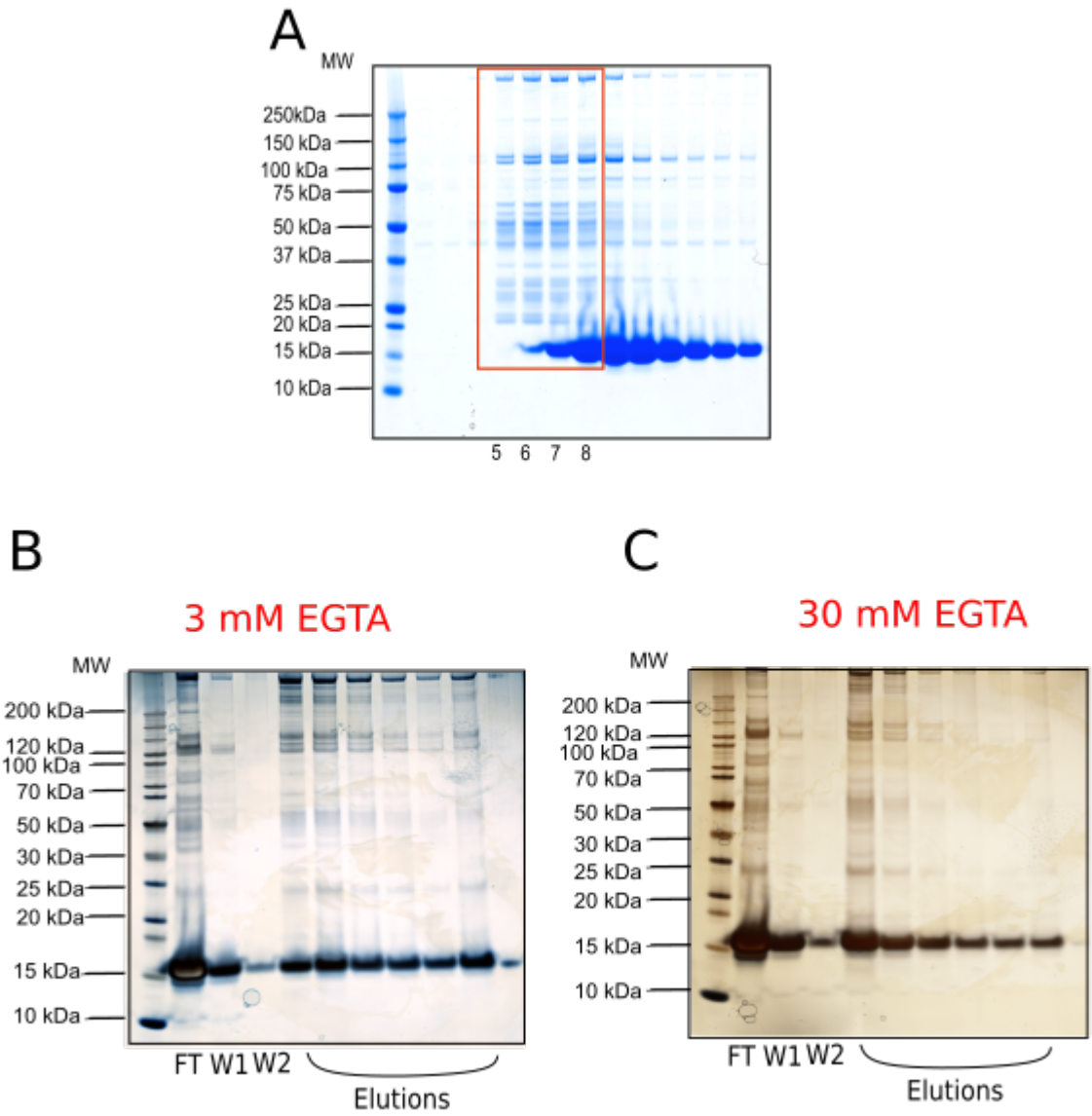


Fig. 3.10: Calmodulin affinity purification. A) Samples 5-8 (red box) were taken from a first streptavidin affinity chromatography purification step where the SDS-PAGE gel and incubated with calmodulin beads. B) In the calmodulin affinity chromatography experiment, flow through (FT) and washes 1 and 2 (W1, W2) were collected and samples were eluted with either 3 mM EGTA (B) or 30 mM EGTA (C).

In conclusion, the glycerol gradient centrifugation protocol provided the best separation of proteasome subunits with UBR4 against other non-interacting proteins (Fig. 3.8). To further characterise UBR4 with the 19S obtained from glycerol gradient centrifugation, size-exclusion in combination with multiple angle light scattering (SEC-MALS) was conducted. This technique allows for highly accurate mass measurement of macromolecules. SEC-MALS separates molecules based on their hydrodynamic volume,

and then uses angular dependence and intensity of scattered light to measure the absolute molar mass and size of the molecules in solution (Sahin and Roberts, 2012). For these experiments, I performed an initial streptavidin affinity purification, followed by glycerol gradient centrifugation. Fractions containing UBR4-19S (lanes 26-30 from Fig. 3.11) were then concentrated and used for SEC-MALS. The experiment was repeated several times and results obtained reported a molecular weight ranging between 1.4-1.8 MDa (Fig. 3.11). The range obtained was a result of the low signal observed likely due to the low concentrations of protein sample. However, the consistent larger molecular weights observed, which were larger than the 20S and 19S complexes molecular weights (approximately 700 kDa and 900 kDa respectively) suggest the presence of a complex that contains UBR4. This further supported the observation that UBR4 was stably-associated with the 19S complex presumably through direct binding.

To investigate this further, fractions containing both UBR4 and 19S particles were taken, concentrated, and imaged using negative stain electron microscopy (Fig. 3.12). Electron microscopy is a powerful technique used to visualise and solve structures of molecular complexes. Electron microscopes are composed of a range of lenses, an electron source and a detection system that detects the electron scatter (Bozzola and Russell, 1999). The electron source emits electrons and these are accelerated by an electric field (Bozzola and Russell, 1999). In high resolution cryo-EM, the typical accelerating voltage is 300 kV whereas a standard 120 kV microscope is commonly used for negative stain electron-microscopy. The electron beam is converged and focused by the condenser and objective lens system, illuminating the sample. The objective aperture helps to eliminate any highly scattered electrons, increasing the phase contrast. Finally, after further rounds of magnification by the projector lens, the electrons are detected and an image that is a projection of the 3D object is produced (Murata and Wolf, 2018).

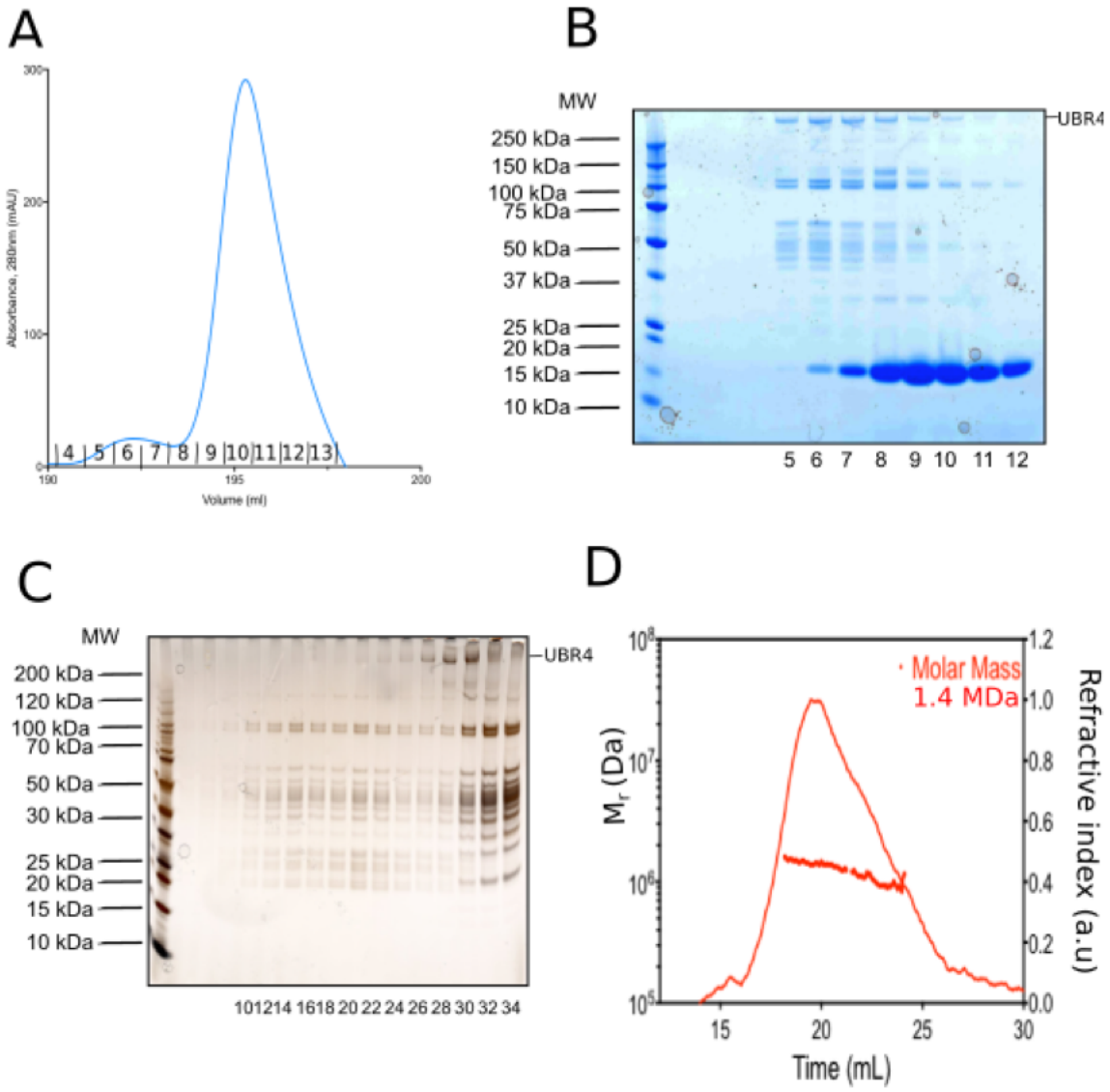


Fig. 3.11: Investigating UBR4-19S molecular weight. UBR4-19S were purified by strep-affinity chromatography and the chromatogram and SDS-PAGE gel is shown in A) and B) respectively. C) Step two of the purification was a glycerol gradient centrifugation and the SDS-PAGE gel is shown in panel C. D) Samples containing UBR4 were then concentrated and run on SEC-MALs with an estimated molecular mass of 1.4 MDa.

Samples are visualised with negative stain to understand the composition, concentration and size of the particle. Negative stain is a method in which the background, rather than the sample is stained. A droplet of protein sample is placed on an EM grid and incubated with a heavy metal salt solution, such as uranyl acetate, that can scatter electrons (Sharp et al., 1950; van Bruggen et al., 1960). The salt precipitates out of solution due to evaporation sedimenting around the sample. Excess stain is then blotted using filter paper, with the remaining stain coating the grid and therefore allowing imaging of the protein surface

(Sherman et al., 1981). Negative staining is used as a quality control check of samples as the stain used limits the resolution of structures, usually in the range of 20 Å, making it unsuitable for higher resolution studies (Harris and Scheffler, 2002). Nevertheless, negative staining of samples allows for qualitative understanding of structure and distribution of macromolecules present in the sample (Booth et al., 2011).

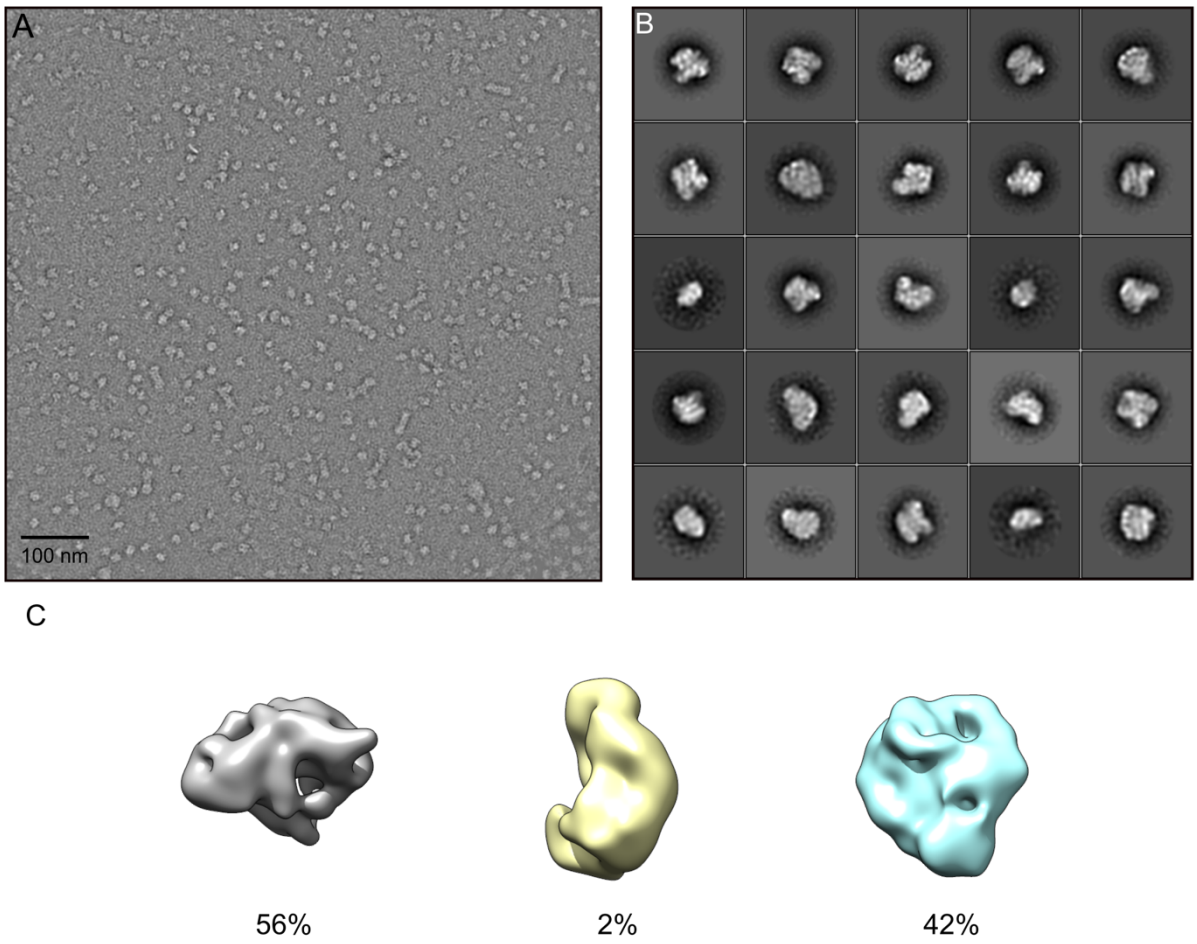


Fig. 3.12: Negative stain electron microscopy of UBR4-19S samples. UBR4-19S samples from glycerol gradient centrifugation were negatively stained and detected under an electron microscope (A). Representative negative stain EM micrograph. B) Representative class averages generated from 31,222 particles. Image processing showed globular shapes for 2D classes and these were further classified to show different populations within the sample as shown in C). C) Initial 3D models were generated from 31,222 particles with the most detailed model shown in a box (model has a dimeter of ~240 Å).

Negative stain grids were prepared of fraction 30 as shown in Fig. 3.8. Negatively stained samples containing UBR4 and 19S complexes showed so far uncharacterised globular shapes with clearly distinct appearances to usual proteasome particles (Fig. 3.12A).

Importantly, this indicated that I was likely imaging UBR4 alone or in complex with the 19S subunit but not solely the proteasome complex. The samples observed were heterogeneous and processing of these images using 2D and 3D classification, as explained in Chapter 5, was limited in resolution and subsequent 2D class averages and 3D initial models were difficult to interpret. To obtain higher quality complexes and to investigate the UBR4-proteasome interaction further, I decided alternative approaches to purifying UBR4 with the 26S proteasome, or 19S sub-complexes, described below in Section 3.4.3.

As a test, I decided to mix purified UBR4-19S sample with purified recombinant 20S (as discussed in Chapter 2 and 5) at a 1:1 molar ratio, to determine whether the UBR4 samples containing UBR4 and 19S could form 26S complexes when incubated with recombinant 20S (Fig. 3.13). Usually, if 19S and 20S are mixed, 26S proteasomes are able to form. However, the results obtained showed mainly 20S proteasomes whereas no 26S proteasomes were detected (Fig. 3.13), suggesting that UBR4-19S may either not be able to bind to 20S proteasomes *in vitro* or might require additional cofactors and/or conditions in order to be able to do so. Alternatively, the procedure of sample preparation for negative stain microscopy may disrupt the interaction, perhaps due to the very low pH of the staining solution.

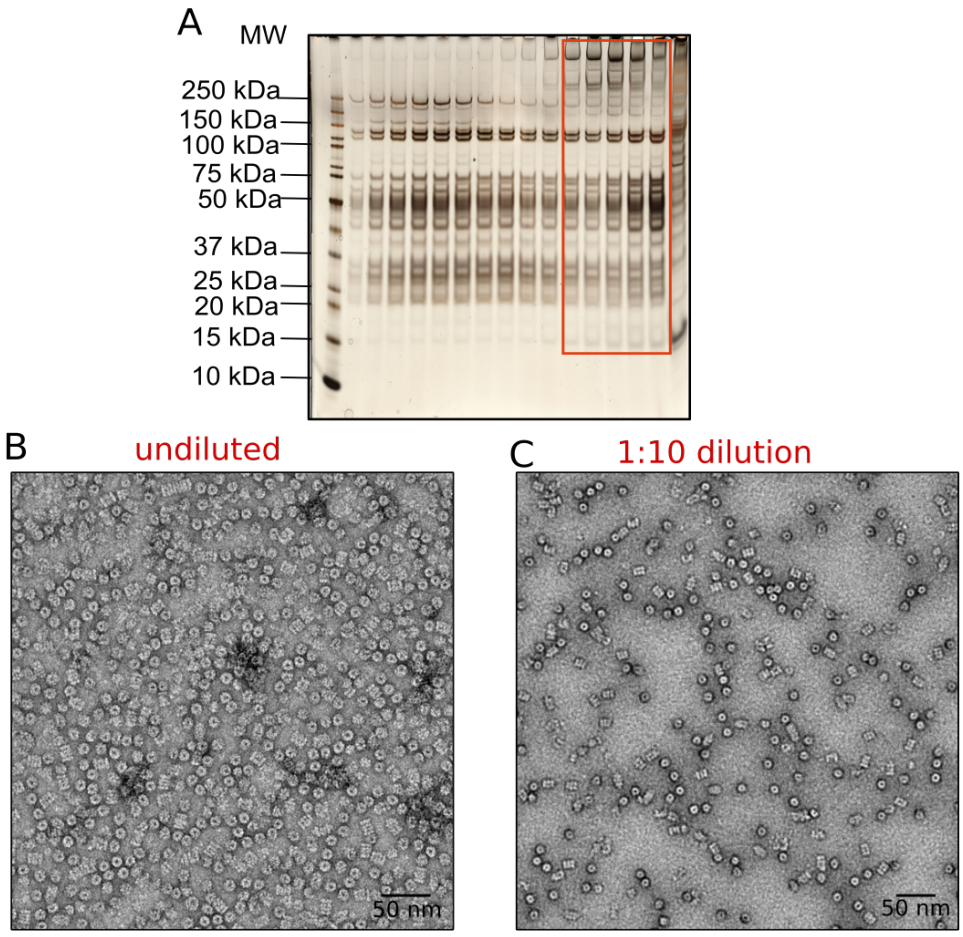


Fig. 3.13: 19S-UBR4 samples and 20S proteasome mix. A)19S-UBR4 fractions from a glycerol gradient centrifugation step (red box) were pooled and mixed with at a 1:1 of 1.1 µM recombinant 20S purified using method outlined in Chapter two. The samples incubated at room temperature and negative stain samples of undiluted (B) and 1:10 dilutions (C) were prepared.

3.4.3 Tagging of endogenous human UBR4

The UBR4 containing sample obtained during the endogenous 26S purification, described in section 3.2.1, allowed for some analysis on UBR4. However, due to the heterogeneous nature of this sample, broader characterisation was hampered. I thus wanted to find a way to purify UBR4 to enable further biochemical and structural analysis. As mentioned, UBR4 is large, almost 600 kDa protein and there are currently no available cDNA constructs. Cloning and recombinant expression of UBR4 is thus not a trivial task due to the intensive cloning requirement to patch together several hundred exons from genomic DNA. I therefore decided to adopt an endogenous tagging system using clustered regularly interspaced short palindromic repeats (CRISPR) technology, as this would allow direct purification of UBR4 and could also be a helpful *in vivo* tool for later experiments.

CRISPR is a relatively novel genome editing tool used to change part of an organism's DNA, such as the removal or addition of genetic material (Cong et al., 2013; Jinek et al., 2012; Mali et al., 2013). This is conducted by an enzyme, termed Cas9, which is directed to the site in the genome that requires alteration by a small guide RNA (gRNA) that itself binds to a specific target DNA sequence in the genome (Cong et al., 2013; Jinek et al., 2012). Cas9 then makes an incision on the DNA backbone recruiting the cellular DNA repair machinery that usually introduces changes to the adjacent genetic information (Ran et al., 2013). The repair of the Cas9 induced breaks can be exploited to either create small deletions or to introduce additional genes or genetic elements by homologous recombination, a donor piece of DNA that include areas of homology surrounding the desired site and exogenous DNA sequences (Ran et al., 2013).

In my case, the CRISPR strategy involved cloning two plasmids: one with a fragment of the UBR4 gene acting as a guide for the Cas9 enzyme the other, a donor plasmid providing the template sequences to be inserted (Fig. 3.14A). The tag (in this case Twin-Strep tags at the C-terminal end of UBR4) along with left and right arms of homology to the gene (situated before and after the UBR4 stop codon) were cloned in a donor plasmid. Both donor plasmid and gRNA plasmids were then co-transfected into human HEK293F cells using a non-lipid based chemical transfection method (Koleva et al., 2006). Several colonies of transfected cells were grown to improve the chances of gene incorporation into the genome. Cells were then selected using flow-cytometry, via a fluorescent marker as an initial step for selection of gRNA plasmid uptake, followed by colony growth using neomycin drug selection. HEK293F cells were initially adapted to adherent cells for efficient transfection grown from single cells to colonies and then converted back to suspension cells when scaling up to large volume cultures. This required optimisation by gradually varying the percentage of FBS when adapting the cells from adherent growth back to growing in suspension. Upon successful colony growth from single cells, different colonies were taken and screened using PCR (Fig. 3.14B). PCR screening primers were designed so that part of the gene and the tag were amplified. The results showed that the Twin-Strep tag had been incorporated into the genome for most colonies (Fig. 3.14B).

Upon successful tagging of the UBR4 gene at the C-terminus, UBR4 was purified using the same purification strategy (streptavidin affinity chromatography followed by glycerol

gradient centrifugation) as described in Section 3.4.2 (Fig. 3.15). During the purification procedure, UBR4 was successfully detected, however, very small yields of protein were obtained and other proteins co-purified (Fig. 3.15). These additional proteins could be either contaminants or bona fide interacting proteins. The presence of UBR4 in the purified extracts was confirmed by MS which also identified several proteasomal subunits, providing further evidence of a direct interaction between UBR4 and the proteasome. Attempts were made to increase the yield of UBR4, by increasing the size of the cultures grown. However, UBR4-tagged cell lines did not survive for more than ten passages, suggesting that the cell lines grown were not very stable and that UBR4 genetic manipulation may have played a role in this cell growth instability.

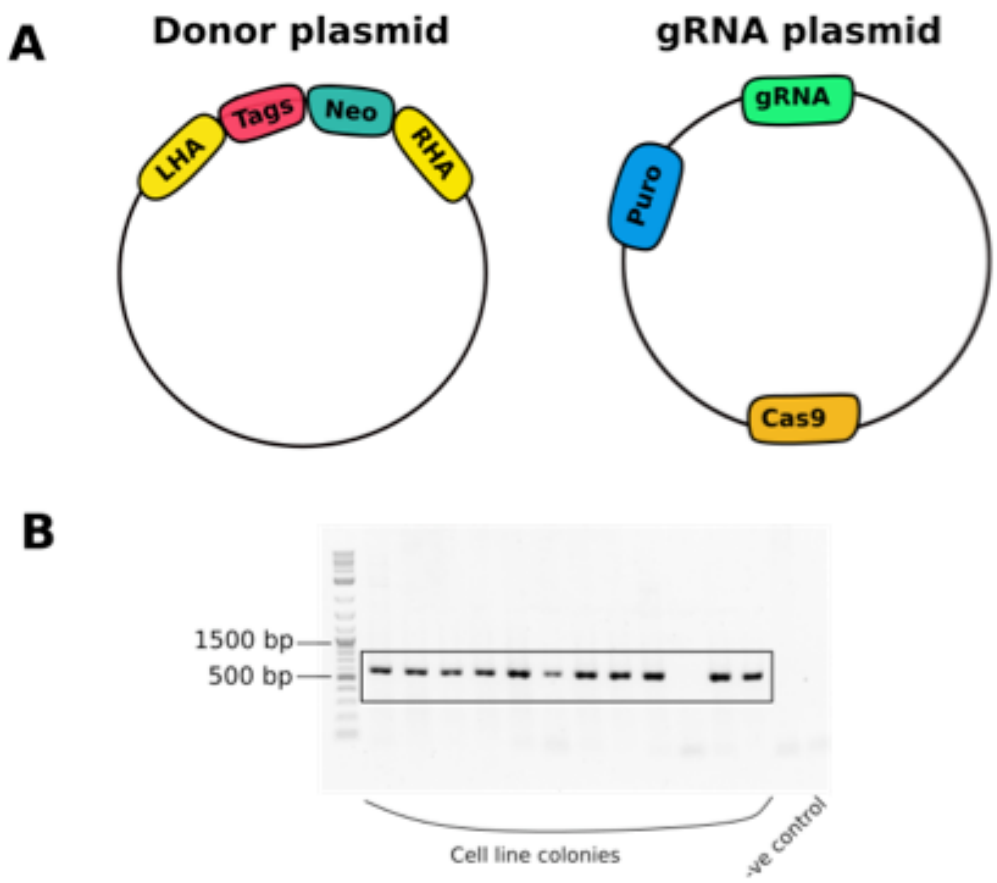


Fig. 3.14: Endogenously tagging the UBR4 gene in HEK293F cells using CRISPR. A) Donor plasmid containing the streptavidin tags, resistance marker and left and right homology arms (LHA and RHA, respectively, see appendix) and the gRNA plasmid which contains the Cas9 and guide RNA (gRNA) sequence. B) a PCR was conducted to screen for positive clones in which the double streptavidin tag had been introduced. Untransfected HEK293F cells were used as a negative control.

Purified UBR4-tagged samples were examining using negative stain electron microscopy (Fig.1.16). In the sample, both globular like particles and some 20S-like components were observed. Due to the heterogeneity of the sample seen from the presence of many proteins in both steps of the purification (see Fig. 3.15), it was difficult to unambiguously identify UBR4 particles. However, the fact that proteasome components were pulled down and identified by MS provided further support for a UBR4-proteasome specific interaction, as this time the UBR4 tag was used for the initial affinity step, rather than using a proteasome specific bait.

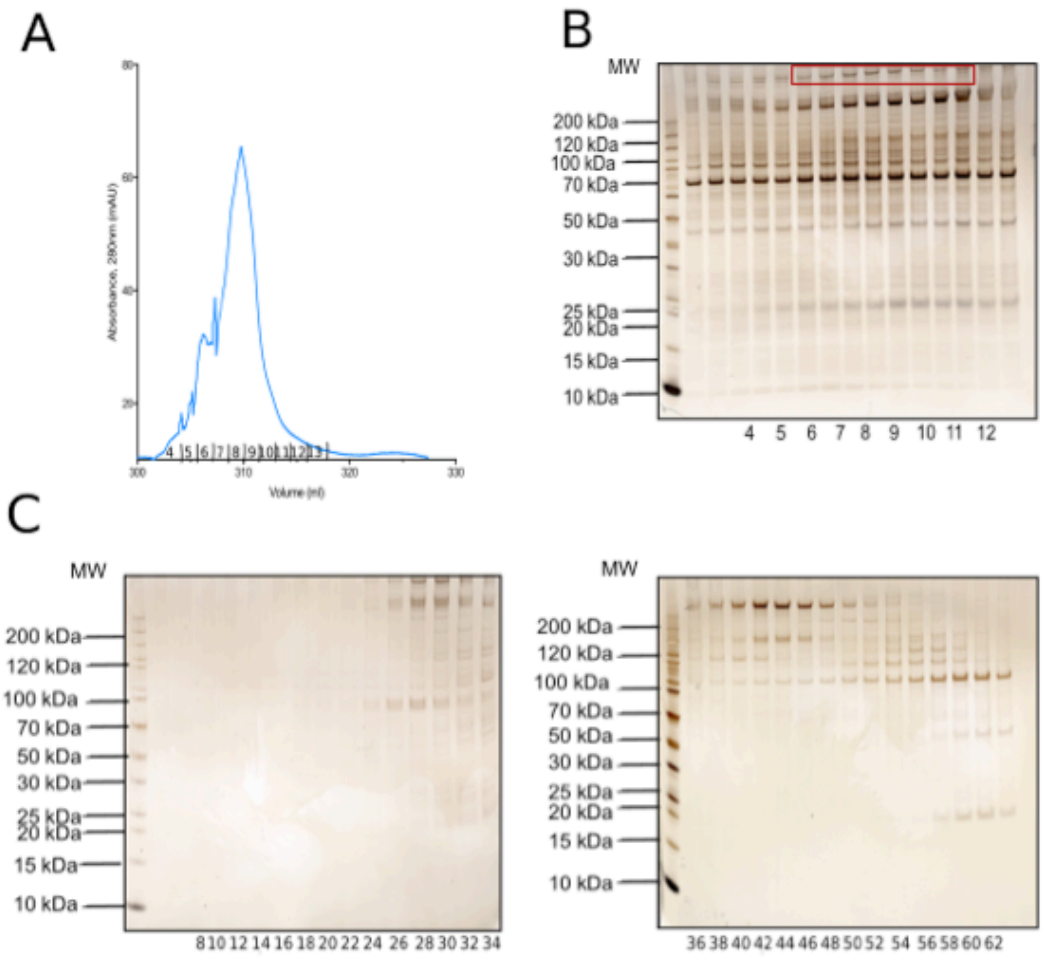


Fig. 3.15: Endogenous streptavidin-tagged UBR4 purification. Streptavidin- tagged UBR4 (red box, panel B) was purified by affinity chromatogram and a A) representative chromatogram B) and SDS-PAGE gel visualised by silver staining is shown. C) A glycerol gradient centrifugation was conducted with the heaviest fractions from fraction 8 continuing to the lightest fractions (fraction 62). Fractions 25-32 from panel C were pooled and analysed by MS.

The traces of proteasomal subunits observed in the glycerol gradient centrifugation fractions (Fig 3.15 panel C fractions 25-32) provided further support for a direct interaction between UBR4 and the proteasome. Indeed, some 20S proteasome particles could be identified in negative stain EM images of the purified UBR4-tagged samples (Fig. 3.16) and peptides corresponding to several proteasomal subunits, such as 20S α and β subunits, were detected by MS. Consequently, I attempted to further characterise these samples. To do this, the proteasome peptidase activity of undiluted UBR4-tagged samples (concentrated fractions 25-32 from Fig. 3.15C) was determined. Different concentrations of purified 20S proteasome (as described in Chapter two) were used in the fluorescent degradation assay to estimate the activity of the UBR4-tagged samples. This allowed me to conclude that the activity in UBR4-tagged purified samples contained comparable

proteasomal activity to known 20S proteasome samples. The UBR4-tagged samples contained similar levels of ß1 and ß5 activity as 0.1 µM 20S proteasome samples, even surpassing activity levels of diluted 0.01 µM 20S sample (Fig. 3.17A and B). The ß2 activity in the UBR4-tagged sample, however, was not observed at high levels (Fig. 3.17C). The presence of proteolytic activity in the UBR4-tagged purifications provides further evidence of the association of UBR4 with the proteasome. The observation of 20S-like particles (Fig. 3.16) in negative stained grids of UBR4-tagged samples with partially functional proteasomes represents an additional layer of evidence and are likely at least partially functional proteasomes.

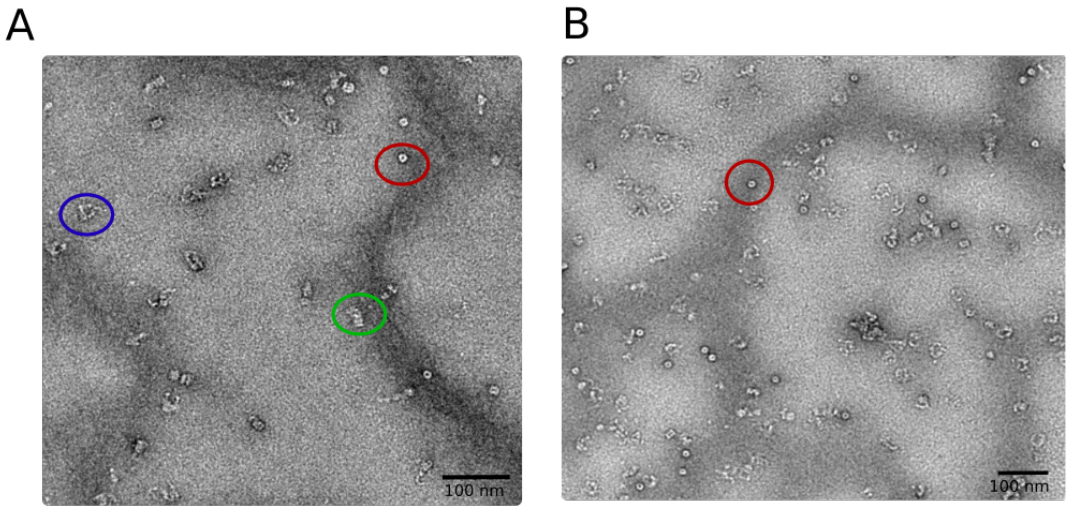


Fig. 3.16: Micrographs of streptavidin-tagged UBR4 samples. A and B) Fractions 25-32 from (see Fig. 3.15C) were concentrated and visualised using negative stain electron microscopy. 20S proteasomes were observed in the same sample (red rings) as well as some other 26S-like complexes (green ring). There were also fibril-like structures (blue ring) that were not characterised further.

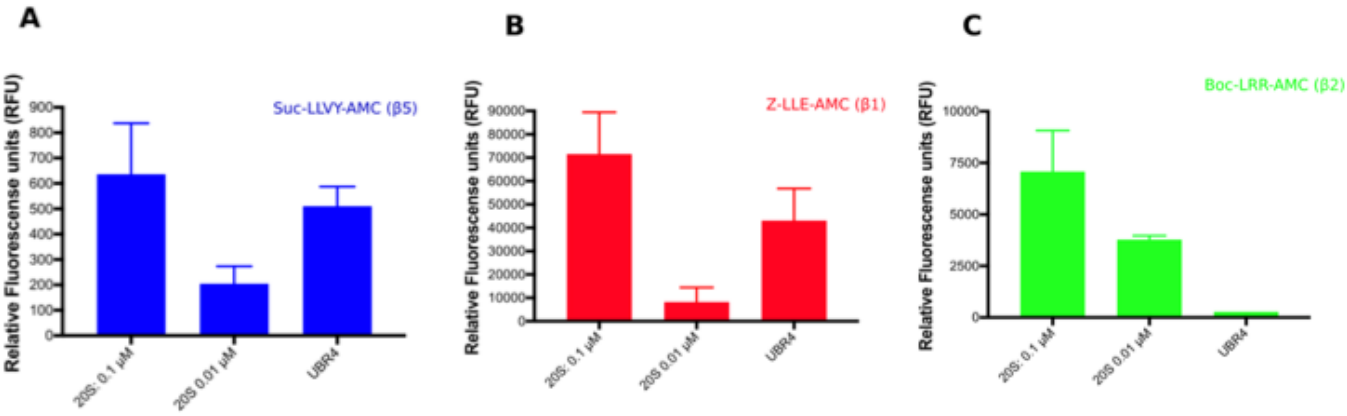


Fig. 3.17: Activity Assay of UBR4 and 20S samples. All three proteasome ß-subunits were tested for protease activity for 20S sample compared to UBR4 sample from Fig. 1.8, in which fractions 25-32 were collected and concentrated. A) Blue B) red and C) green bar graphs show ß5, ß1, and ß2 activity respectively. Experiments were done in triplicates and error bars represent the standard error of the mean (SEM).

3.4.4 UBR4 and the proteasome sample characterisation

Finally, I decided to investigate the ubiquitination state of the proteasome purified with UBR4. To this end, different proteasome samples were purified from HEK293F cells as described previously, incubated with an anti-Ub antibody and detected using a chemiluminiescence technique. Here, the HRP enzyme bound to the ubquitin antibody catalyses the oxidation of luminol, which can be detected by a change in light emisison (Thorpe and Kricka, 1986). Three different samples were compared: 1) UBR4-19S samples as described above 2) 26S proteasomes purified as described and shown in Fig. 3.4.1, and 3) UBR4-tagged purified samples (Fig. 3.15). Samples were loaded onto the gel to check for the presence or absence of ubiquitination and visualised using the western blot technique. The results showed that 19S-UBR4 samples exhibited high ubiqutination levels whereas the 26S sample showed low ubiquitination levels and UBR4-tagged samples showed no ubiqutination (Fig. 3.18). The differences observed between 19S-UBR4 samples and UBR4 from tagged cell lines suggests that having high levels of proteasomal subunits stimulates the level of ubiquitination, because it is most likely the proteasome subunits themselves that are ubiquitinated. The fact that these proteasome subunits are

ubiquitinated is interesting as it suggests some level of regulation, however, whether this is caused by UBR4 needs to be further investigated.

Ubiquitination of purified proteasome samples was also tested by a UBI-capture pull down, where proteins are run through a matrix that captures ubiquitinated substrates, as described in Chapter two (Benvegnu et al., 2017) (Fig 3.19). Here, samples that were concentrated after glycerol gradient purificaiton (as described previously and shown in Fig. 3.19) were tested. Ubiquitinated protein lysate served as a control reference (Fig. 3.19). The results show that the UBR4 and 19S purified samples contained ubiquitinated proteins and interacted with ubquitin-associating resin (Fig. 3.19). However, the presence of ubiquitinated proteins in the sample does not directly imply that this is due to UBR4 presence or activity. It is possible that the proteasome may be preferentially found in an ubiquitinated state when bound to UBR4, and the reason for this needs to be further investigated. My results, consistently showing ubiqutin signal associated with proteasomal proteins in the presence of UBR4, suggest the possibility that such an association might have some physiological relevance.

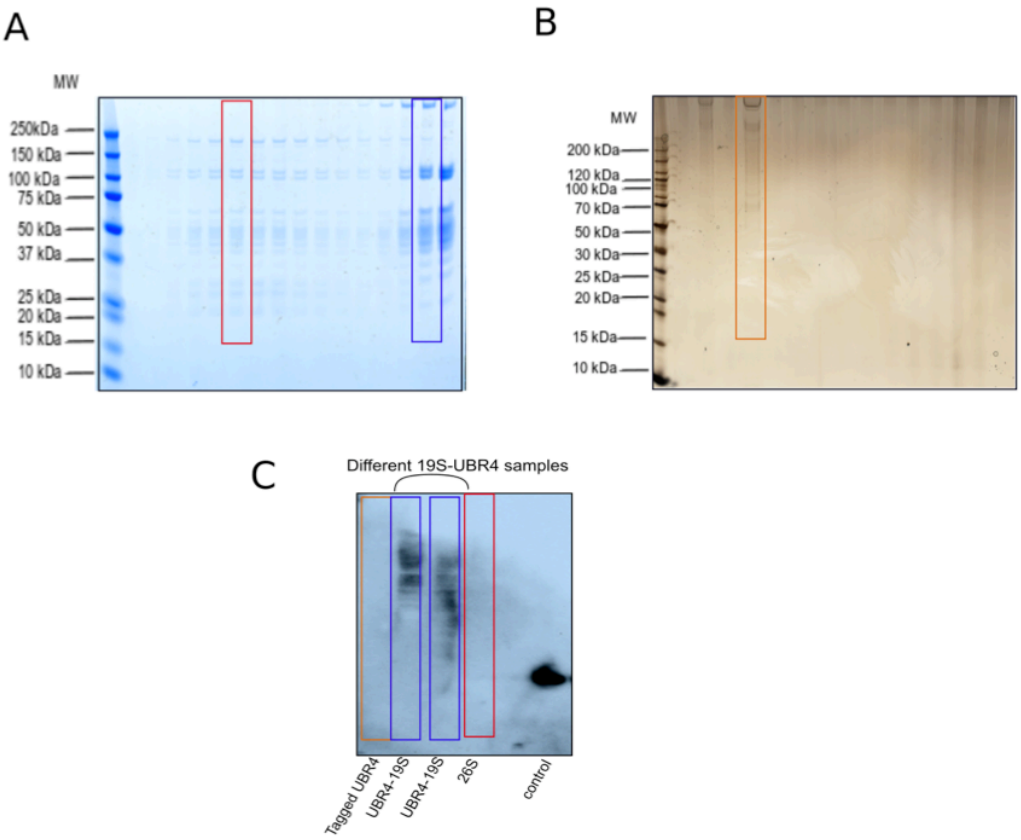


Fig. 3.18: Anti-ubiquitin western blot of UBR4 samples. Streptavidin-tagged UBR4, UBR4-19S and 26S samples were purified and separated by gel electrophoresis. The levels of ubiquitination were qualitatively measured with an anti-ubiquitin antibody and visualised by western blotting. (A) 26S and UBR4-19S were taken from a glycerol gradient centrifugation step as shown in a representative SDS-PAGE gel (B) and streptavidin-tagged UBR4 was taken from the glycerol gradient step as shown in a representative SDS-PAGE gel (C). The sample corresponding to the 26S proteasome (red box) showed low levels of ubiqutination, however UBR4-19S samples (blue boxes) showed stronger levels of ubiquitination. The streptavidin-tagged UBR4 sample only showed little to no levels of ubiqutin signal (orange box). Ub-40, a ubiquitinated protein, served as a control sample.

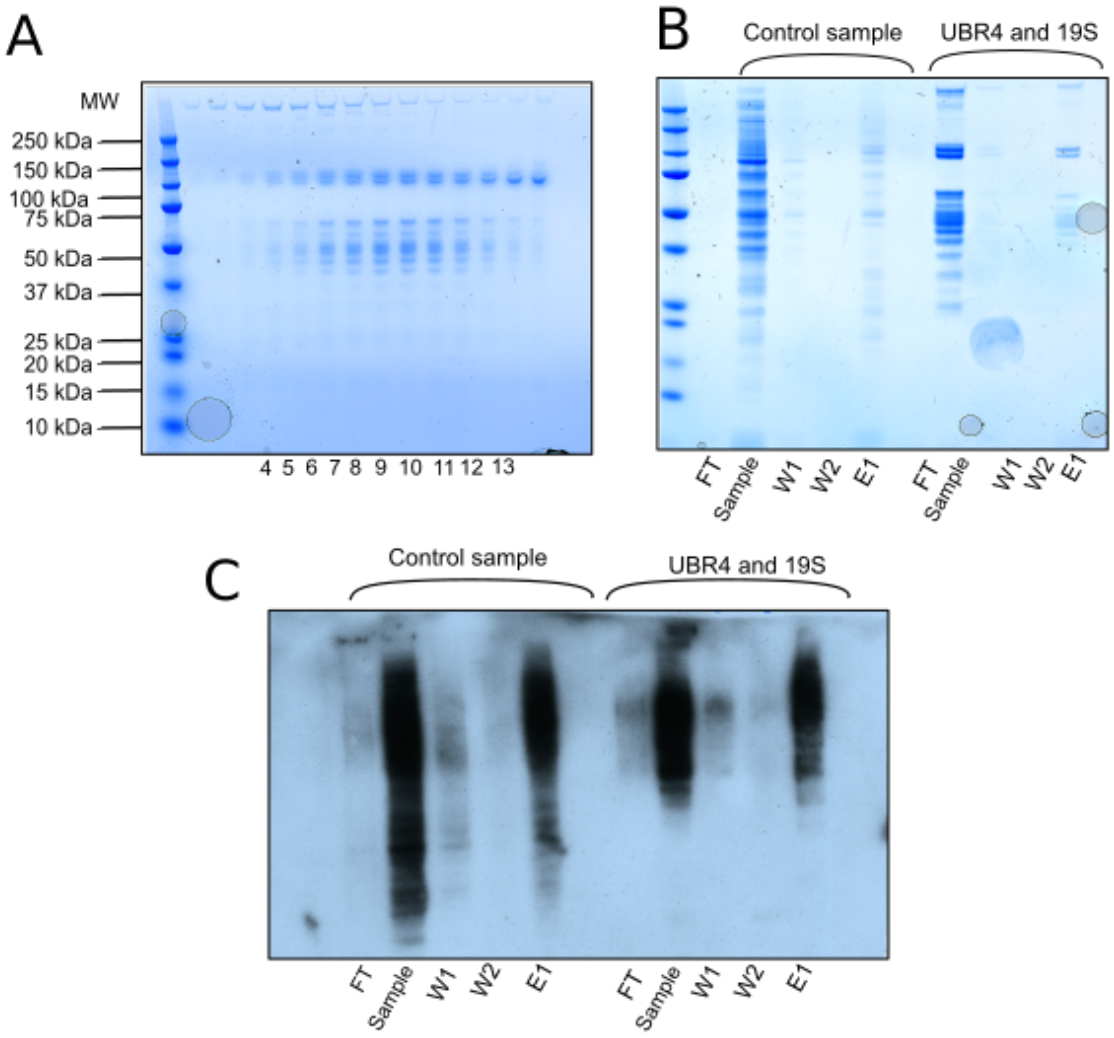


Fig. 3.19: UBI-capture experiments of UBR4-19S samples. A) UBR4-19S samples were purified as normal with the representative glycerol gradient SDS-PAGE gel shown. B) Samples were concentrated and put through a UBI-capture matrix that binds ubiquitinated proteins and subsequently visualised using SDS-PAGE. C) a western blot of wash steps and elutions showed ubiquitinated levels of the UBR4-19S samples. FT= Flow through, W1= wash 1, W2= wash 2, E1= elution 1.

3.5 Discussion

My results show that UBR4 co-elutes with the 19S proteasome subunits and that separating these two species proves to be very difficult biochemically. This strongly supports the notion that UBR4 may be forming a complex with 19S subunits. Interestingly, glycerol gradient fractions from glycerol gradients that predominantly contain 19S and UBR4 (see Fig. 3.8) do not appear to contain 20S proteasome subunits or UBL. Experiments by others also provide evidence that UBL does not interact directly with UBR4 but rather UBR4 is

associating with the proteasome. These include the fact that UBR4 has been co-purified with proteasomes from cell lines that have a tagged proteasome and therefore do not require UBL in the purification strategy (Besche et al., 2014). Further, treating proteasome fractions containing Rad23 and UBR4 with Usp-2, a Rad23 deubiquitinating enzyme, does not affect UBR4 binding to the proteasome (Besche et al., 2014). This data is therefore highly suggestive that UBR4 is not binding to UBL proteins but instead may be forming a complex with 19S proteasome specific proteins.

When UBR4-19S samples are visualised using negative stain EM, there are predominantly uncharacterised globular proteins that are, however, compatible with the size and plausible architecture of UBR4-19S particles (Fig. 3.12). Interestingly, there are also some 20S particles present, purified from tagged UBR4 cell lines (Fig. 3.16). Further, some double capped complexes from UBR4-19S samples from endogenous proteasome purifications can be observed and are likely a result of overlapping peaks and poor separation of all proteasome species in currently existing separation techniques (Fig. 3.12). It is puzzling why 20S proteasomes would be seen with tagged UBR4 purifications, whereas mostly 19S RPs are seen with untagged UBR4 samples, and this requires further exploration. It may be that UBR4 is binding at the interface of both the 20S and 19S particles, however, only 20S proteasomes are strongly recovered with tagged UBR4. However, these UBR4 could also be binding separately to both the 19S and 20S and the UBR4-19S or the UBR4-20S complexes could be subpopulations in which the UBR4 is acting as a scaffold and binding to both complexes. Tagging the N-terminus of UBR4 instead and probing any differences in co-eluted proteasome species would test the effect of UBR4's tag position. Subsequent image processing of negative stain images is difficult to interpret but initial 3D models yielded globular like structures that could be representative of UBR4-19S protein particles (Fig. 3.12).

Attempts to increase the yield and homogeneity of sample preparation by endogenously tagging UBR4 also showed proteasomes co-eluting with tagged UBR4, including 20S proteasomes (Fig. 3.15). Negative stain images of these samples similarly showed traces of proteasome complexes but the obtainable low yield remained a problem (Fig. 3.16). Future experiments should look into overexpressing tagged UBR4 in either mammalian or insect cells to facilitate increased yields of pure and homogenous UBR4 sample. Furthermore, overexpression of UBR4 would allow more intricate biochemical experiments such as

binding assays to proteasome complexes and consequently provide more insights into the UBR4-proteasome interaction.

Due to the sample impurity and low yield of UBR4, it is difficult to obtain functional data on UBR4. However, the SEC-MALS provide some clues that the putative UBR4-19S complex obtained indeed forms a complex with an overall molecular weight lower than the 26S proteasome (Fig. 3.11). It is worth noting here, that this could also be due to the elongated shape of the proteasome and anomalous behaviour in the size-exclusion column.

Western blot analysis indicates that UBR4 and ubiquitinated proteins co-eluted. However, the reasons the UBR4-19S samples are found in ubiquitinated states needs to be further investigated. For example, repeating such experiments with purer, and more concentrated sample is one requirement to obtain further insights into which proteins are interacting with UBR4. The UBI-capture experiments were conducted to identify which proteins were ubiquitinated on 19S and UBR4 samples with results showing high ubiquitination levels in UBR4-19S samples (Fig. 3.19). To identify whether individual proteasome subunits are ubiquitinated, western blots against different proteasome subunits, such as Rpn1 for example, could be conducted with the eluted fractions from the UBI-capture experiment. Further, my MS results indicating that another E3 ligase, Huwe1, is found in samples of UBR4-19S is interesting and worth investigating further, as UBR4 may be acting in concert with other E3 ligases that have ubiquitin ligase activity. There are examples of E3 ligases, such as the RING E3 ligases MdmX and Bmil (see Section 3.1), that require formation of heterodimers with other E3 ligases to confer E3 activity. This may also be the case for UBR4 and the fact that we see Huwe1 in UBR4-19S samples may be indicative of Huwe1 playing a supporting role with UBR4. To further investigate this, it will be important to demonstrate that Huwe1 indeed directly interacts with UBR4. Further, it would be interesting to conduct a native MS experiment on the purified UBR4-19S sample to approximately assess stoichiometry, composition and subunit identities of the complex.

Proteolytic activity assays on tagged UBR4 samples show ß5 and ß1 activity within the sample and this is comparable to ß5 and ß1 activity of 0.01 µM of 20S (Fig. 3.17). The basis of ß5 and ß1 activity within UBR4 corresponds to the observation of 20S proteasomes within the UBR4 sample identified in negative stain. This result also suggests that the proteasomes

seen within the UBR4 sample might be in an active form due to their activity levels, and therefore UBR4 does likely not have an inhibitory role on proteasome complexes.

The biological significance of UBR4 binding to the proteasome is intriguing. Whether UBR4 acts as a scaffold of ubiquitinated proteins at the proteasome and improves efficiency of their degradation, or whether it plays a role in the regulation of the proteasome itself are all interesting ideas but have yet to be elucidated. Further, what remains intriguing is why some E3 ligases (Ube3a/E6AP, Ube3c/Hul5, Rnf181, Huwe1, and UBR4) are found to co-elute with the proteasome as mentioned earlier. In addition, investigations into other E3 ligases found co-eluting with UBR4, such as Huwe1, will allow to dissect UBR4's role as an active E3 ligase. These experiments will provide a more precise molecular understanding of the role of UBR4 as an enzyme and a cofactor of the 26S proteasome. The experiments I have performed, have established a foundation to further investigate these hypotheses. UBR4 is a challenging protein to work with due to the unavailability of its cDNA and the lack of its biochemical role in the cell. My experiments have laid a foundation to overcome such experimental difficulties and provided a challenging step towards understanding UBR4.

Chapter 4 P97 and the 26S Proteasome

4.1 Introduction

4.1.1 P97 overall cellular roles

P97 is a highly abundant member of the AAA+ ATPase superfamily that is conserved in eukaryotic cells, including in humans and yeast (Peters et al., 1990). AAA+ ATPases are members of a protein family that contain canonical motifs required for ATP binding and hydrolysis (Hanson and Whiteheart, 2005). Members of this family are associated with diverse cellular roles which include membrane fusion, chromatin-associated processes, and endoplasmic reticulum-associated degradation (ERAD) (Stach and Freemont, 2017). P97, also known as vasolin-containing protein (VCP), Cdc48 (cell division cycle protein 48) in *S. cerevisiae*, and Ter94 (transitional endoplasmic reticulum ATPase) in *Drosophila melanogaster,* has been implicated in many cellular pathways since its discovery more than 25 years ago (Peters et al., 1990). These pathways include DNA repair, cell cycle regulation, the UPS, ERAD and the unfolded protein response (UPR) (Hwang and Qi, 2018). In all of its cellular roles, P97 converts energy through ATP hydrolysis to mechanical energy in order to extract ubiquitinated proteins from complexes or membranes (Stach and Freemont, 2017). Although P97 on its own has low affinity to ubiquitin, the interaction can be mediated through a range of cofactors (Stach and Freemont, 2017).

4.1.2 Structure of P97

P97 is a stable homo-hexamer, with each monomer comprising a globular N-terminal domain and two tandem ATPase domains, D1 and D2 (Fig.4.1 and Fig. 4.2) (van den Boom and Meyer, 2018).

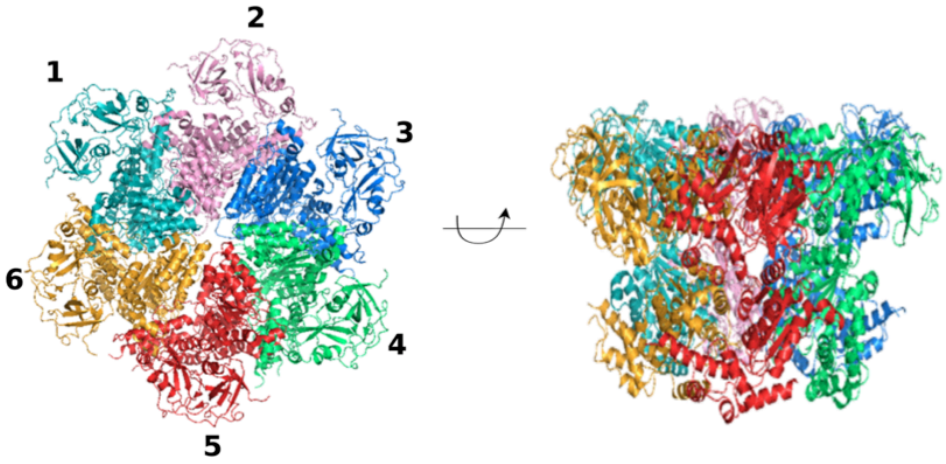


Fig. 4.1: P97 structure. P97 forms hexamers. A top view of the hexamer is shown in the left panel, and each monomer is numbered. A side view (right panel) of the P97 hexamer is also shown (PDB: 5FTK).

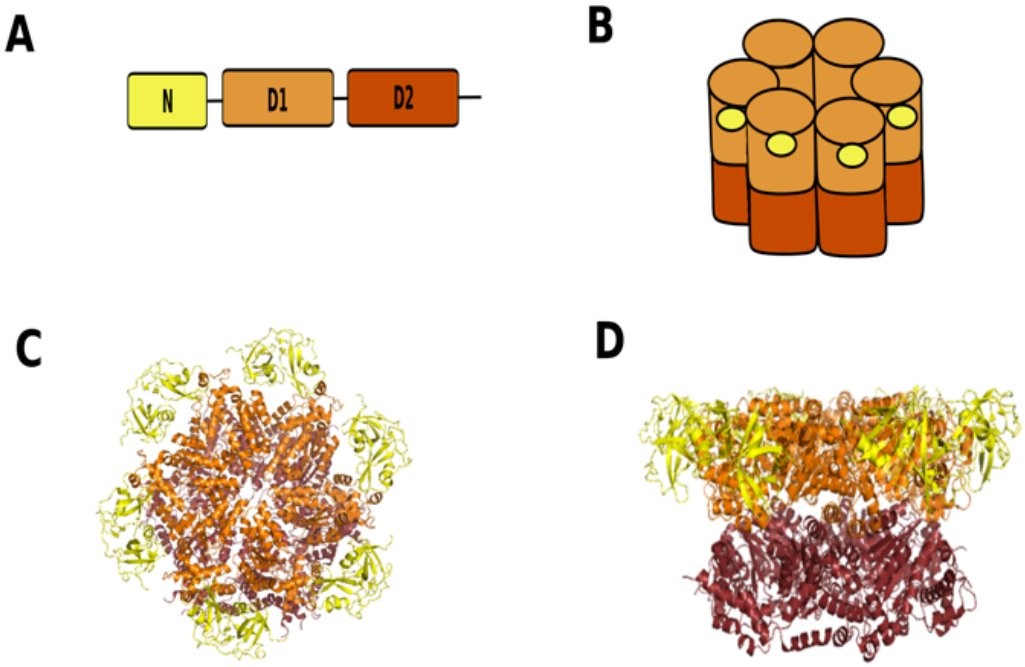


Fig. 4.2: P97 domain architecture. P97 is composed of three domains: A) N-terminal domain (yellow), and the ATPase domains D1 (orange) and D2 domain (dark orange). N, D1 and D2 domains are joined by linkers (black lines). B) A schematic representation of the P97 hexameric assembly. C) and D) The human P97 hexameric structure top view and side view, as obtained by cryo-EM (PDB: 5FTK).

D1 and D2 domains are similar both in structure and sequence and are connected by a short linker, called the D1-D2 linker (Fig. 4.2) (Ye et al., 2017). However, despite their

structural relatedness, D1 and D2 domains are functionally different. The N-terminal domain is linked to the D1 domain via a flexible polypeptide linker (named N-D1 linker) and there is a short C-terminal tail at the C-terminus of the D2 domain (Fig. 4.2) (Ye et al., 2017). Similar to other AAA+ ATPases, the D1 and D2 domains of P97 consist of a characteristic helical domain and a conserved RecA-like domain that features a nucleotide binding site (Fig. 4.3) (Ye et al., 2017). The D1 and D2 ATP binding sites are located within the RecA-like domain (Fig. 4.3) (Ye et al., 2017). The structure of full-length P97 has been solved by both X-ray crystallography and cryo-EM (Banerjee et al., 2016; Davies et al., 2008). The initial X-ray crystallography structures were limited to 3.5 Å but more recent cryo-EM studies improved the resolution to about 2.4 Å (Banerjee et al., 2016).

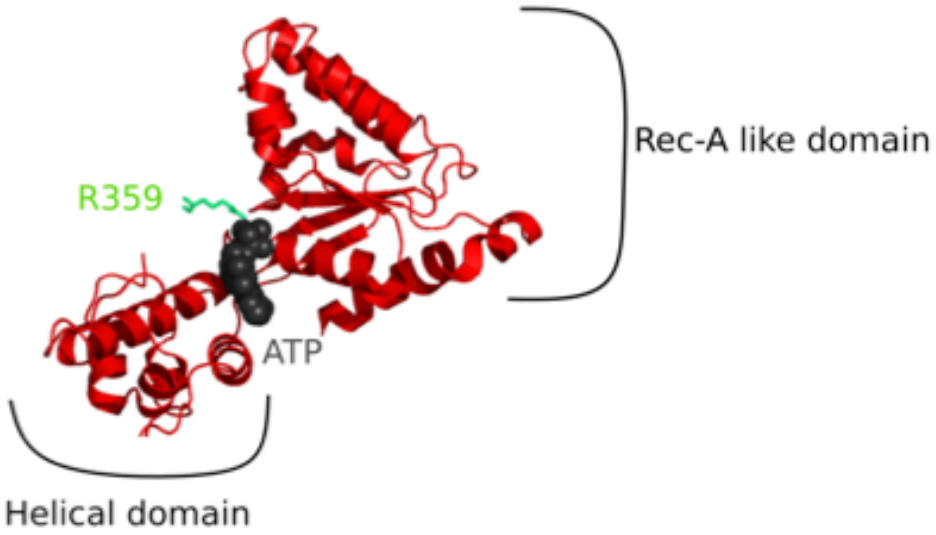


Fig. 4.3: Domain architecture of D1 AAA+ ATPase. P97 is an AAA+ ATPase in which each monomer consists of ATPases present in the D1 and D2 domains. The ATPases have a characteristic helical domain and a Rec-A like domain. The Rec-A like domain harbours a nucleotide binding site and binds ATP (grey circles). R359 from the neighbouring N-terminal domain promotes ATP hydrolysis by interacting with the ATP phosphates (PDB: 5FTK).

Earlier lower resolution cryo-EM studies, which had resolved the overall architecture of P97, demonstrated the rotational movement between the ATPase domains upon ATP hydrolysis (Rouiller et al., 2002). More recently, higher resolution P97 cryo-EM maps have revealed additional conformational changes and functions (Banerjee et al., 2016). The structures, solved bound to non-hydrolysable ATP analogues (ATPγS) at physiologically relevant concentrations, revealed three different conformational changes of P97 that occur when ATP binds. These conformations suggest that the mechanism of P97 activation is a

two-step process involving the D2 and D1 hexameric layers. ATP first binds to D2 and then to D1, resulting in significant movement of the N-terminal domain and thus a large conformational change (Fig. 4.4) (Banerjee et al., 2016). For purified wild-type P97, ADP binds to the D1 and D2 domains with a K_d of approximately 1 μM and 80 μM respectively, however, the ATP affinity for both domains is the same (approximately 2 μM) (Ye et al., 2017). The D2 domain is the major contributor for overall ATP activity, whereas the D1 domain acts as an assembly chaperone for the P97 hexamer and allows it to form (Ye et al., 2017). This has been shown in several studies where the D2 exhibits higher ATP activity than the D1 domain (Song et al., 2003). The N-terminal domain, on the other hand, mediates most of the interactions of P97 to its cofactors/binding proteins (Ye et al., 2017).

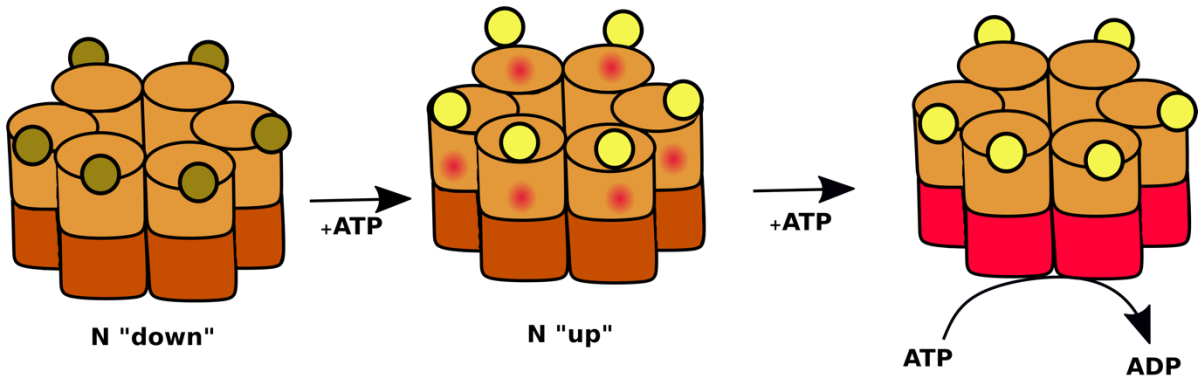


Fig. 4.4: Schematic diagram of P97 conformational changes. P97 N-terminal begins in the "down" conformation (dark yellow). Upon ATP nucleotide binding to the D1 domain (red circles), the N-domain adopts the so called "up" conformation and the D2 domain can subsequently hydrolyse ATP, thereby allowing for substrate unfolding and translocation through the pore present in the middle of the hexamer.

4.2 P97 and its cofactors

P97 associates with a wide variety of interactors, allowing it to participate in highly diverse cellular processes (Buchberger et al., 2015). Co-factors found to interact with P97 can help to recognise ubiquitinated substrates or, for example, contain catalytic activity for substrate processing (Hänzelmann and Schindelin, 2017). To date, there have been 30 known P97 cofactors identified and many more are thought to remain undiscovered (Arumughan et al., 2016). Described below are some examples of the role of P97 cofactors involved in a range of cellular processes.

4.2.1 The Ufd1-Npl4 (UN) Complex and ERAD

P97 is well-known for its role in ERAD. Here, P97 acts as a segregase in the protein-quality control process within cells, allowing exit of misfolded proteins from the ER (Torrecilla et al., 2017). Misfolded proteins are then tagged with ubiquitin and targeted for proteasomal degradation. The ER is a highly dynamic membrane-enclosed organelle that hosts several metabolic processes and plays an important role in protein synthesis and quality control (Lev, 2012). Failure to remove misfolded proteins can be detrimental for the cell; intracellular accumulation of toxic protein aggregates underlie several human diseases (Guerriero and Brodsky, 2012). In fact, more than 60 human diseases have been linked to the dysfunction of the ERAD pathway, such as neonatal diabetes (Type 1) and X-linked Charcot-Marie-Tooth disease and (Guerriero and Brodsky, 2012).

One aspect of ERAD that remained elusive for years was the way in which P97 was able to generate the energy necessary for protein unfolding (Ye et al., 2017). Force generation remains a highly contested issue. However, one of the most agreed mechanisms proposes the force is created by the cooperation between D1 and D2 rings (Ye et al., 2017). The most observed conformational changes are the D2 rotation and the swing movement of the N-terminal domain from the “down” into the “up” conformation (Fig. 4.4). These movements are linked to ATP hydrolysis in the D1 domain (Fig. 4.4) (Ye et al., 2017). The force exerted by P97 on a substrate is used to channel and unfold it through a central tunnel (Fig. 4.5). In ERAD, proteins that have failed to fold in their correct native state are degraded by a P97-dependent mechanism (Stevenson et al., 2016).

During ERAD, P97 has to bind specific cofactors (Brodsky, 2012). Two P97 cofactors that are integral to the ERAD pathway are the 68 kDa nuclear protein localisation protein 4 (Npl4) and the 34.5 kDa ubiquitin recognition factor in ER-associated degradation protein 1 (Ufd1), which together form the UN heterodimer (Fig. 4.5). The UN binds to P97 and acts as an adaptor to regulate the P97 ubiquitin-dependent substrate processing (Isaacson et al., 2007). Ufd1 and Npl4 are both essential for viability in eukaryotes (Ye et al., 2001). Ufd1 contains an N-terminal ubiquitin binding domain and a P97 binding domain, called BS1. Npl4, conversely, contains an N-terminal domain that binds to P97 and a C-terminal domain that binds ubiquitin (Meyer et al., 2012).

In the current model for P97 substrate unfolding, the UN complex binds to the N-terminal domain of P97 (Fig. 4.5). Upon binding of the UN complex, the P97-UN complex becomes competent for ERAD substrate processing. Poly-ubiquitinated substrates bind first to the UN complex and are then inserted into the central pore of the P97 hexamer, where ATP hydrolysis by the P97 D2 domains provides the necessary energy for substrate unfolding and movement through the P97 central pore (Fig. 4.5) (Bodnar and Rapoport, 2017). Subsequent ATP hydrolysis in the D1 domain coupled with trimming of the polyubiquitin by a deubiquitinase completes the translocation process through the P97 central pore (Bodnar and Rapoport, 2017). A likely candidate for the deubiquitinase activity is the ubiquitin thioesterase (YOD1) deubiquitinase (Ernst et al., 2009; Stein et al., 2014). It is currently not exactly known how a polyubiquitinated substrate is passed on to downstream components. One model proposes that the polyubiquitinated substrate is deubiquitinated during P97 processing and then subsequently re-ubiquitinated for proteasome binding (Liu and Ye, 2012). A second model proposes that substrates are not deubiquitinated during ERAD and are instead sent to the proteasome (Richly et al., 2005). And finally, a third model suggests that substrates are only partially deubiquitinated so remaining ubiquitins can interact with downstream factors, and mark the substrate for proteasomal degradation (Bodnar and Rapoport, 2017).

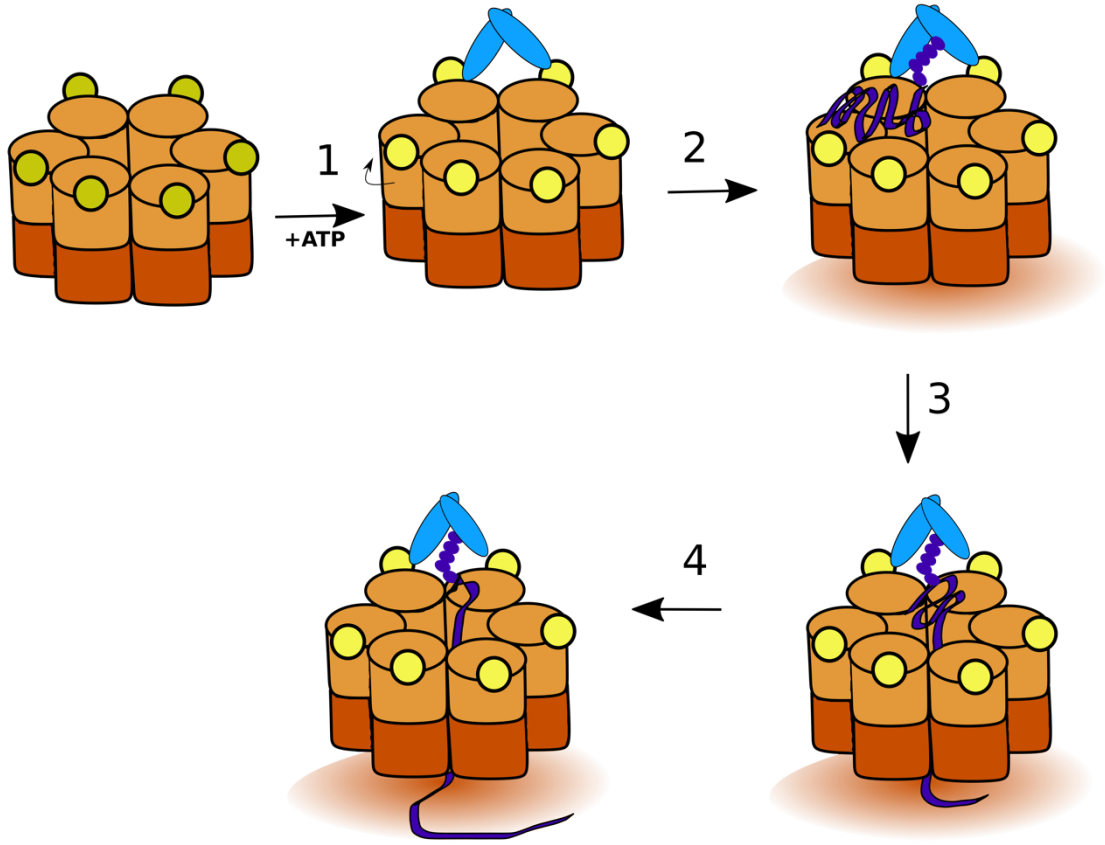


Fig. 4.5: Current model for P97-UN in ERAD. P97 N-domains begin in the "down" conformation and transition to the "up" conformation upon D1 ATP binding. Once in the "up" conformation, the heterodimeric UN complex (composed of the Ufd1 and Npl4 heterodimer) (blue) binds to the N-domains (yellow) (Step 1). Ubiquitinated substrates (purple) can then bind (Step 2). In step 2, the N-domain stays in the "up" conformation with the D1 domain bound to ATP. Meanwhile, the ATPase rate of the D2 domain increases (orange outline) and this causes the substrate to be inserted into the central pore as shown in step 3. ATP hydrolysis by the D2 domain promotes unfolding of the substrate and translocation through the pore to the other side.

4.2.2 Cofactors of P97 with UBX domains

The ubiquitin regulatory X domain (UBX) is a ubiquitin-like fold found in some proteins and shown to bind to ubiquitin receptors on the proteasome (Buchberger et al., 2001). Known P97 cofactors with these domains include the Fas-associated factor 1(FAF1)-UBX complex and the UBX domain-containing 7 (UBXD7)-UBX complex, which have roles in cell cycle pathways and mitophagy respectively (Hänzelmann et al., 2011; Li et al., 2017).

A common theme amongst these cofactors is that they interact with the N-terminal domain of P97 through a conserved signature motif, however, the functional significance of this remains elusive (Hänzelmann and Schindelin, 2017).

4.2.3 Cofactors of P97 with VIM and VBM domains

VCP interacting motif (VIM) and VCP binding motif (VBM) domains consist of mainly positively charged amino acids that adopt a helical conformation (Hänzelmann and Schindelin, 2017). P97 has been shown to bind to the VBM domain of Rhomboid-related protein 4 (RHBDL4) through its N-terminus, which plays an important role in the retro-translocation of ubiquitinated proteins in ERAD (Fleig et al., 2012).

4.3 P97 and the 26S Proteasome

Many proteasome-associated proteins have been co-purified with the 26S proteasome, including P97 (Besche et al., 2009). For example, P97 has been shown to immunoprecipitate with proteasomes in both human B cell lines and yeast (Dai et al., 1998; Verma et al., 2000). More recently, Bersche and colleagues also identified P97 co-eluting with the proteasome using the UBL-proteasome purification described in Chapter two (Besche et al., 2009). The purification method involving the UBL domain to purify the 26S proteasome is also discussed in detail in Chapter two. With this method, P97 and other P97 adapters such as the UN complex, p47 and FAF1 were observed to co-purify (Besche et al., 2009). P97 and adaptor proteins were also identified to co-migrate with the 26S proteasome during glycerol gradient purification (Besche et al., 2009). Although both P97 and the 26S proteasome interact with poly-ubiquitin chains, no ubiquitin conjugates were detected in the glycerol gradient purification but both proteasomes and P97 were observed. Moreover, studies showing archaeal 20S activation by a related P97 AAA+ ATPase and P97-20S complexes have prompted discussions on whether or not P97 is a possible alternative cap of the proteasome (Barthelme and Sauer, 2012; Pick and Berman, 2013). However, in humans only 26S proteasomes have been shown to co-elute with P97 and not 20S proteasomes (Isakov and Stanhill, 2011; Meyer et al., 2012). Matouschek and Finley have suggested two potential models that would explain P97's roles within the eukaryotic proteasome (Matouschek and Finley, 2012). The first model suggests P97 could be acting as an alternative AAA+ ATPase for a specific subset of proteasomes that has yet to be

discovered, while the second hypothesis postulates that the P97-20S interactions have been lost over time despite it being an initial part of the degradation machinery (Matouschek and Finley, 2012).

More recently, research revealed two endogenous substrate proteins for the yeast ortholog Cdc48-20S proteasome complex (Islam et al., 2020). These substrates, the copper-zinc superoxide mutase Sod1 and the threonyl tRNA synthase Ths1, are essential for the normal functioning of yeast cells, and are degraded by a ubiquitin-independent mechanism without the need for any post-translational modifications (Islam et al., 2020). The recognition signal could instead be one or more signal motifs. This study is one of the first to identify Cdc48-20S specific substrates, however, it is reasonable to speculate that there are many others. Exactly how this complex is formed and its significance remains elusive, however, it has been proposed that Cdc48-20S complexes could serve as a backup proteasome needed under cellular stress conditions (Islam et al., 2020).

It remains unclear how exactly ubiquitinated substrates are transferred to the proteasome after ERAD. This may either be mediated through a direct interaction between P97 and the proteasome or via a P97 cofactor bridging the two together (Raasi and Wolf, 2007). For example, P97 may be directed to the proteasome via shuttle factors, such as Rad23 (or other UBL-UBA containing proteins) (Raasi and Wolf, 2007). It has been shown that P97 can bind to both Rad23 and other ubiquitin conjugates suggesting that there may be an interplay between Rad23 and ubiquitin conjugate delivery to the proteasome via P97 (Besche et al., 2009). However, it has also been shown that P97 possesses a higher affinity for the proteasome than for the UBL domain of Rad23 (Raasi and Wolf, 2007). Interestingly, P97 cofactors pulled down with the proteasome using the UBL method are upregulated in Drosophila during proteasome inhibition (Lundgren et al., 2003). Finally, these cofactors vary with the cell type; proteasomes isolated from HEK293F cells using a Rad23 bait did not pull down P97-cofactor complexes or ubiquitin conjugates, despite showing a 20% increase in proteasome isolation yield (Scanlon et al., 2009). Thus, it is important to account for differences between different cell lines, tissues and the experimental conditions used in those pull-down assays.

4.4 P97 in human disease

Due to the fundamental and various roles P97 fulfils within cells, it is not surprising that there are serious clinical implications associated with P97 mutations. Generally speaking, P97-associated diseases are degenerative (Kimonis et al., 2008). In mouse models, the most common mutation, R155H, leads to diseases such as inclusion body myopathy, Paget's disease of bone, frontotemporal dementia and amyotrophic lateral sclerosis (ALS) (Badadani et al., 2010). Further, knock out of P97 is embryonic lethal for mice underpinning P97's crucial role in embryonic development (Badadani et al., 2010). Interestingly, no mutant described so far impedes P97's ability to oligomerise, implying that oligomerisation is mandatory and such mutations are not tolerated. However, some mutations do affect the rate of ATP hydrolysis (Niwa et al., 2012; Weihl et al., 2006), while others affect the ability of P97 to bind to its cofactors (Fernández-Sáiz and Buchberger, 2010; Ritz et al., 2011). In general, diseases with P97 mutations reduce or impair its function, rather than abrogate it.

4.4.1 P97 and neurodegenerative diseases

Most mutations in P97 that give rise to neurodegenerative disorders are caused by the inadequate removal of substrates during protein degradation, leading to an increase in protein aggregate formation (Huryn et al., 2019). Among these, more than 20 autosomal dominant mutations in the gene encoding P97 cause a rare degenerative disorder called inclusion body myopathy with early-onset Paget disease and frontotemporal dementia (IBMPFD)/ALS (Meyer and Weihl, 2014). These mutations (which include I27V, R93C, G97E) are found in the N-terminal region of P97, with a strong accumulation in the N-D1 linker (Meyer and Weihl, 2014). Interestingly, most diseases caused by P97 mutations are located in the N-D1 interface. However, it remains contested whether P97 inhibitors or activators would be most therapeutically beneficial for IBMPFD/ALS patients. That is, whether these mutations cause an increase in P97 activity (dominant active effect) or reduced P97 activity (dominant negative effect) (Huryn et al., 2019). Other examples of neurodegenerative diseases caused by P97 mutations include Charcot-Marie-Tooth disease (CMT2Y) and Huntington's disease, however, this remains a poorly understood area and warrants further research (Fujita et al., 2013; Gonzalez et al., 2014).

4.4.2 Other diseases involving P97

P97 mutations have also been implicated in some cancers and viral infections. Examples of cancers linked to P97 mutations include human melanomas and breast carcinomas (Woodbury et al., 1980). Cancer cells are known to depend more on protein degradation systems than normal cells, and P97 has been shown to be overexpressed in many cancers types (Yamamoto et al., 2005; 2004). P97 has also been shown to contribute to viral replication, as in the case of poliovirus (Arita et al., 2012).

4.4.3 P97 inhibitors

Given the strong correlation of P97 mutations in the development of certain cancer types and other diseases, P97 is considered an important potential therapeutic target. As mentioned, certain cancerous cells upregulate protein degradation and rely heavily on P97. Inhibiting P97 leads to elevated apoptosis of cancerous, rather than healthy cells (Deshaies, 2014). So far, a range of inhibitors of P97 ATPase activity have been interrogated with promising therapeutic potential in cancer therapy (Huryn et al., 2019). However, concerns about specificity and cytotoxicity from off target effects have been expressed (Tang et al., 2019). Moreover, P97 inhibitors may also be promising treatments for antiviral therapy but this has yet to be explored.

4.5 Results

4.5.1 Initial observations of P97 in 20S preparations

In order to understand and thoroughly investigate the proteasome interaction network, one must be able to prepare these proteins in relatively high concentrations and obtain relatively homogenous samples that can be used in a range of biochemical and structural analyses. Multi-subunit complexes can be purified directly from endogenous sources such as mammalian cells, however, this usually results in low yields, depending on the natural protein abundance (Jarvis, 2009; Stowell et al., 2016). Recombinant protein expression in insect cells offers a practical alternative to produce high yields of target complexes, as previously described in Chapter two. Moreover, the recombinant expression of target

complexes allows for easy manipulation, such as generation of point mutations, compared to the complex modifications of endogenous samples.

Based on the observation by Dr. Toste Rêgo from the da Fonseca lab, that recombinant immuno-proteasomes (i20S) co-purify with endogenous P97 from insect cells, when over-expressed using the baculovirus over-expression system, I decided to further investigate this interaction. MS analysis consistently identified P97 as a proteasome interacting partner alongside many other known interactors, such as 19S subunits, 20S chaperones and other binding partners, including PA200 and PA28 (Rêgo and Fonseca, 2019). The main goals of this project were to establish a complete and systematic study of the proteasome interaction network, and to understand the underlying mechanism of the P97-proteasome interaction. Specifically, I wanted to understand whether P97 interacts directly with the proteasome and whether other components are needed. Understanding the components involved in this complex formation might indicate whether this interaction was independent of the ERAD pathway.

Firstly, I wanted to re-investigate previous preliminary observations from our lab, where P97 co-purified with recombinant i20S proteasomes. To do so, I cloned and generated a bacmid of untagged P97 using human P97 plasmid as a template (described in Chapter 2) and co-infected insect cells with three baculoviruses encoding untagged P97, 20S chaperones and Twin-Strep tagged i20S. My results were consistent with the previous observations in the lab that P97 co-elutes with the 20S proteasome, which was confirmed using MS analysis (see Fig. 4.6B).

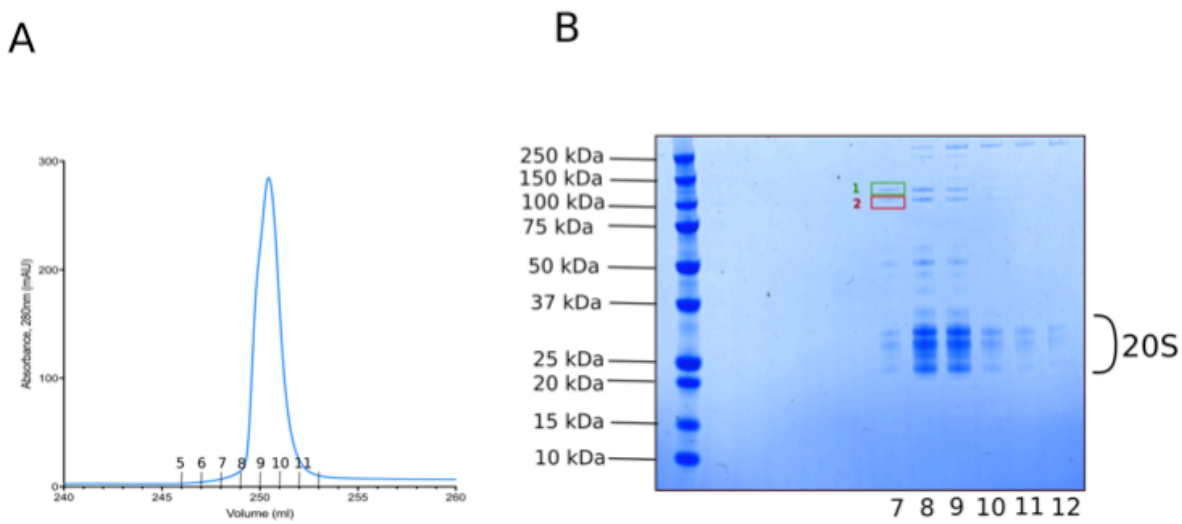


Fig. 4.6: Purification test of i20S proteasome. Streptavidin-tagged i20S proteasome was purified using streptavidin affinity chromatography in which A) a representative chromatogram and B) a SDS-PAGE is shown. The two bands corresponding to P97 size (green and red boxes) were sent for MS analysis, positively identified P97.

4.5.2 Co-expression of P97 with constitutive and immunoproteasomes

Next, I decided to investigate whether P97 only selectively co-eluted with i20S proteasomes, or whether constitutive 20S proteasomes also co-eluted with P97. I therefore co-expressed untagged P97 with either Twin-Strep tagged 20S and 20S chaperones or Twin-Strep tagged i20S-chaperones in similar quantities. I then compared P97 presence and the difference in yield obtained (Fig. 4.7). The yield of proteasomes was similar in both experiments (Fig. 4.7). SDS-PAGE did not show evidence for P97. However, MS analysis identified the presence of P97 in both 20S and i20S preparations, suggesting that P97 binding to proteasomes, or proteasome complexes involving additional factors, is not specific to either the constitutive 20S or i20S proteasome. Due to the similarity of yields of 20S proteasomes seen in the presence of P97, I decided to conduct further experiments with constitutive 20S proteasomes only as this is expressed more abundantly within cells (Morozov and Karpov, 2019).

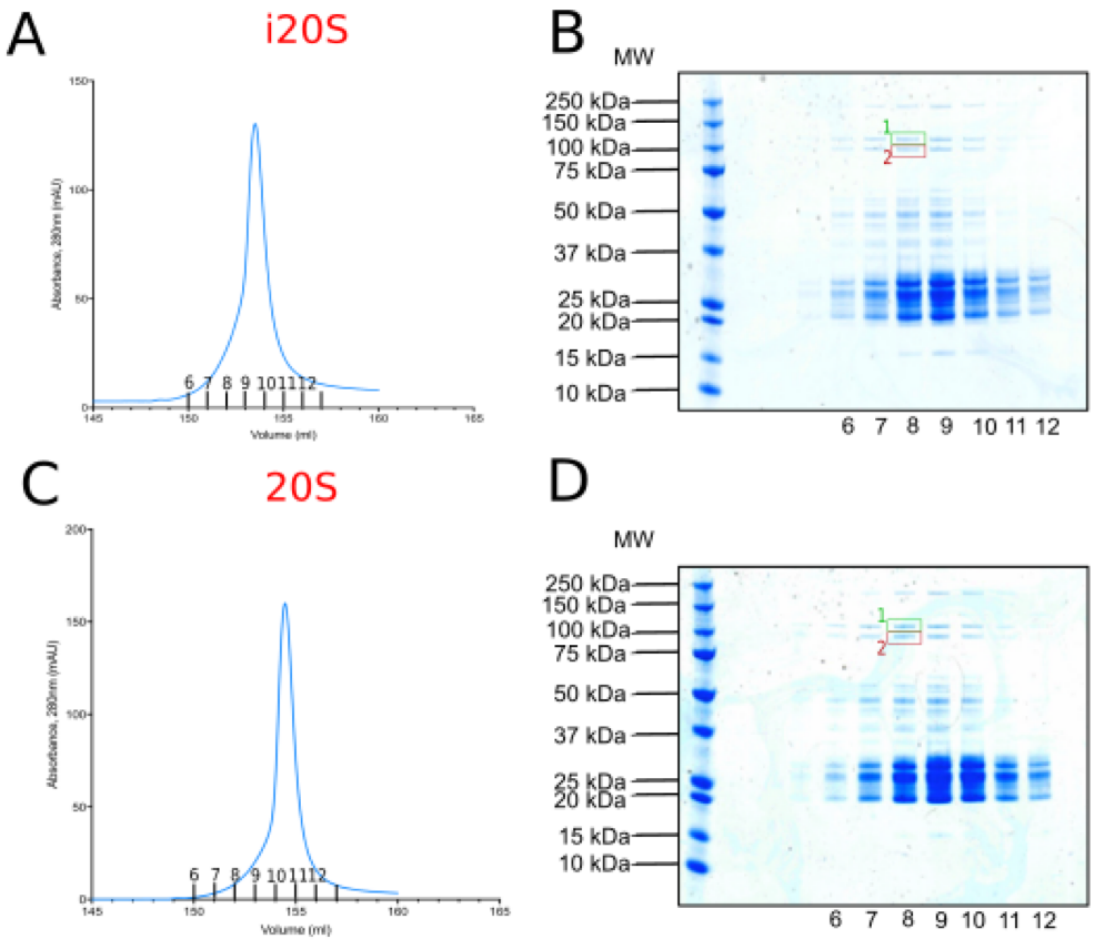


Fig. 4.7: Purification of tagged i20S and 20S proteasomes co-expressed with P97. i20S (in which the β7 subunits were Streptavidin- tagged) was purified using streptavidin affinity chromatography in which the A) representative chromatogram and B) representative SDS-PAGE are shown. Upon MS analysis, P97 was shown to co-elute with i20S protein (green and red boxes). Panels C and D are the same as Panels A and B, except this time recombinant 20S was used instead of i20S.

As mentioned, the P97 and proteasome co-elution has been reported by other labs (Besche et al., 2009). However, the true nature and significance of this interaction remains elusive, and there are no current complex structures in mammals available. I, therefore, wanted to investigate if addition of ATP to the purification was needed to improve the yield of P97-20S complex formation. Since P97 is a AAA+ ATPase and ATP is needed for a variety of its functions, it is conceivable that ATP may also be required for proper P97-20S complex assembly. To test this hypothesis, I performed the standard P97, 20S and 20S chaperone co-expression and attempted to purify the complex in the presence of 5 mM ATP using Twin-Strep tagged 20S proteasome as bait. As shown in Fig. 4.8, I could not detect any significant difference between the amount of 20S proteasome and P97 captured. While P97 still co-eluted with the 20S proteasome in the presence of additional ATP (confirmed

by MS analyses), no enrichment in P97 concentration was observed. This suggested that P97 may be associating with the 20S in an ATP-independent manner.

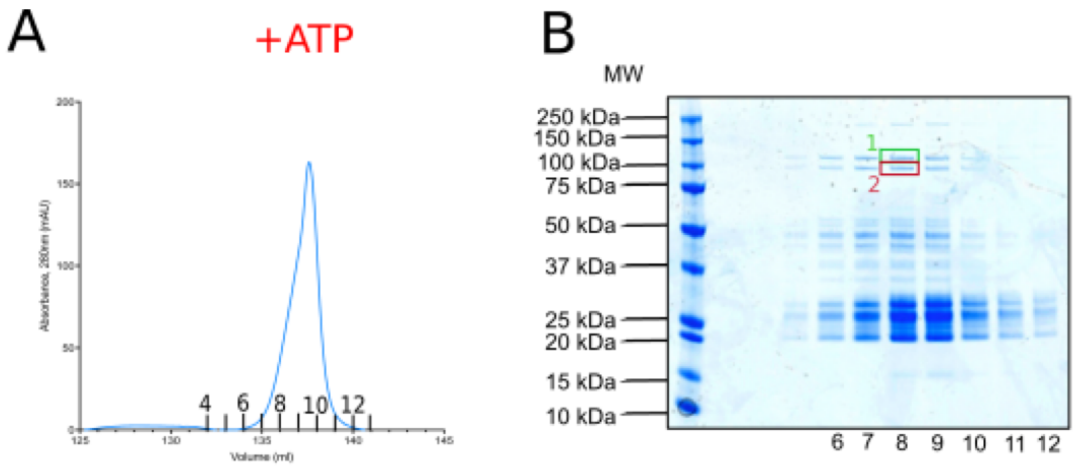


Fig. 4.8: Purification of P97 and streptavidin-20S with ATP. P97 was co-expressed with streptavidin-tagged 20S and 20S chaperones and ATP. Panels A and B show a representative chromatogram and SDS-PAGE gel respectively. No difference in the amount of P97 could be observed compared to previously used purification conditions (red and green boxes and see Fig. 4.6).

4.5.3 P97 purification

Since the yield of P97 that co-purified with both 20S and i20S was low, I tested the efficiency of P97 overexpression on its own. To do this, P97 was fused to an N-terminal Twin-Strep tagged and captured on a streptavidin column, which was followed by size exclusion chromatography (Fig. 4.9). The rationale for an N-terminal tag location was that the C-terminal contained regions that could be important for binding to the proteasome and so tagging the C-terminus may prevent binding to the proteasome. Using this affinity purification method, I was able to capture large amounts of P97 (approximately 2.5 mg of total protein from 2L of SF9 insect cell culture), with relatively high purity as assessed by SDS-PAGE (Fig. 4.9). In subsequent size exclusion chromatography, P97 eluted at a volume in agreement with hexamer formation (approximately 534 kDa). As discussed previously, physiologically P97 forms hexamers and thus this experiment suggested that the purified P97 retained its functional and biologically relevant form.

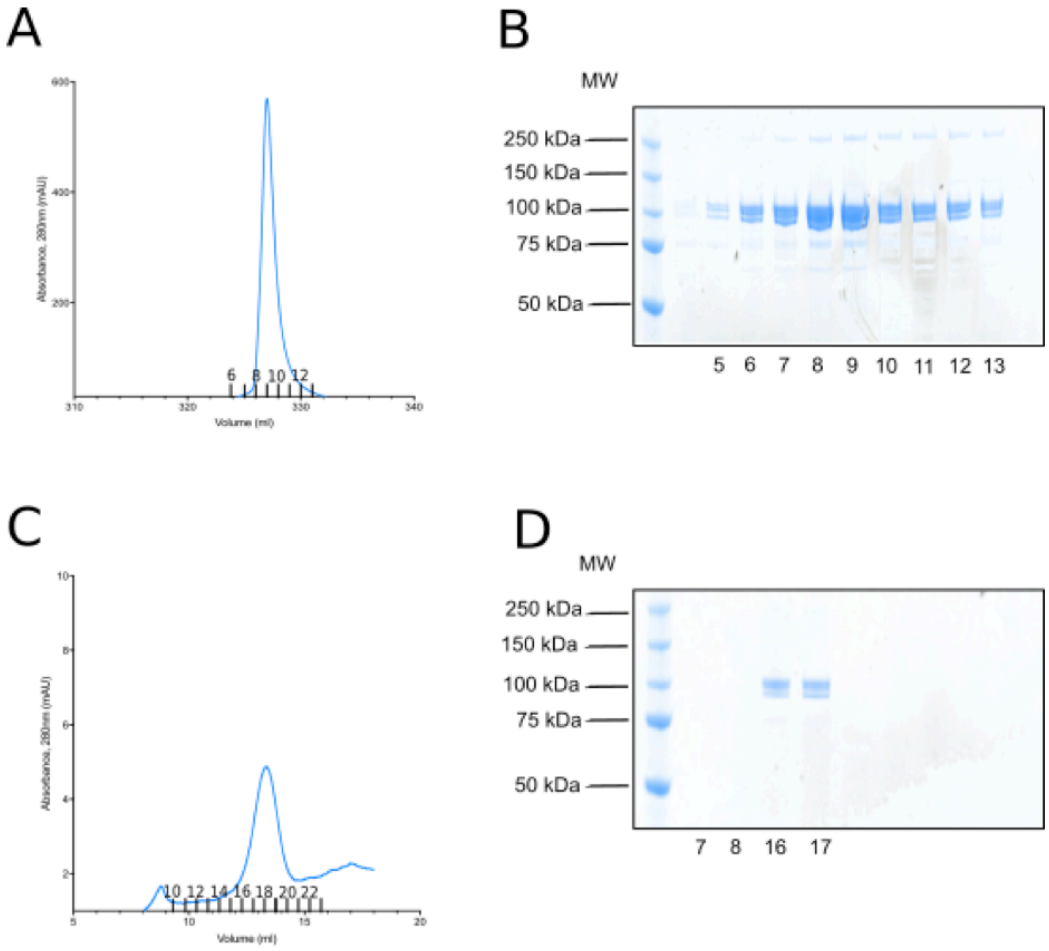


Fig. 4.9: Purification of P97. Streptavidin-tagged P97 was purified using streptavidin affinity chromatography. Panels A and B show the chromatogram and representative SDS-PAGE gels from the first step of purification, respectively. Size exclusion chromatography was then performed on fraction 9 from A), in which the representative C) chromatogram and D) SDS-PAGE gel are shown. The sample eluted from the size exclusion column at a volume expected for a hexamer (approximately 534 kDa).

4.5.4 Co-expression of proteasomes with affinity tagged P97

So far, MS analysis had shown evidence of P97 co-purifying with the 20S and i20S. As mentioned, these experiments involved pulling on either affinity-tagged 20S or i20S co-expressed with untagged P97. However, several other proteins were also seen in MS analysis and a low yield of P97 was observed. In an attempt to improve the yield of potential P97-20S proteasome complexes, N-terminal Twin-Strep tagged P97 was co-expressed with untagged 20S and 20S chaperones instead. The rationale behind this approach was, that pulling down P97 rather than 20S would reduce other proteasome-interacting proteins seen previously in SDS-PAGE gels. As an initial attempt, an affinity purification was conducted using Twin-Strep tagged tagged P97 as bait (Fig. 4.10A). SDS-

PAGE of the eluted sample showed some smears of typical proteasome size, but the presence of proteasomes was difficult to judge (Fig. 4.10B). The samples (fractions 6-12 shown in Fig. 4.10B) from the first purification step were therefore concentrated to try and improve the identification of 20S subunits (Fig. 4.10C). After sample concentration, some clearer protein bands were observed, which were later shown by MS to be proteasome subunits. However, at this point it was clear that complex formation yield was still very low and that a change of strategy was required.

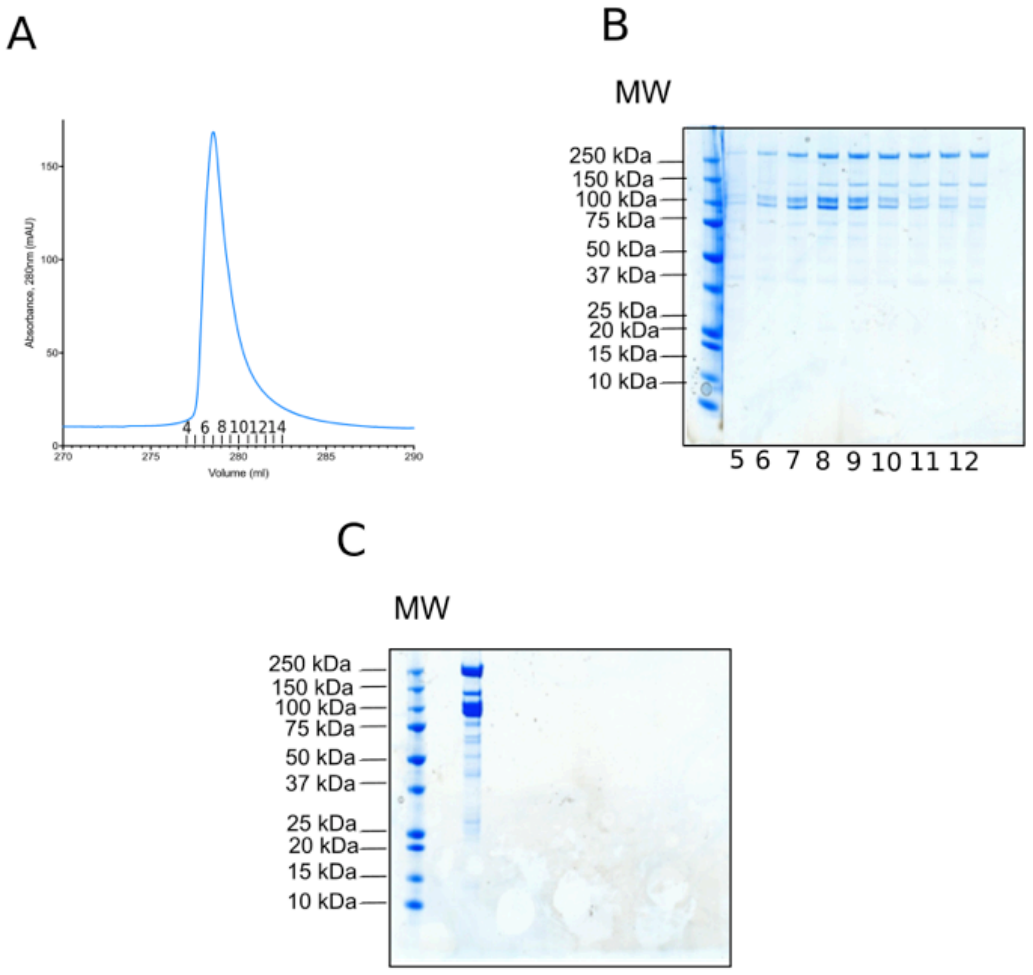


Fig. 4.10: Purification of streptavidin-tagged P97 co-expressed with 20S. Streptavidin- tagged P97 was co-expressed with 20S proteasome and 20S proteasome chaperones and pulled-down by streptavidin affinity chromatography in which A) a representative chromatogram and B) SDS-PAGE gel is shown. C) Fractions 6-12 from Panel B were pooled, concentrated and separated on SDS-PAGE to identify other proteins by MS analysis. There were other interacting proteins identified and all 20S proteasome subunits with some 19S subunits were identified in the MS data.

4.5.5 Codon optimisation of P97

Having shown that P97 and 20S proteasomes can co-purify from cells, albeit at low concentrations, I wanted to establish conditions that facilitated P97-20S complex formation within cells. This could also allow the identification of other protein partners that may be part of the P97-20S complex *in vivo*. Here, improving the yield and quality of P97 sample seemed to be key. As seen in Fig. 4.9, two distinct bands were present in the P97 sample and both had been identified as P97 by MS, with the lower band representing a P97 fragment. Possible reasons for the P97 truncations observed could be the presence of endogenous proteases that cleave unstructured N- or C- termini or premature termination during protein synthesis. I, therefore, decided to codon optimise the P97 gene. Some codons are favoured in specific organisms over others (Kurland, 1991). Although the underlying reasons for codon bias are not very well understood, it has been shown that codon optimisation can significantly improve the protein yield in recombinant protein expression. This is especially important when using a heterologous expression system, such as utilising insect cells to express the human 20S-P97 complex, where codon usage is optimised towards the host system. Codon-optimised human P97 gene was therefore used for expression in insect cells.

I also decided to modify my strategy for infection of insect cells. This involved using the cre-lox recombination system to enable both the 20S proteasome chaperones and P97 to be incorporated into a single baculovirus (from here on, this is referred to as P97-20S chaperones) (Ray et al., 2000). As mentioned in Chapter two, cre-recombinase is an enzyme from bacteriophages that enables the site-specific recombination of DNA between lox P sites. By incorporating P97 and the 20S chaperones into one viral vector, two viruses rather than three would be needed for infecting insect cells, thereby increasing the chances of co-infection and co-expression of the desired proteins.

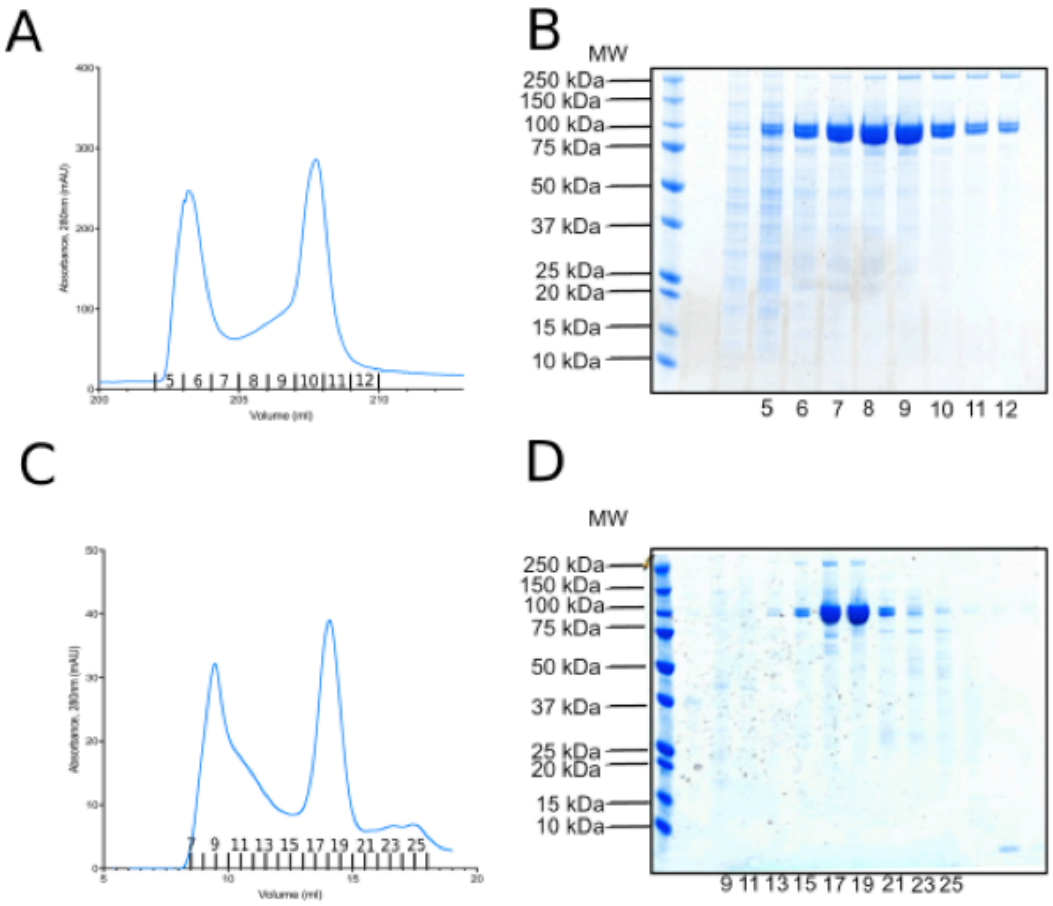


Fig. 4.11: Purification of codon-optimised streptavidin-tagged P97 co-expressed with 20S. P97 was codon optimised and cloned with 20S chaperones into a single baculovirus. Proteins were purified by pulling on the streptavidin-tagged P97 protein using affinity chromatography as a first step. A and B) show the representative chromatogram and SDS-PAGE gel. Fractions 5-9 were pooled and run on a size-exclusion column where C and D) show the representative chromatogram and SDS-Page gel respectively. Fractions 10, 13 and 19 were sent for MS and proteasome subunits (20S proteasome α subunits and RPTs from the 19S RP) were identified. However, despite improved levels of expression of P97, proteasome subunits were still difficult to identify on the gel.

Using codon-optimised P97 and incorporating the 20S chaperones in the same vector, I co-expressed untagged P97-20S chaperones with Twin-Strep tagged 20S proteasomes in insect cells using the same purification protocol as described in Section 4.4.2 (Fig. 4.11). After a streptavidin affinity chromatography step, the yield of P97 significantly increased, when compared to previous experiments. An increased amount of proteins co-eluting with P97 could also be observed (Fig. 4.11A and B). Some of these proteins migrated at molecular weight levels typical for proteasome subunits and could also be confirmed by MS. To further separate these components, fractions 5-9 were pooled, concentrated and

SEC was conducted; fractions 10, 13, and 19 were subsequently sent for MS analysis (Fig. 4.11C and D). Although the MS results identified some proteasome subunits (20S proteasome α subunits and Rpts from the 19S RP), there was still no formation of P97-20S complex yielded at near-stoichiometric quantities.

4.5.6 Proteolytic activity of P97-20S samples

To check whether P97 could modulate 20S activity, through potential interactions, activity assays in which 20S and P97 were incubated together or separately were conducted. Due to P97's role as an AAA+ ATPase, and the elusiveness of P97's role in relation to the 20S proteasome, I decided to carry out activity assays in the presence and absence of ATP, to investigate the role of ATP on protease activity. Any background signal from the buffer was corrected with wells only containing buffer. Approximately 20 times more P97 (4 μM) was used compared to 20S (0.2 μM) to ensure full saturation of 20S with P97. Since the same concentrations were included in both samples (P97-20S versus 20S), any difference observed in activity would be a direct result of P97 ability to modulate 20S activity. My results show that, overall, P97 has negligible effects on the 20S proteolytic activity *in vitro* (Fig. 4.12). Slightly lower activity was observed for all three proteases for the P97-20S samples compared to 20S proteasome alone, in which the same concentrations of 20S proteasome and P97 were used. Addition of ATP did not change the results, except for the β5 proteases which showed slightly increased activity for the P97-20S samples compared to the 20S complex. However, there was also relatively high variability in P97-20S samples observed, and consequently this may not be a significant result. This result does not exclude the possibility that P97 may affect the *in vitro* activity of 20S proteasome. In these assays, a low proteolytic activity could also be detected in the P97 sample. This background proteolytic activity could be due to the presence of other active proteases in the sample or minor proteasome contaminations. In general, it is difficult to draw a firm conclusion on whether P97 affects 20S proteasome activity and whether other co-factors, such as ATP, or proteins are involved.

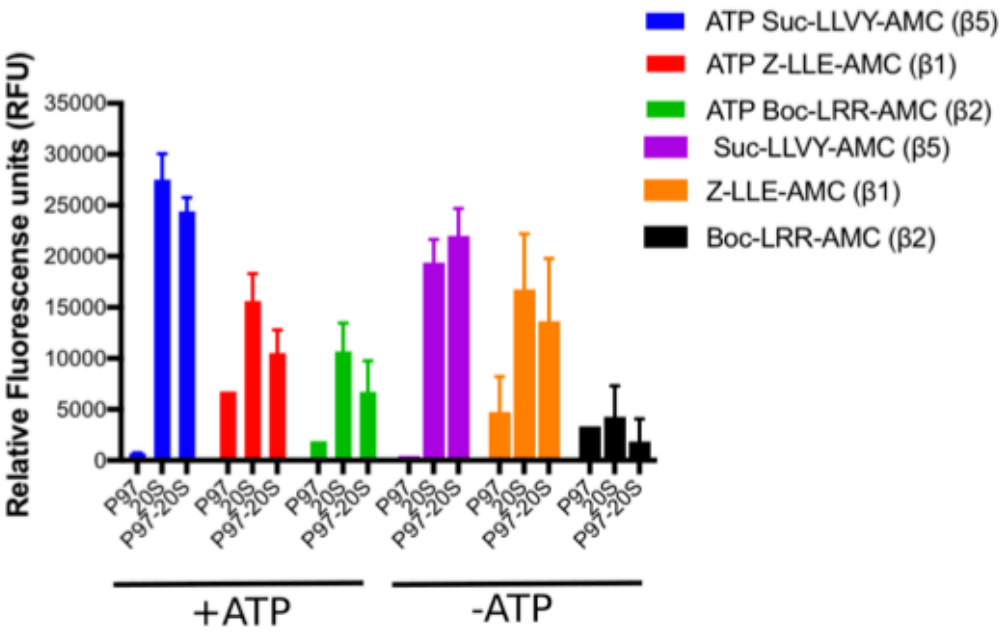


Fig. 4.12: Activity assays of 20S proteasomes in the presence of P97 with and without ATP. All proteolytic active proteasome ß-subunits were tested. The activity assays were conducted in buffers either containing ATP or no ATP present and substrates processed by different proteasome subunits (Suc-LLVY-AMC, Z-LLE-AMC, Boc-LRR-AMC for ß5, ß1 and ß2 respectively). Blue, red and green show ß5, ß1 and ß2 activity respectively, with ATP added to the buffer and purple orange and black show ß5, ß1, and ß2 activity respectively without ATP addition. Error bars represent the standard error of the mean (SEM).

4.5.7 Electron microscopy of P97-20S samples

In order to test binding of P97 to the 20S proteasome, negative stain grids were prepared of purified 20S proteasomes co-expressed with P97 and 20S chaperones (Fig. 4.13). The P97 protein levels are much lower compared to the 20S complex. This may be in part attributed to the purification strategy in which tagged 20S proteasome complexes have been used as bait. Despite the presence of faint, visible bands in the SDS-PAGE gel that correspond to P97, neither the expected P97 hexamer nor particles that could correspond to the P97 hexamer bound to the 20S complex could be identified by negative staining microscopy (see Fig. 4.13B). This indicates that other factors may be required for *in vitro* complex formation and that the interaction between P97 and the 20S proteasome without

other binding partners may only be transient. Alternatively, the complex formation may be disrupted during the stringent conditions of EM grid preparations.

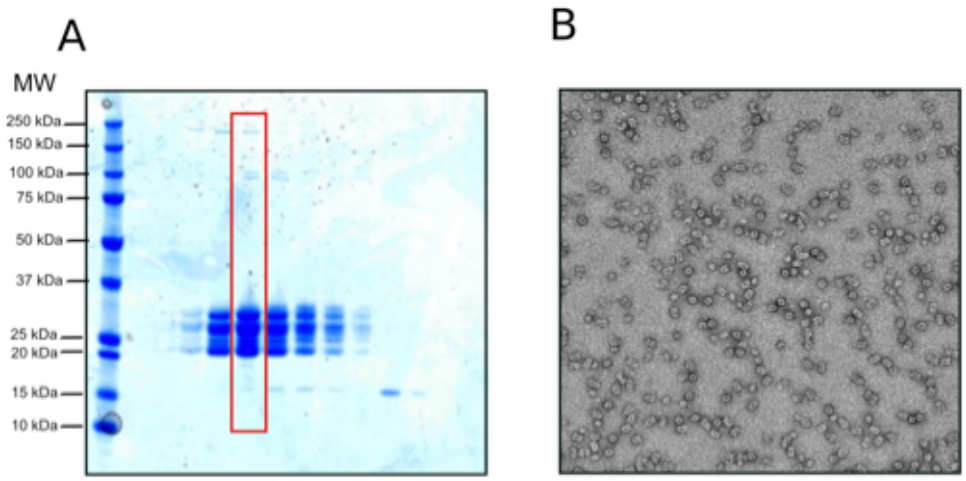


Fig. 4.13: Electron microscopy of P97-20S proteasome samples. A) a representative SDS PAGE of tagged-20S coexpressed with P97 and 20S chaperones purification. B) Sample form A) (red box) was visualised using negative staining technique.

4.6 Discussion

Understanding the complete proteasome interaction map remains a challenging aspect in the proteasome field. The proteasome interacts with a variety of different proteins, as discussed in Chapter one. P97 plays an important role in ERAD and the UPR whereas the proteasome is involved in the later stages of this pathway. However, if the highly abundant P97 protein binds directly to the human proteasome remained unknown so far. From our lab's initial observation of MS data and a small subset of scientific papers, there was initial evidence that P97 might directly bind to the proteasome (Barthelme and Sauer, 2012; Isakov and Stanhill, 2011; Meyer et al., 2012; Pick and Berman, 2013). My work sought to investigate this question and I showed that indeed P97 and the proteasome do interact.

My early experiments attempted to pull down the P97 protein in insect cells co-expressing P97 and 20S proteasome and, although there were traces of proteasome subunits, the yield

obtained of formed P97-20S complexes was very low. Using 20S as a bait instead, unfortunately did not improve the yields. However, P97 was consistently found co-eluting with 20S proteasomes indicating that it binds either directly or indirectly to the 20S proteasome. Due to the unknown nature of this interaction, several different experimental conditions were explored, including the addition of ATP or the use of immunoproteasomes versus constitutive proteasomes in affinity chromatography experiments. As mentioned above, no significant differences were observed. There are a variety of factors that may be contributing to the lack of identifiable P97-20S proteasome complexes in insect cells. For example, P97 may be depending on a cofactor or the P97-20S interaction may be transient or a low-affinity interaction and consequently difficult to purify at high quantities. Additionally, post-translational modifications (PTMs) differ between insect and human cells and are important for protein solubility, stability, and hydrophobicity (Klenk, 2002; Tokmakov et al., 2012). PTMs may therefore play an important role in the P97-20S interaction. Furthermore, species specific localisation of P97 and proteasomes within insect cells could also cause inefficient binding. While P97 localises to both the cytosol and nucleus in humans, it has been shown that proteasome localization in lower eukaryotes is mainly found in the nucleus compared to the cytosol (Takeda et al., 2011; Woodman, 2003). In addition, I showed that P97 eluted as a hexamer, suggesting that the lack of efficient binding is not due to P97 functionality.

My decision to continue optimising P97 and 20S proteasome expression in insect cells was due to the need to explore any other cofactors that may be co-purifying with P97 and the 20S proteasome. By using a codon optimised P97 version for expression in insect cells, I could significantly improve the obtainable yield of purified P97 protein. However, it was still difficult to produce high yields of P97-20S complexes. To overcome low yield in P97-20S complex, P97 and the 20S chaperones were incorporated into a single virus, reducing the number of viruses needed for transfection from three to two and thereby potentially promoting more stoichiometric complex formation. Unfortunately, this did not improve the yield of desired P97-20S complexes, suggesting that either other co-factors may be required for complex formation or the interaction is too transient to capture.

It is very difficult to predict other co-factors that could be interacting with the P97-20S complex especially because there are only few well-known co-factors and none of these have been observed in my MS analyses. The fact that I do not see these co-factors in MS

results may be due to the fact that these genes are not annotated in insect cells. As mentioned above, it is well known that the UN complex (involving co-factors Ufd1 and Npl4) is required in the unfolding process and is regularly observed in complex with P97. Since this process is upstream of the proteasome degradation step, they may continue playing a significant role and this could potentially be a good starting point for further investigation. Although I obtained these plasmids, due to time restraints I was unable to further pursue this avenue. Activity assays conducted with P97 and 20S showed high background of activity of P97 and variability in the activity assay made it difficult to draw comparisons. It would be useful to repeat these experiments with 26S proteasomes as well and to see if there is any difference observed when adding 26S.

Using crosslinking agents, such as BS3, during purification could also help to improve P97-20S complex stability, if the interaction is transient and difficult to capture. This exact approach proved essential when purifying P97-20S in archaea (Barthelme and Sauer, 2012). It is also worth trying to purify P97-20S complexes in the presence of Rad23 as this may be the key bridge between the two complexes as, previously described in Chapter one; it is a key shuttling factor that sends ubiquitinated proteins to the proteasome. Finally, recent tomography data have identified non-membrane-bound micro-compartments consisting of P97 and 26S complexes (Albert et al., 2020). The significance of these compartments is thought to be associated with compartmentalising ER-associated degradation and co-localising machinery at specific ER hotspots (Albert et al., 2020). It would, therefore, be interesting to conduct further experimental work using cryo-electron tomography to investigate the true nature of this interaction and identify which substrates P97 and the 26S machinery are binding and degrading to. Tomography allows for identification of heterogeneous structures and would allow P97 and the 26S proteasome to be visualised in their native, cellular environment with any other interacting proteins (Koning et al., 2018). A series of tiles of 2D projection images that are obtained and used to generate a 3D tomogram using a similar processing framework described in Chapter 5 (Beck and Baumeister, 2016). However, tomography requires thin volumes for suitable imaging conditions which is achieved by FIB-milling (Koning et al., 2018). These conditions would therefore need to be met if this strategy is pursued with the proteasome and P97.

It is clear that there is an interplay between P97 and 20S as shown as shown by the co-elution of P97 and 20S proteasomes in cells. The P97-20S interaction has not been seen before with human proteasomes, however, my experiments clearly show there is an interaction worthy of further investigation. My *in vivo* purification methods provide a foundation for investigating this difficult, and potentially transient interaction, with further P97 co-factors so we can fully understand the role P97 plays with the proteasome.

Chapter 5 PI31 and the 20S Proteasome

5.1 Introduction

Proteasome inhibitor PI31 subunit (PI31) is a 31 kDa protein that plays a role in controlling proteasome function and was first identified in bovine red blood cells fractions (Chu-Ping et al., 1992). PI31's exact role in the regulation of proteasome function, however, remains elusive (Shang et al., 2015). Compared to other known proteasome regulators, such as PA28 and PA200, PI31 has been observed to inhibit rather than activate 20S proteasome activity (Zaiss et al., 2002). However, this inhibitory activity has not yet been observed *in vivo* (Zaiss et al., 2002). PI31 is also the least studied of all the currently known proteasome regulators (Fort et al., 2015). PI31 is highly conserved through metazoan evolution and orthologues have been identified from yeast to mammals (Botelho-Machado et al., 2010). In Drosophila, PI31 function is essential, suggesting that PI31 may be involved in basic cell functions (Bader et al., 2011). Drosophila PI31 has also been shown to interact with the F-box protein Nutcracker, a component of the SCF complex type E3 ubiquitin ligases involved in caspase activation during spermatogenesis (Bader et al., 2011) (Fig. 5.1). Finally, PI31 has been linked to axonal proteasome transport and neuron maintenance (Liu et al., 2019; Minis et al., 2019). In general, several studies have shown PI31 involvement in a range of cellular roles which will be discussed below.

5.1.1 PI31 structure

PI31 contains a 150 amino acid N-terminal Fbxo7/PI31 domain (FP domain) that was first identified based on the sequence similarity between PI31 and the F-box protein (FBP) Fbxo7, also known as PARK15/Ncp (Fig. 5.1) (Shang et al., 2015) (Fig. 5.1). The FBP family of proteins is a diverse protein family with numerous members that do not necessarily have any functionality in common, but all possess this characteristic 45 amino acid F-box motif (Fig. 5.1) (Bai et al., 1996). A well-known example of proteins in the FBP family are the SCF E3 ubiquitin ligases, in which the FBPs act as the substrate-recognition subunit of the SCF complex. Such substrates are ubiquitinated by SCF for

degradation by the 26S proteasome (Goldenberg et al., 2004). The PI31-Fbxo7 interaction is conserved from Drosophila to mammals (Bader et al., 2011; Kirk et al., 2008). The structure of the FP domain has been solved and exhibits an α/β fold that associates as a dimer in both solution and in crystals (Fig. 5.1) (Kirk et al., 2008). The homodimerisation of PI31 occurs through a preferentially helical interface, whereas heterodimerisation with the Fbxo7 protein occurs through contacts between β sheets (Kirk et al., 2008). PI31 also contains a C-terminal proline-rich domain, a predicted disordered region in which 26% of amino acids are prolines (Fig. 5.1) (McCutchen-Maloney et al., 2000). The C-terminus of PI31 contains a HbYX-motif that can bind to the 20S proteasome. This motif is also found in the Rpt subunits of the 19S and other regulatory proteasome-binding proteins (such as P97) as discussed in Chapter one (Bader et al., 2011; McCutchen-Maloney et al., 2000). Inhibition of the 20S by PI31 could be mediated by restricting access to the enzymatic core of the proteasome through a substrate-occlusion mechanism by binding of the C-terminal tail, however, this has not been shown directly (Bader et al., 2011; McCutchen-Maloney et al., 2000). The N-terminus of Fbox7 is structurally similar to PI31. However domains in Fbxo7 that are are not found in PI31 include the UBL and F box domains (Kirk et al., 2008) (Fig 5.1).

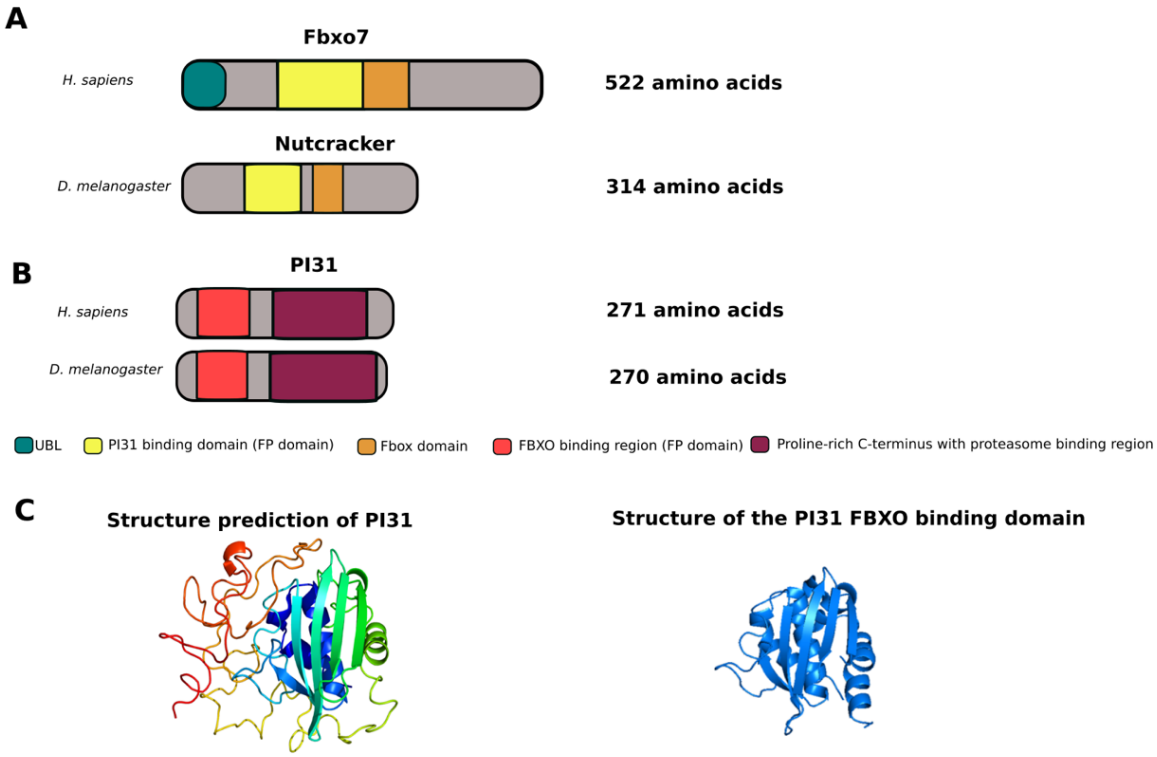


Fig. 5.1: Schematic of the structural features of PI31 and FBXO proteins. Domain scheme organisation of the human and fruit fly of A) F-box proteins Fbxo7 and Nutcracker and B) PI31. Both human and *D. melanogaster* proteins contain homologous FP domains at their N-terminus. C) Tertiary structure prediction of the full length human PI3I as determined by Phyre (Kelley LA et al. (2015), shown alongside the structure of its FP domain (PDB: 4OUH) (Kirk et al., 2008).

5.1.2 PI31 and the proteasome

PI31 has been shown to inhibit proteasome activity (Chu-Ping et al., 1992). However, this observation is limited to *in vitro* studies and has not been observed *in vivo* in mammalian cells (Zaiss et al., 2002). On the other hand, studies have shown that PI31 promotes protein degradation *in vivo* in yeast, plants, and Drosophila (Bader et al., 2011; Yang et al., 2016; Yashiroda et al., 2015). Additionally, Drosophila PI31 expression upregulates the *in vivo* 26S activity under conditions that require increased regulated proteolytic activity, as seen in the terminal differentiation of sperm (Bader et al., 2011). PI31 is also thought to have a role in regulating the immunoproteasome (Zaiss et al., 2002). As discussed in Chapter one, immunoproteasomes play an important role in regulating proteolytic processes involved in MHC class 1 antigen presentation (Rock and Goldberg, 1999;

Shastri et al., 1998). During antigen presentation, the γ-inducible immunoproteasome cleaves MHC class I ligands from antigenic sequences in a highly specific manner (Zaiss et al., 2002). PI31 seems to also be involved in this process by interfering with the maturation of the immunoproteasome (Zaiss et al., 2002). Overexpressing PI31 reduces MHC class I surface levels as well as the MHC class I presentation of cytotoxic T-lymphocyte epitopes in IFN-γ treated mouse embryonic cultured cells (Zaiss et al., 2002). The exact mechanistic role of PI31 in abrogating immunoproteasome function is unknown. However, it has been suggested that PI31 may be interfering with processing of the proteasome subunit pre-peptides and preventing mature immunoproteasome formation (Zaiss et al., 2002). In terms of localisation, both PI31 and immunoproteasomes and immunoproteasome precursors localise at the nuclear envelope and ER membrane providing further evidence of a potential link (Zaiss et al., 2002). Moreover, PI31 has been shown to compete with PA28-mediated activation of immunoproteasomes (Zaiss et al., 1999). Through kinetic analyses, it was found that PI31 exhibits higher affinity for the 20S proteasome than PA28 and, consequently, a 50-fold higher PA28 concentration is needed to abrogate PI31-mediated 20S proteasome inhibition (Zaiss et al., 1999).

PI31 has also been linked to 26S proteasome assembly (Cho-Park and Steller, 2013). Drospholia PI31 has been shown to bind to and be regulated by the ADP-ribosyl transferase tankyrase (TNKS), which is linked to a variety of cellular processes including telomere maintenance (Cho-Park and Steller, 2013; Hsiao and Smith, 2008). PI31 is ADP-ribosylated by TNKS to stimulate 26S proteasome function (Cho-Park and Steller, 2013). Ribosylated, modified PI31 possesses lower affinity for the 20S proteasome and thus allows 20S proteasomes to bind to 19S complexes and to form 26S proteasomes (Cho-Park and Stellar, 2013). In fact, in both mammalian and Drosophila, inhibiting TNKS reduces 26S proteasome activity (Cho-Park and Steller, 2013), although, evidence suggests that mammalian PI31 is not ADP-ribosylated (X. Li et al., 2014).

Besides 26S proteasome assembly, PI31 has also been linked to 26S proteasome transport in neuronal cells (Liu et al., 2019). Proteosome transport in neurons has been a long outstanding question in the field, due to the critical function of proteasomes at synapses and the long distances proteasomes have to travel within neurons. To solve this transport problem, proteasomes employ a microtubule-based transport mechanism. However, much

of this pathway remains unknown (Gorbea et al., 2010; Hsu et al., 2015; Kreko-Pierce and Eaton, 2017; Otero et al., 2014). Recent studies have shown that modified PI31 is involved in this process by binding to the dynein complex, a molecular motor component of the microtubule network (Liu et al., 2019). Phosphorylation of PI31 at a conserved site by P38 microtubule-associated protein (MAP) kinase was shown to promote binding of PI31to the dynein light chain LC8-type (DYNLL1/2) proteins, which are involved in proteasome transport by the microtubule network (Kreko-Pierce and Eaton, 2017). In the study conducted, PI31 mediates axonal transport of proteasomes by binding to DYNLL1/2 and acting as an adapter in Drosophila motor neurons (Liu et al., 2019).

5.1.3 Other cellular PI31 roles

As mentioned, PI31 binds to Fbxo7, but surprisingly this interaction leads to increased PI31 stability rather than ubiquitin-mediate degradation (Bader et al., 2011; Vingill et al., 2016). Loss of Fbxo7 causes impaired proteasome function, which is consistent with the observation that PI31 reduces proteasome function (Liu et al., 2019). There is, however, a lack of supporting evidence to explain the exact cellular consequences of this interaction.

PI31 also interacts with the CD2BP2 protein, originally discovered as a binding partner of the T cell adhesion protein CD2 which is part of the T cell signalling pathway (Kofler et al., 2005). The interaction occurs through the glycine-tyrosine-phenylalanine (GYF) domain of CD2BP2, which recognises the proline-rich sequence within the C-terminus of PI31 (Kofler et al., 2005). This interaction was shown by yeast two-hybrid and spot analysis, but its significance remains unknown and has yet to be demonstrated *in vivo*.

5.1.4 PI31 in disease

There are disease-related implications of the PI31-Fbxo7 interaction. Mutations in Fbxo7 have been shown to impair proteasome function and cause Parkinson Disease (PD)-like symptoms in mice and early onset form of PD in humans (Conedera et al., 2016; Di Fonzo et al., 2008; Paisán-Ruiz et al., 2010; Vingill et al., 2016). These symptoms are referred to as Parkinson's plus syndromes (PPS). PPS are groups of illnesses that attack the brain and nerve cells and are linked to PD, but may include other underlying pathologies. In fact,

unlike more standard models of PD, PPS affect neurons in the pyramidal area of the brain, causing earlier symptoms, such as loss of muscle control (Mitra et al., 2003). The reasons for such links are unknown, but it has been proposed that mutations in Fbxo7 could impair normal protein clearance, resulting in protein aggregate formation that are typically found in many neurodegenerative diseases (Ballatore et al., 2007; Irvine et al., 2008; Li and Li, 2011; Ross and Poirier, 2004; Tai and Schuman, 2008). Using genome-wide association studies, PI31 has also been linked to Alzheimer's Disease (AD) (Sherva et al., 2011).

The Fbxo7 homologue, Nutcracker, has been studied in Drosophila, which has proven a useful model organism to study human diseases *in vivo.* This has been especially important in understanding neurological systems due to the similar level of complexity of neurons and neurotransmitter systems in Drosophila compared to more complex organisms such as mammals (Bilen and Bonini, 2005). Both Nutcracker and a *Drosophila melanogaster* homologue of PI31 were studied by Merzetti et al., (2017) who showed that Nutcracker disruption leads to reduced proteasome activity, but has no reducing effect on the proteosomal machinery. Drosophila cells that have a mutated PI31 homologue show increased levels of poly-ubiquitinated proteins indicating lack of a functional proteasome (Bader et al., 2011).

5.1.5 Proteasome inhibitors

Taken together, it is clear that PI31 seems to have an important role in basic cellular homeostasis and in functions relating to proteasome regulation. The fact that PI31 has been shown to have inhibitory effects on proteasome function raises its use as a potential therapeutic agent for the treatment of cancers, with some proteasome inhibitors already in clinical use as discussed in Chapter one (Crawford et al., 2006; Cusack, 2003; Park et al., 2018). Currently, three proteasome inhibitors, bortezomib, carfilzomib, and ixazomib, are routinely used in clinical settings and have received approval from regulatory bodies (Manasanch and Orlowski, 2017). These treatments, mostly confined to myeloma and mantel-cell lymphoma, have dramatically improved patient outcomes (Manasanch and Orlowski, 2017). However, relapses and resistance are common to these drugs, indicating the need for further therapies (Manasanch and Orlowski, 2017). PI31 is one of several mammalian proteasome inhibitor proteins, which also include the HIV-1 transcriptional

activator, Tat, and a heptatitis B virus protein, HBx (Rechsteiner and Hill, 2005). Other known proteasome inhibitors include the proline and arginine-rich polypeptide (PR39) and proteasomal-associated factor 1 (PAAF1). Structural studies of PR39 have shown its structural rearrangement of both the 20S and 26S proteasome. This arrangement has been shown to allosterically inhibit the proteasome peptidase activity (Gaczynska et al., 2003). PR39 was initially discovered as an antimicrobial peptide secreted by macrophages (Agerberth et al., 1991). Finally, 26S purification from HeLa cells identified PAAF1, with subsequent experiments showing its proteasome inhibitory function (Park et al., 2005). PAAF1 is thought to destabilise the 20S CP-1 19S RP interaction and overexpression of PAAF1 results in ubiquitinated protein accumulation (Park et al., 2005). The modulation of its interaction with the proteasome can therefore, in principle, be explored in the development of new therapies.

5.2 Results

The current literature on PI31 and 20S proteasome provided many potential avenues to explore the PI31-20S interaction and understand its biological significance. Despite the large amounts of data available on PI31 and the 20S proteasome, there is not yet any structural information showing the direct interaction between these proteins and I wanted to address this. As mentioned in previous chapters, our lab established the recombinant assembly of the human 20S proteasome in insect cells. To develop a wider understanding of the PI31 physiological function, I sought to investigate the PI31 and proteasome interactions through protein biochemistry and a structural biology approach. I did so by using the recombinant 20S expression system in which insect cells were co-infected with baculoviruses containing 20S proteins and PI31. This approach also allows other proteins that may be interacting with both PI31 or the proteasome from the insect cells to co-elute, providing a wider understanding of the PI31 physiological function.

5.2.1 Co-expression of PI31 and the 20S proteasome

Two streptavidin tags were initially added at the N-terminus of PI31, based on reports from the literature which suggested that PI31 binds to the 20S proteasome through the C-terminus. Baculoviruses containing this Twin-Strep tagged PI31 and the 20S chaperones

were produced with bacmids prepared by cre-lox recombination as discussed in Chapter four to improve the efficiency of co-infection. Upon expression, PI31 was pulled-down using streptavidin affinity chromatography (Fig. 5.2A and B). From the first examination, there were several other proteins pulled down with PI31, and the typical 20S proteasome smear was seen co-eluting with PI31(Fig. 5.2B). As a second step, fractions from the affinity chromatography were pooled (fractions 6-12, Fig. 5.2B) and size-exclusion chromatography was conducted, which showed PI31 co-eluting with 20S proteasomes (Fig. 5.2C and D). The majority of PI31 eluted in later fractions, as expected for a 31 kDa protein (fractions 20-22, Fig. 5.2D). However, a population of PI31 also eluted in the earlier fractions (see red box in Fig. 5.2D). The PI31-20S proteasome complex as obtained from the size exclusion chromatography experiment (fraction 16, Fig. 5.2D) was then observed under negative staining EM. From first observations, it was clear that 20S was co-eluting with PI31 in the sample but further structural work was needed to identify any stable complexes forming (Fig. 5.2E).

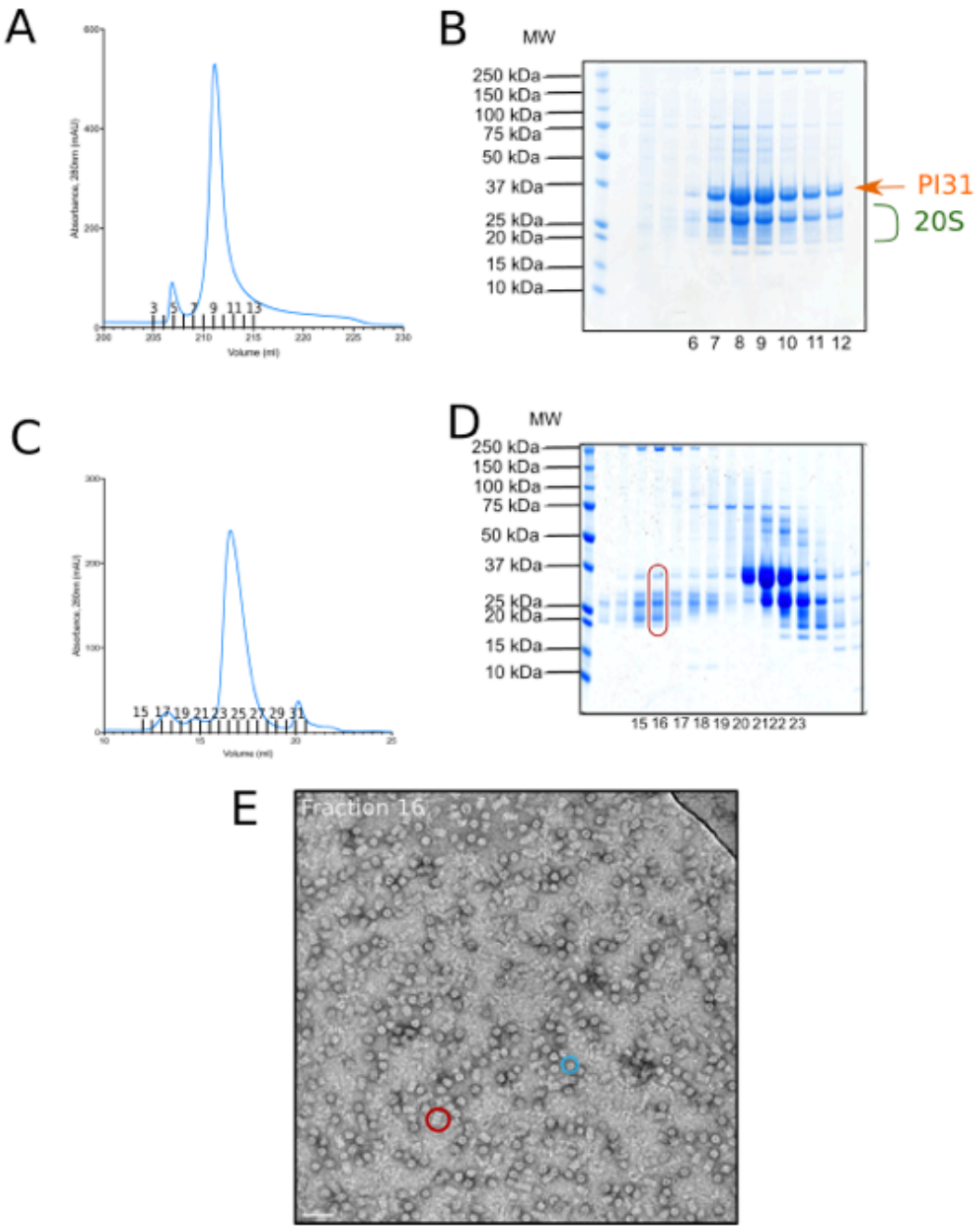


Fig. 5.2: Purification of streptavidin tagged PI31 co-expressed with the 20S proteasome. Streptavidin-tagged PI31 was co-expressed with untagged 20S proteasome in insect cells and purified in two steps. The first step involved affinity chromatography in which A) a representative chromatogram and B) SDS-PAGE gel are shown. C and D) Fractions 6-12 from B were pooled, concentrated and run on size-exclusion chromatography in which a representative chromatogram and SDS-PAGE gel are shown respectively. E) A representative micrograph of fraction 16 from panel D was visualised by negative stain EM. 20S proteasomes side views (red) and top views (blue) are seen.

5.2.2 Biochemical analysis of PI31

Before further experiments were designed and conducted involving PI31 and 20S proteasome, the oligomeric state of PI31 was investigated. Previous literature had shown that PI31 forms a dimer *in vitro* and that this its preferential oligomeric state (McCutchen-Maloney et al., 2000). However, I decided to test whether this held true for my PI31 construct and whether other oligomeric states could be identified. Twin-Strep tagged PI31was therefore expressed using the baculovirus expression system and purified by streptavidin affinity chromatography followed by SEC (Fig. 5.3A and B). This resulted in two main peaks corresponding to the different protein bands seen by SDS-PAGE (Fig. 5.3B). Fraction 18 from the SEC was then further analysed by a SEC-MALS (Fig. 5.3C). The SEC-MALS data showed a main peak with a molar mass of 38.3 kDa, which was close to that expected molecular weight for a monomer of PI31 (31 kDa) (Fig. 5.3C). There was also a minor peak present at a molecular mass of 58.3 kDa, suggesting a monomer-dimer equilibrium. Collectively, this shows that PI31 in my hands was observed as both a dimer and a monomer as seen by others (McCutchen-Maloney et al., 2000).

The two peaks of PI31 observed in size-exclusion chromatography were of interest as it was unclear whether both proteins were PI31 or whether it was a contaminant (Fig. 5.3B). The two bands were thus analysed using MS (Fig. 5.4). The MS analysis showed that both of these bands were PI31. However, the lower band corresponded to PI31 with an N-terminal truncation (Fig. 5.4B). While the significance of this truncation was unknown, increasing the concentration of protease inhibitors present in the lysis buffer of the PI31 sample decreased the lower band, (as seen in later purifications). In addition, negative stain grids were prepared of PI31 alone, however, these were of poor quality due to high background signal. This was likely due to either noise or PI31 monomers that were too small to be visualised through negative staining.

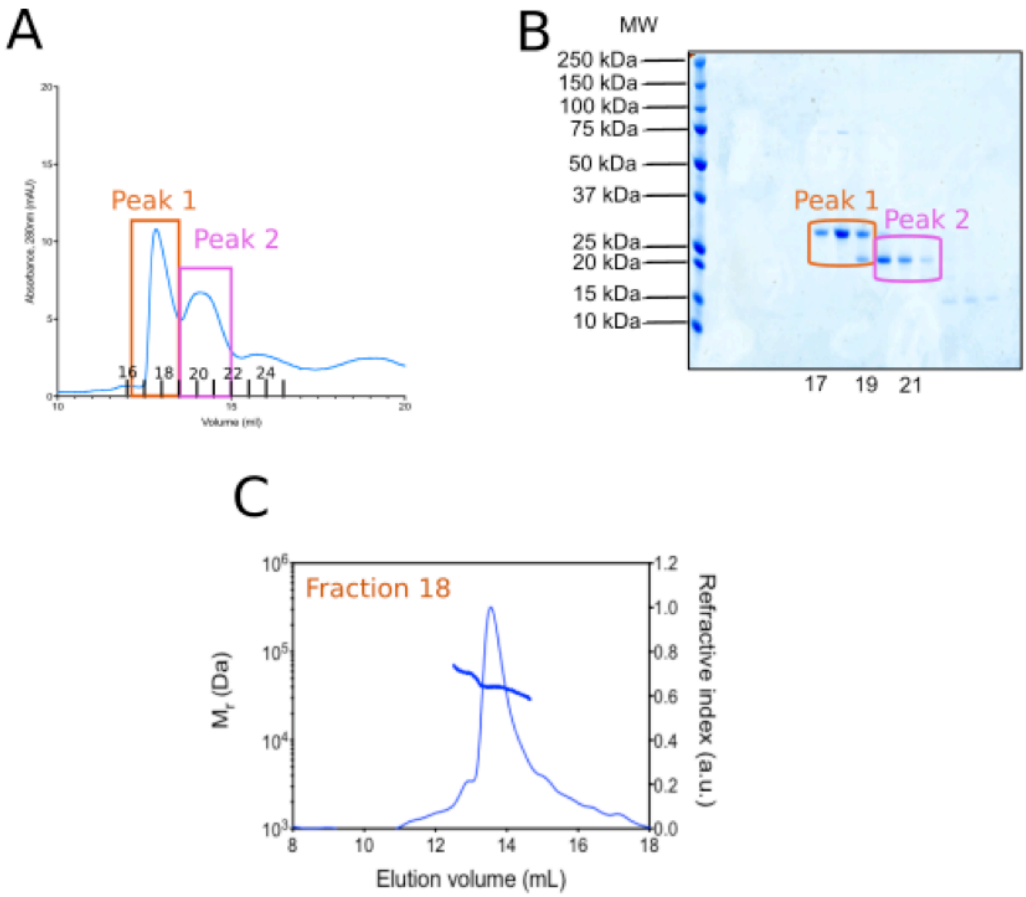


Fig. 5.3: Characterisation of PI31. Streptavidin- tagged PI31 was expressed in insect cells and purified using streptavidin affinity chromatography followed by size-exclusion chromatography (SEC) in which A) a representative chromatogram and B) SDS-PAGE gel are shown. C) Fraction 18 from peak 1 on the SEC (orange box) was then investigated using SEC-MALS in which both dimer and monomer forms were observed.

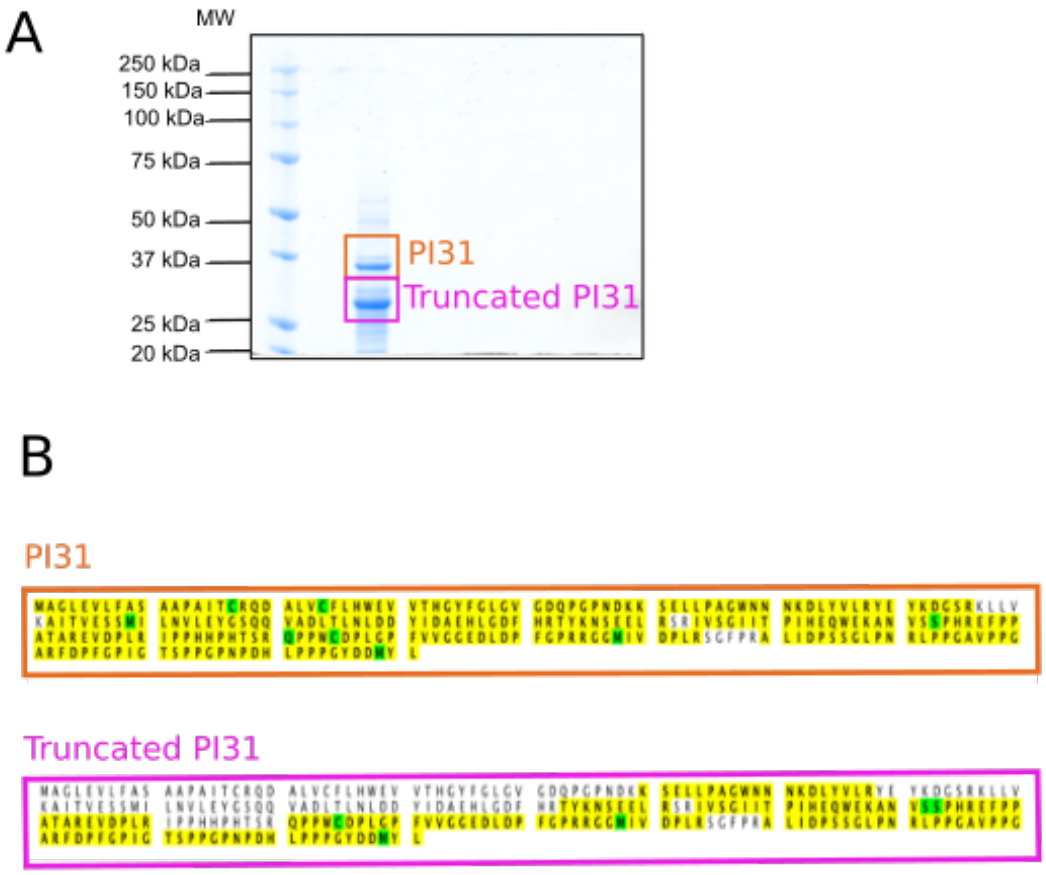


Fig. 5.4: PI31 purification analysis. A) The two bands observed in PI31 purification, after streptavidin affinity chromatography and size exclusion chromatography shown in a representative SDS-PAGE gel. B) Peptides derived from the two bands were analysed by mass spectrometry and the corresponding peptides identified in each sample are highlighted (yellow and green). The results showed both bands were PI31 however the lower band, with the smaller molecular weight, was a truncated form of PI31. The truncation is in the N-terminal region of PI31. The truncations could be a result of proteases in the sample as adding protease inhibitors decreased the concentration of the lower-truncated band (seen in later purifications).

5.2.3 Preparation of tagged 20S-PI31 complexes

Having determined the oligomeric states of PI31, I proceeded to further investigate the PI31-20S proteasome interaction. From my Twin-Strep PI31 purifications, endogenous 20S proteasomes from insect cells were observed, suggesting that PI31 that was solely expressed was also capturing insect 20S proteasomes (Fig. 5.5). This further suggested that during co-expression of 20S proteasomes and tagged-PI31 in insect cells, the PI31-20S complexes (as shown in Fig. 5.2D) were likely a mixture of recombinantly expressed 20S

proteasomes and endogenous 20S proteasomes from insect cells. Due to the unknown sequence of the 20S proteasome subunits from Sf9 cells, obtaining a mixture of 20S populations binding to PI31 could complicate the structural analysis and this needed to be avoided. Twin-Strep tagged-20S proteasomes and untagged PI31 were instead co-expressed in insect cells. This involved re-cloning and removing the tag on PI31 and making further baculoviruses. The same procedure then followed (as shown in Fig. 5.1), except with Twin-Strep tagged 20S and untagged PI31 (Fig. 5.6). Due to the absence of a tag on PI31, it was difficult to distinguish whether PI31 was co-eluting with 20S proteasome. However, by running 20S sample next to PI31-20S samples (Fig. 5.10A) it was clear that a top band was present in the PI31-20S samples, and missing from the 20S proteasome. In addition, a western blot with anti-PI31 antibody further confirmed this result with other samples (Fig. 5.10C).

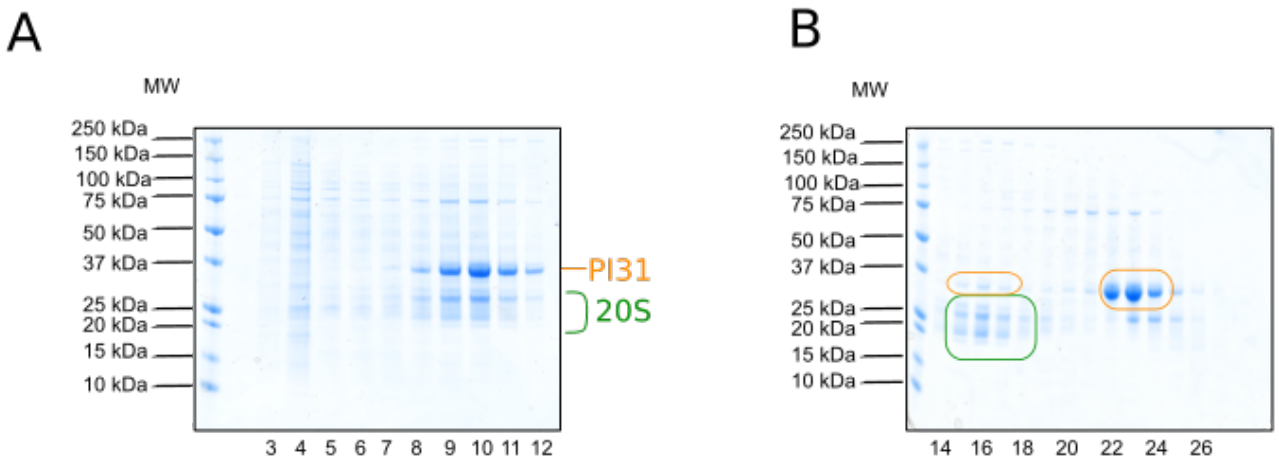


Fig. 5.5: Purification of streptavidin tagged PI31 co-eluting with insect 20S proteasome. Streptavidin-tagged PI31 was expressed in insect cells and purified in two steps. A) Streptavidin affinity chromatography pulled down on the tagged PI31 in which the representative SDS-PAGE is shown. B) Fractions 8-11 from A were pooled, concentrated and separated by size-exclusion chromatography in which the representative SDS-PAGE gel is shown. 20S and PI31 are marked in green and orange, respectively.

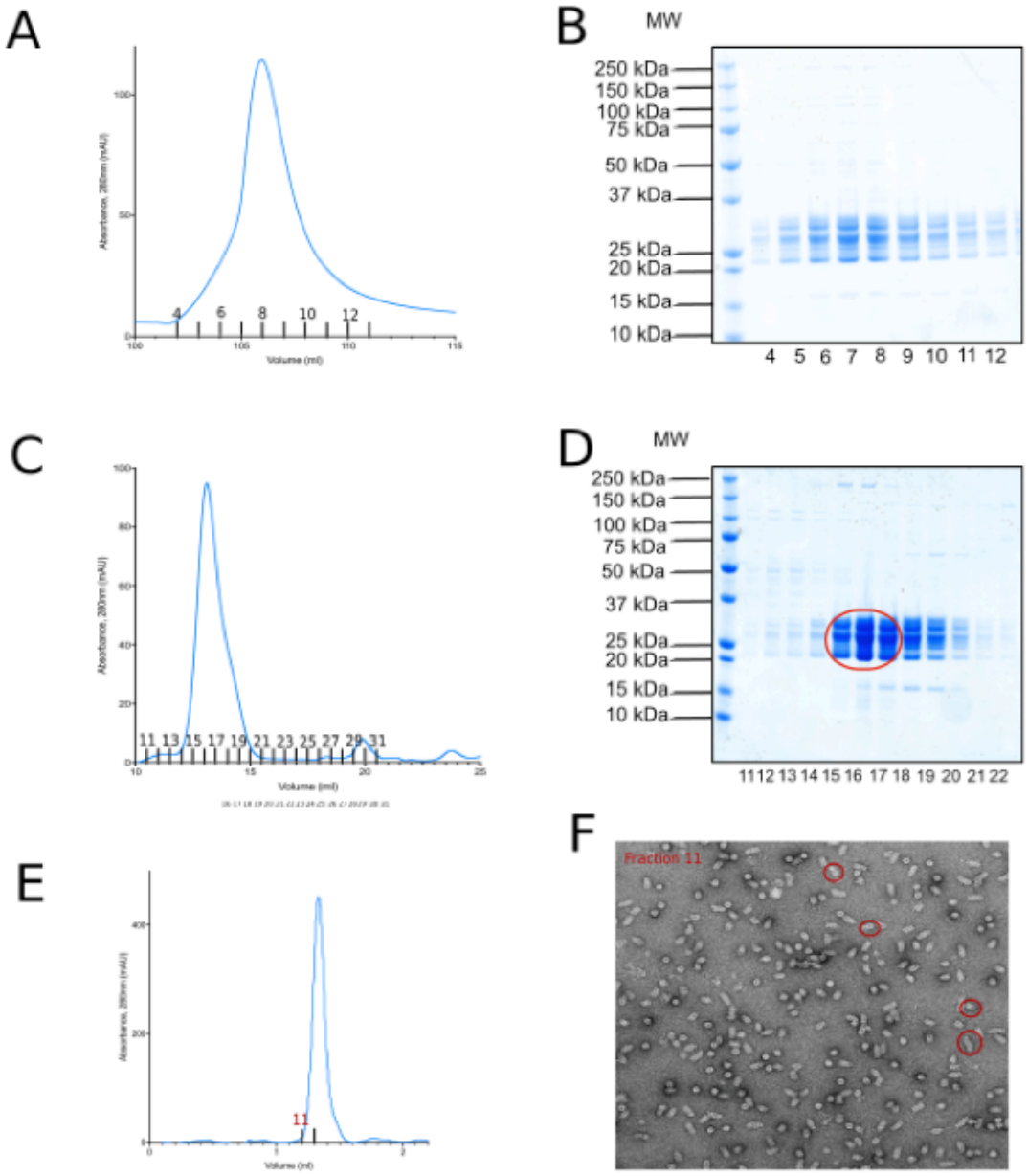


Fig. 5.6: Multi-step purification of PI31-20S complexes. Streptavidin- tagged 20S proteasomes were co-expressed with untagged PI31 in insect cells and purified. A) 20S proteasomes were pulled down using streptavidin affinity chromatography in which representative chromatogram and B) SDS-PAGE gel is shown. Fractions 4-12 were from B were pooled, concentrated and separated by several rounds of size-exclusion chromatography in which C) the representative chromatogram and D) SDS-PAGE gel is shown from the first round (red circle denotes pooled fractions), followed by E) the representative chromatogram from the final purification step. F) A representative negative stain micrograph of PI31-20S sample from the final purification step with examples of PI31-20Scomplexes circled in red.

The purification of tagged-20S and PI31 consisted of several steps. In my optimised protocol, a streptavidin-affinity chromatography purification step was followed by three rounds of size exclusion chromatography (Fig. 5.6C and E). At each step of size-exclusion

chromatography, the very beginning of the peak (fractions 15-17 from Fig 5.6) were selected to minimise the overlap with the free 20S proteasome. The fact that PI31 has a small molecular weight meant that free 20S proteasomes eluted very close to PI31-bound 20S and it was therefore important to minimise this overlap as much as possible. The corresponding negative stain grids from the PI31-20S samples showed the most 20S proteasomes with a cap-like structure (Fig. 5.6F). This required further structural analysis.

To better understand the PI31-20S complex, further experiments were conducted to explore its purification and stability. Firstly, during initial purification trials, I had difficulties concentrating the PI31-20S fractions before performing a size-exclusion chromatography. The fractions tended to form aggregates which were seen in SEC profiles (aggregating at fractions 8-9 on SEC profiles). However, after increasing the glycerol concentration in the lysis buffer from 5% to 10%, the protein concentration became efficient, no precipitation was observed and a greater yield of complex was obtained (Fig. 5.7A and B). The PI31-20S complex was observed as a small "hump" on a chromatogram and, upon increasing glycerol concentration in the lysis step, this part of the chromatogram was better defined (Fig. 5.7A). As well as accounting for the glycerol concentration, different concentrations of EDTA in the lysis buffer were also tested. Previously in the lab, experiments conducted by Dr. Migle Kisonaite had shown that including EDTA in all purification buffer solutions is important to minimise $MgCl_2$ availability, in order to obtain stable complexes of 26S proteasomes that are structurally intact as observed by cryo-EM. PI31-20S samples were, therefore, purified with both 1 mM and 10 mM EDTA. However, in my case, I saw no clear difference in yield of the P131-20S complex, as observed by SDS-PAGE (Fig. 5.7). This suggested that the complex formed stable complexes even in higher concentrations of EDTA present (Fig. 5.7C and D, green box). 1 mM EDTA was therefore used in future protein purifications.

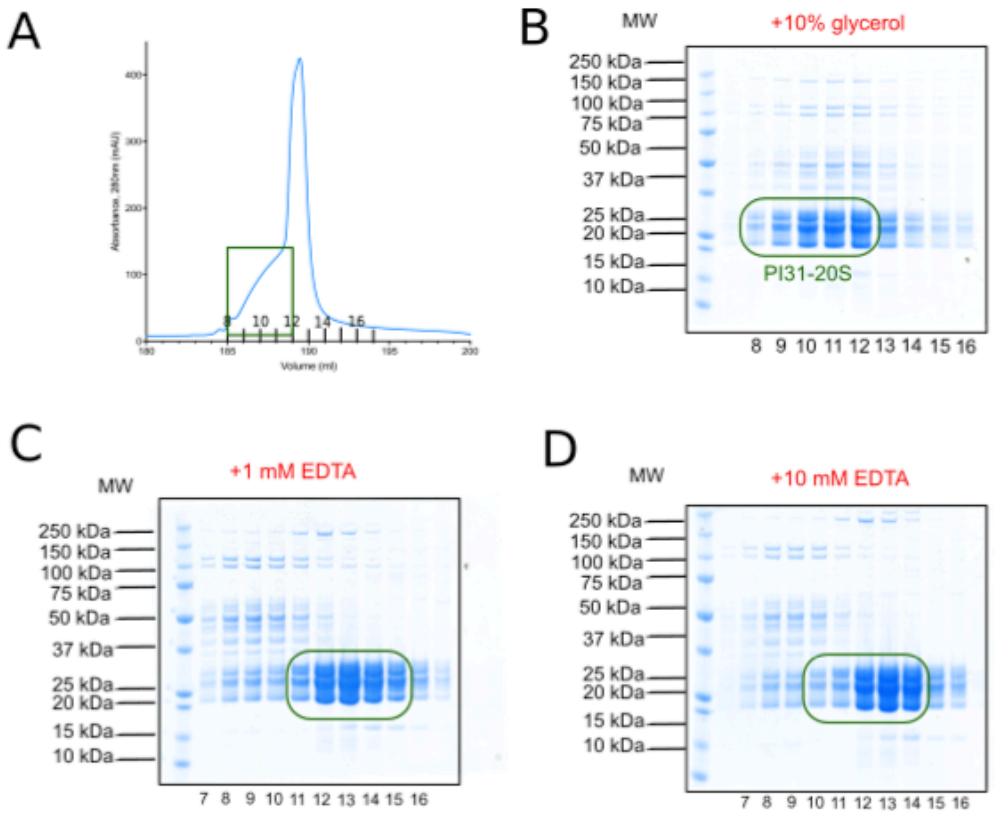


Fig. 5.7: Optimisation of the PI31-20S purification. Different aspects of the PI31 purification protocol were changed to understand the stability and improve the purity of the PI31-20S sample. A and B) Doubling the glycerol concentration from 5 to 10% within purification buffers improved the concentration of the sample and defined a better peak shoulder of the complex (green box panels A and B) in which A) a representative chromatogram and B) SDS-PAGE gel are shown. C and D) Different concentrations of EDTA (1 and 10 mM respectively) were also tested in the purification buffers in which representative SDS-PAGE gels are shown, however, this did not make a difference to the yield or purity of the protein obtained.

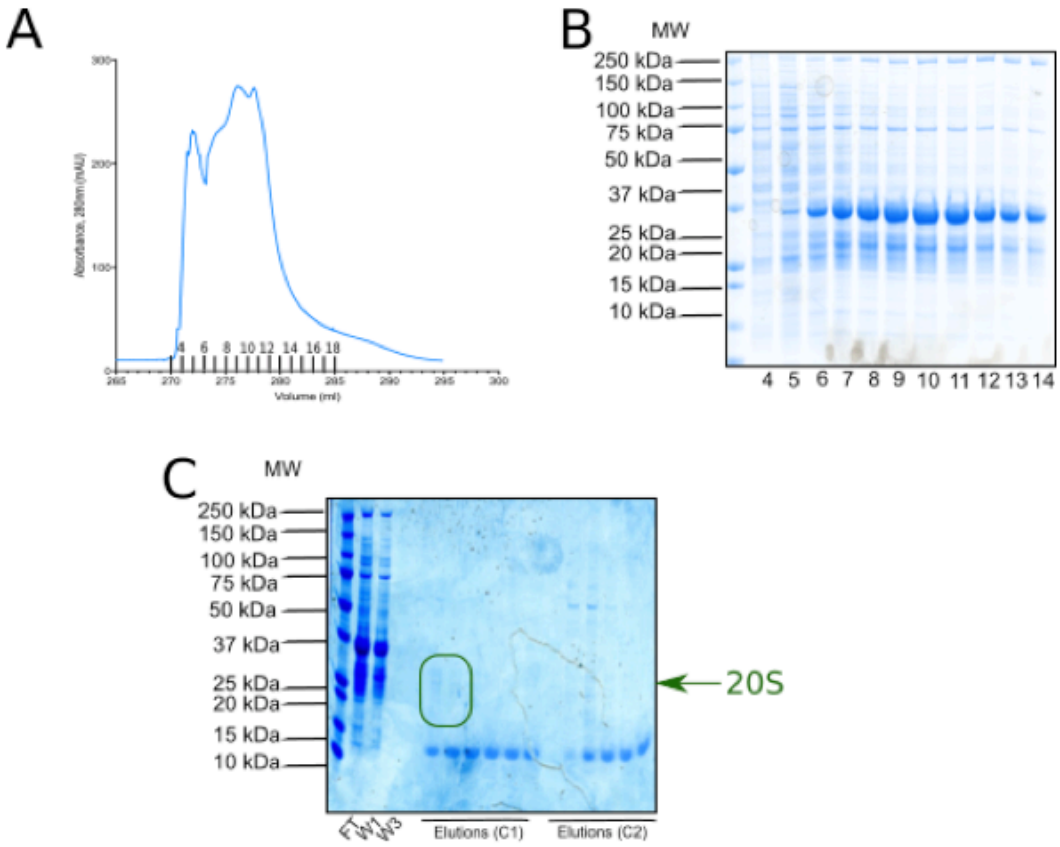


Fig. 5.8: Flag affinity purification of the PI31-20S complex. Streptavidin- tagged PI31 was co-expressed with flag-tagged 20S proteasomes in insect cells and purified in two steps. PI31 was pulled down by streptavidin affinity chromatography in which A) the representative chromatogram and B) SDS-PAGE gel is shown. C) Fractions 5-14 from B were pooled, concentrated and purified using flag affinity purification in which the representative SDS-PAGE is shown. Two columns were used due to the low efficiency of flag beads in capturing proteins (elution from column 1 and 2 referred to as C1 and C2 respectively). The former pulled down on PI31 and the latter pulled down on the 20S proteasome from the initial samples taken from step 1. Although some traces of 20S proteasome were seen in the elution factions, the yield was low and the concentration of complex was too low to work with.

To further optimise the purification procedure, different purification techniques were tested to improve the purification of the PI31-20S complex. Flag-tagged-20S proteasome and Twin-Strep tagged PI31 were co-expressed using the baculovirus expression system. The use of the two labels during purification can solve the potential overlap of free 20S proteasome seen before and enrich for PI31-20S complexes. After co-expressing flag-tagged 20S proteasome and Twin-Strep tagged PI31 in insect cells, a streptavidin affinity column was used to capture Twin-Strep tagged PI31 (Fig. 5.8A and B). The fractions containing PI31 and 20S (fractions 6-14 from Fig. 5.8B) were then combined and

incubated with flag beads to capture flag-20S proteasome within these samples (Fig. 5.8C). Most PI31 and 20S were seen in the wash fractions (green box, Fig. 5.8C). Some traces of 20S were seen in the eluted fractions of the flag affinity capture experiment, however, the yield obtained was very low and this purification strategy was therefore not pursued further (Fig. 5.8C, green box). The poor yield may have been affected by the tag or the affinity of the complex to the beads. For example, the new double flag tag used in β7 of the 20S proteasome may have hindered complex formation.

As an alternative, PI31 and 20S proteasome were purified independently in an attempt to reconstitute the interaction *in vitro*. This would confirm the direct interaction of PI31 to the 20S proteasome *in vitro* and obtain some information on how transient this PI31-20S interaction is. This involved incubating PI31 with 20S, followed by size exclusion chromography of the mixed sample, and elution of the PI31-20S complex. Both Twin-Strep tagged 20S proteasome (Fractions 11-16 Fig. 5.9B) and Twin-Strep tagged PI31 (Fractions 22-25, Fig. 5.9D) were individually purified from insect cells by streptavidin affinity purification; the 20S proteasome tag was cleaved and, after mixing the 20S proteasome with PI31, a further size exclusion purification was performed (Fig. 5.9). 20S proteasomes were incubated with uncleaved PI31 as it had been previously shown that tagged-PI31 does not interfere with binding to the 20S proteasome. The presence of the tag also allowed identification of PI31 in the 20S proteasome factions, as untagged PI31 overlapped with 20S proteasome subunits (Fig. 5.9A and B). The purified 20S proteasome was incubated with 10-fold molar excess of PI3I, and the sample was injected into a size exclusion column in order to separate populations of bound and unbound proteins. The resulting chromatogram showed that PI31 and the 20S proteasome did not co-elute as a complex (Fig. 5.9C).

The results indicated that under these *in vitro* conditions, the interaction between 20S and PI31 may be transient i.e. PI31 has a high K_{off} rate and for the majority of the time the proteins are found in their free state. However, from my previous experiments, it was clear that human PI31 forms a stable interaction with the 20S proteasome upon overexpression in insect cells, which remains at at high concentrations of salt (over 100 mM NaCl) and this complex was preserved over several rounds of purification. It is therefore more likely that an external co-factor, present in insect cells, either another protein, a chaperone, or

small co-factor, could instead be contributing to stabilising the PI31 and 20S proteasome interaction. I therefore decided to proceed working on co-expressed tagged 20S proteasome and untagged PI31complex from insect cells, as this purification strategy yielded the most promising results, which I could work with structurally and biochemically.

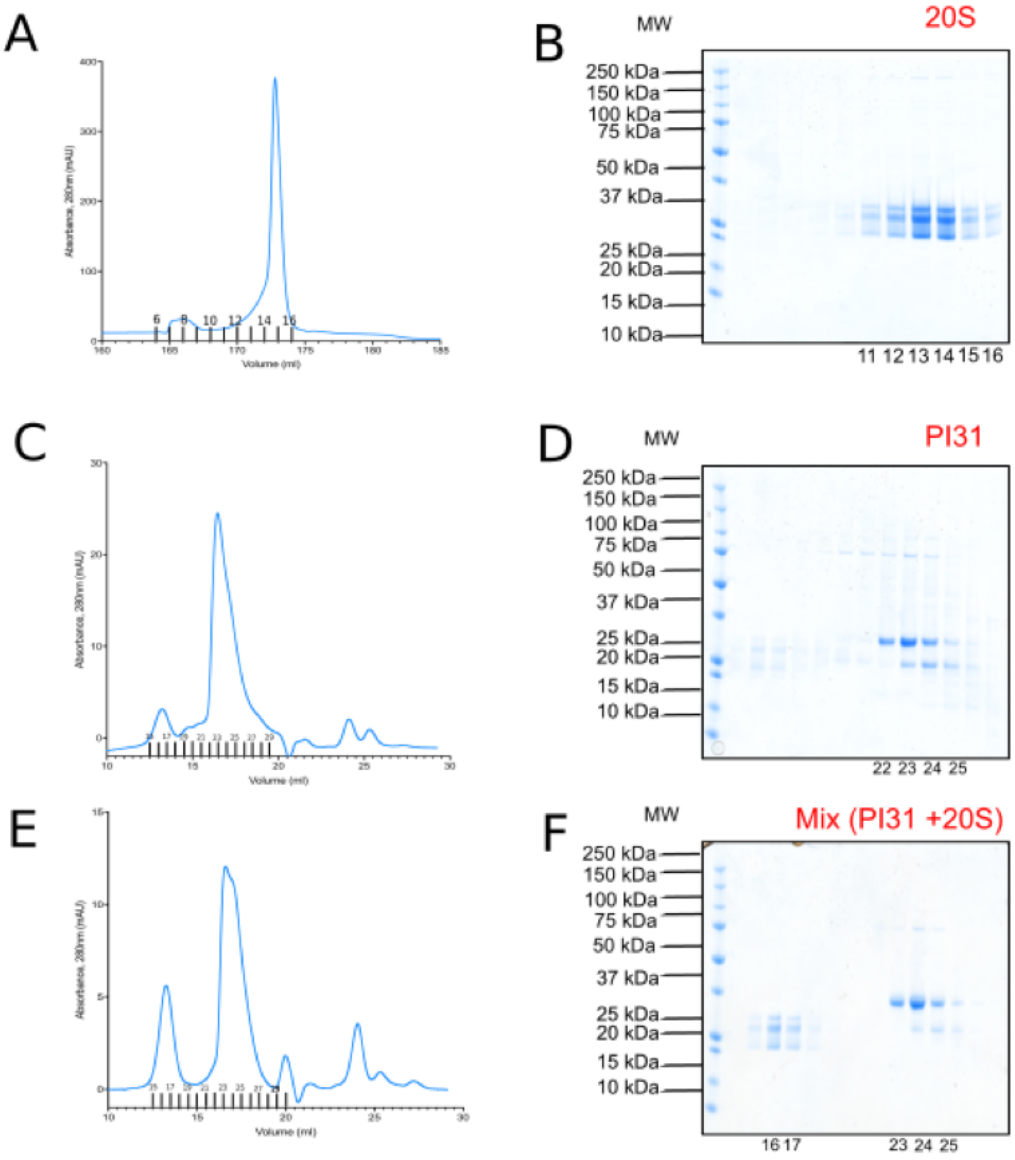


Fig. 5.9: *In vitro* binding of PI31 and the 20S proteasome. Streptavidin-tagged 20S and streptavidin-tagged PI31 were individually purified in which A) and B) show the representative chromatogram and SDS-PAGE gels for the 20S proteasome purification. C) and D) show the representative chromatogram and SDS-PAGE gels for PI31. Fractions 11-16 from B and fractions 22-25 from D were pooled and concentrated for 20S and PI31 respectively. E and F) PI31 and 20S were then incubated together (10:1 molar ratio) and injected into a size-exclusion chromatography column with the representative chromatogram and SDS-PAGE gel shown. From panel F, it is clear that PI31 and 20S are not seen co-eluting (20S and PI31 shown in black and blue box respectively).

5.2.4 Activity assays of PI31 and 20S

In order to evaluate the effect of PI31 on the proteasome activity *in vitro*, the purified PI31-20S samples were compared with free purified 20S proteasome. This required running a SDS-PAGE of the PI31-20S sample and 20S proteasome sample to approximate the relative proteasome concentrations in the samples (Fig. 5.10A). The intensities of the 20S bands were calculated using ImageJ imaging software and appropriate dilutions were made to ensure that both samples used in the assays were of the equivalent 20S concentration. Once sample concentrations had been optimised, an activity assay as described in Chapter three, was conducted. Both the free 20S proteasome and the PI31-20S sample for protease activity of all three proteolytic active beta subunits were tested (Fig. 5.10B). The results showed that PI31-bound 20S had lower activity than free 20S proteasomes. This was consistent with previous literature which demonstrated an inhibitory effect of PI31 on the 20S protease activity. There was a larger difference between PI31-20S and 20S activity observed for ß1 and ß2 proteolytic sites compared to ß5 sbunits. However, in all three cases it was clear that PI31 was inducing an inhibitory effect on the 20S proteasome. However, despite the averages and error bars showing a difference between PI31-20S and 20S activity levels, a Mann-Whitney U-statistical test was also conducted to assess whether these differences were statistically significant. The p-values obtained ($p=0.1$) suggest the differences are not statistically significant, however, the p-values could also be affected by the low sample size ($n=3$) and therefore increasing this would likely increase the confidence of the differences observed. Samples were also incubated with anti-PI31 antibody to confirm the presence of PI31 in the PI31-20S samples (Fig. 5.10C).

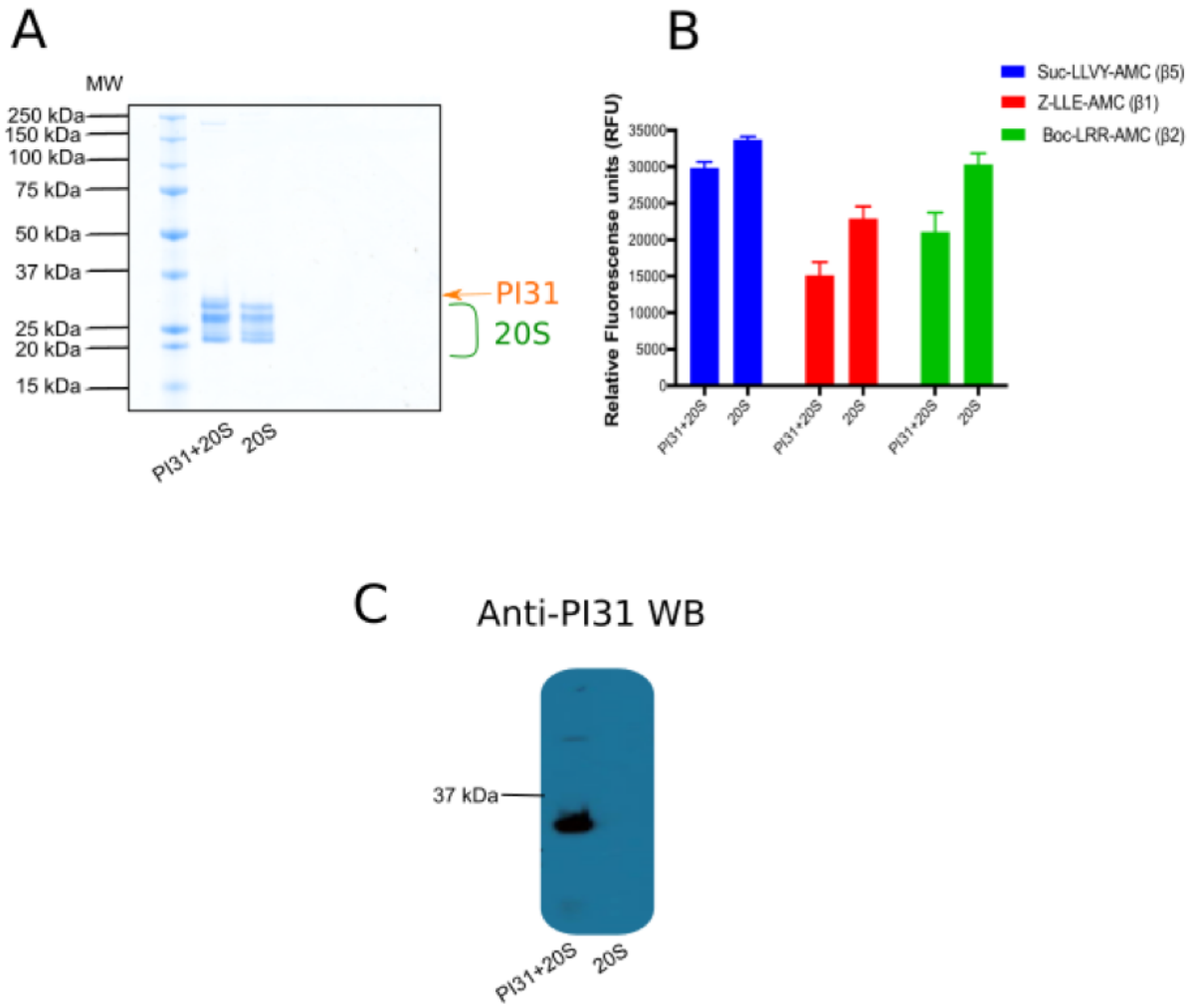


Fig. 5.10: Proteolytic activity of PI31-20S sample. A) A representative SDS-PAGE of PI31-20S and free 20S proteasome is shown. SDS-PAGE was used to compare concentrations of both samples, which was corrected through dilutions based on the intensities of the bands observed. B) The proteolytic activity of PI31-bound 20S was tested against β5, β1, and β2 20S proteasome subunits (blue, red, green respectively in panel A) and compared with free 20S proteasome. The results showed that PI31 had an inhibitory effect on all three β subunits. C) The samples from panel A were treated with Anti-PI31 antibody in a western blot experiment. Experiments were conducted in triplicate and error bars represent the standard error of the mean (SEM). A fold change of proteolytic activity between PI31-20S compared to 20S samples of 1.1, 1.5, 1.4 for β5, β1, and β2 20S proteasome subunits respectively were seen and p-values of 0.1 after performing a Mann-Whitney U- statistical test).

To test the activity of the PI31-20S complex further, a titration assay was conducted. To do this, PI31 and 20S proteasomes were individually purified, as described previously, and an increased concentration of PI31 was incrementally added to a constant 20S proteasome concentration of 30 nM, and the proteolytic activity was measured at each point (Fig.

5.11). These results showed some inhibitory effects observed for the β5 active site only, with β1 and β2 sunits showing no significant change in activity (Fig. 5.11). The variability of the measured activity for β1 was quite large so this result would require further validation. Nonetheless, it is clear that titrating 20S proteasomes with PI31 has little effect on overall activity, despite using molar excess of PI31 compared to 20S proteasome sample. When comparing the results of the titration assays to the previous activity assay in Fig. 5.10, it is clear that mixing PI31 and 20S proteasome *in vitro* results in reduced binding and therefore reduced proteolytic activity change, as compared to co-purified PI31-20S complex from insect cells. The titration results are also consistent with the previous binding experiments of individually purified 20S proteasomes and PI31, where no binding was observed upon incubation and subsequent size-exclusion chromatography purification (Fig. 5.9). These findings therefore highlight the potential transient nature of this interaction or the role of other co-factors or proteins in this complex assembly.

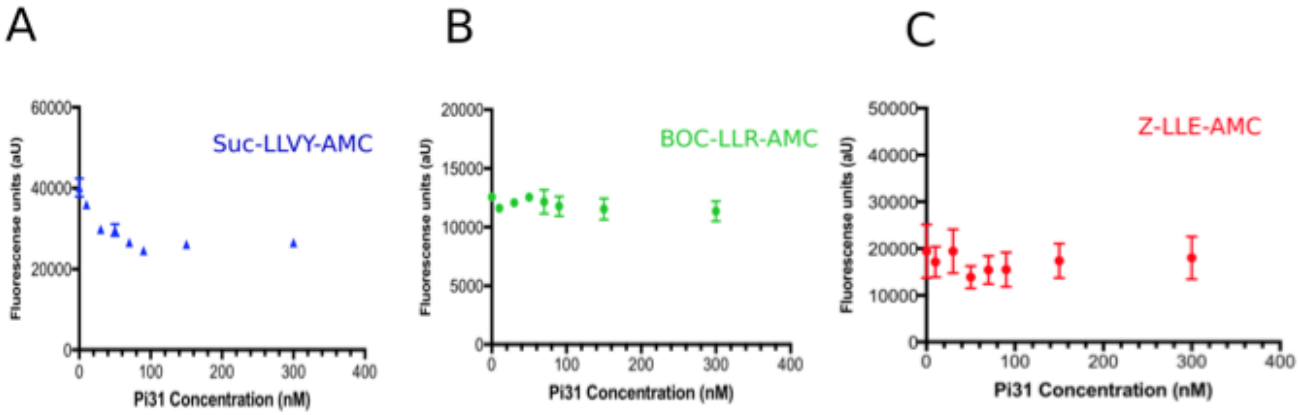


Fig. 5.11: Titration assay of PI31 with the 20S proteasome. Streptavidin-tagged PI31 and streptavidin-tagged 20S proteasomes were individually purified with subsequent removal of tags from both. Increasing concentrations of P131 were added to 0.03 μM of 20S proteasome and measured against the proteolytic activities of the β5, β2, and β1 subunits (shown as blue, green and red respectively. Experiments were conducted in triplicate and error bars represent the standard error of the mean (SEM)

5.2.5 Biophysical analysis of the PI31-20S sample

The PI31-20S sample (as shown in Fig. 5.12) was further analysed by SEC-MALS and interferometric scattering microscopy (ISCAT). ISCAT is a relatively novel technique, that

is complementary to SEC-MALS, and uses light scattered from an object and a reference light field, to measure the molecular weight of the sample in question (Young et al., 2018). The interference between scattered and reflected light is directly proportional to the size of the object. The advantages of using ISCAT compared to SEC-MALS is that very little concentration of sample is required and data collection is faster (Young et al., 2018). Both SEC-MALS and ISCAT were used to investigate the molecular weight of the PI31-20S sample so the same sample could be compared using different techniques. In the SEC-MALS experiment, there was some polydispersity across the peak with an average mass of 896 kDa (Fig. 5.12C). The leading edge of the peak had an average mass of 1.05 MDa, whereas the central region together with the trailing edge had an average mass of 848 kDa (Fig. 5.12C). The ISCAT data showed that the major species can be fitted to a single Gaussian plot with a mass of 745 kDa, and showed that there are other higher and lower species present in the sample (Fig. 5.12D). The data from ISCAT gave similar results to the SEC-MALS and accounts for the polydispersity as there are multiple species present with molecular weights from 748 kDa to approximately 1 MDa (Fig. 5.12). These molecular weights suggest complex formation compared to the approximately 700 kDa and 31 kDa expected molecular weights for the 20S proteasome and PI31, respectively.

The fact that there are several species present in the sample suggested that PI31 is either forming different, higher oligomeric forms with the 20S or that there are other factors within the sample that are binding to PI31 and the 20S, leading to the range of molecular weights observed.

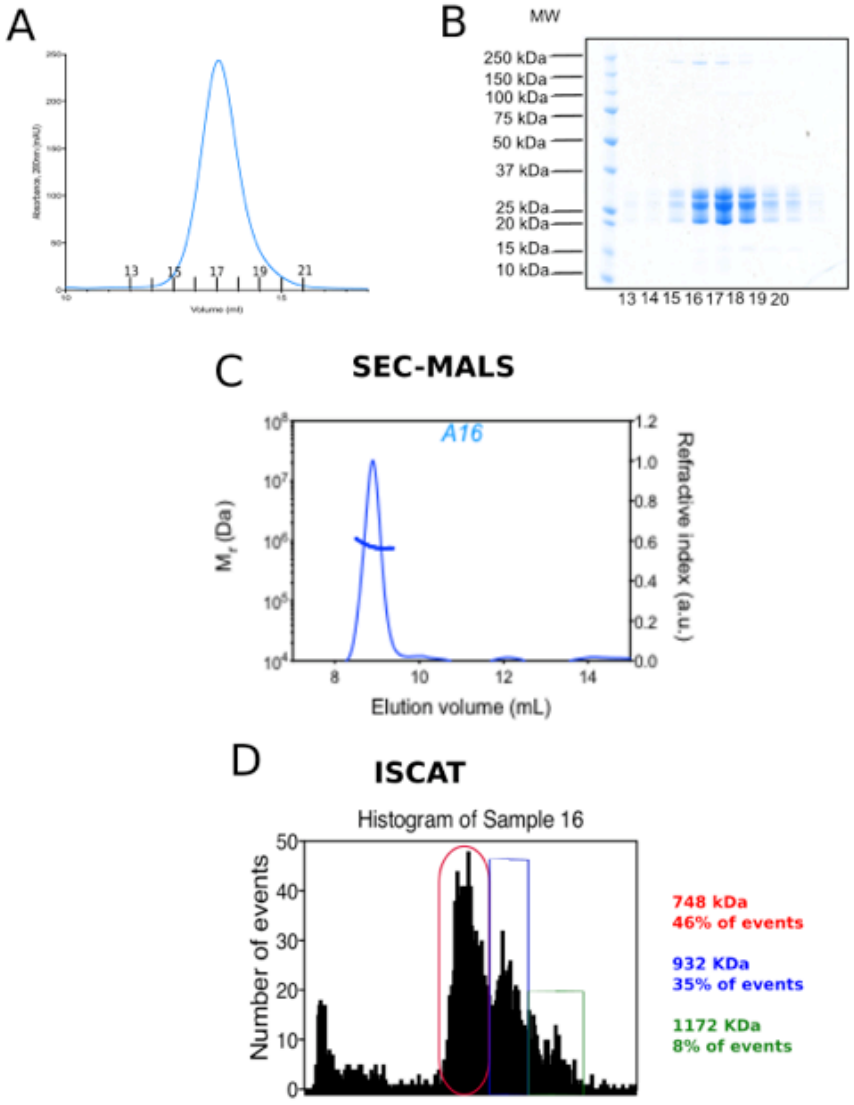


Fig. 5.12: Biophysical analysis of PI31-20S. A and B) PI31-20S samples were purified in which the representative chromatogram and SDS-PAGE gel are shown. Fractions 15-17 from B were pooled, concentrated and subsequent C) SEC-MALS and D) ISCAT experiments were performed. SEC-MALS and ISCAT were in agreement and showed a range of different species present within the sample ranging from 748 kDa to over 1 MDa.

5.2.6 Structural characterization of the PI31-20S complex

5.2.6.1 Cryo-EM background

To investigate the PI31-20S complex further, structural data was obtained using cryo-electron microscopy (cryo-EM). Cryo-EM is used for high resolution structure determination, and provides an orthogonal method of solving structures in addition to x-ray crystallography and NMR spectroscopy. Recently, the advent of the new millennium brought with it greater processing power and camera technology that has since paved the way for a revolution in cryo-EM development and its use as a high resolution structural (Egelman, 2016). High resolution structural analysis had previously been limited to x-ray crystallography and nuclear magnetic resonance (NMR), however, both these techniques have limitations. X-ray crystallography as a structural method is limited to proteins that can form crystals and requires a large input of starting material. Even more critical, the crystals formed are often debated on their physiological resemblance (Garman and Weik, 2017; Snyder et al., 2005). NMR spectroscopy, on the other hand, is limited to small proteins (30-40 kDa) that have been isotope enriched (Wang et al., 2013). The advantages of cryo-EM, therefore, are that protein structures from a range of sizes can be solved (from 50 kDa- 70 mDa), small concentrations of proteins are required, and samples are kept in aqueous solution in their native state and their structures are therefore solved in close-to-native state (Bai et al., 2015; Khoshouei et al., 2017; Kühlbrandt, 2014; McIntosh et al., 2005; Orlova and Saibil, 2011).

In cryo-EM, samples in aqueous solution are applied to EM-grids and then blotted with filter paper to create a thin layer of sample on the grid's surface. Samples are subsequently frozen to preserve the sample's hydrated state. However, to prevent ice crystallization or dehydration, samples are rapidly plunge frozen in liquid ethane resulting in amorphorous ice (below -170 °C) (Fernandez-Moran, 1960). Preservation of the protein's near-native state is important as the structure observed is more likely to be functionally relevant (Taylor and Glaeser, 1974). At temperatures below -170 °C , there is smaller damage inflicted on the sample by the electron beam as the lower temperatures trap free radicals locally, which are made by breaking of chemical bonds, limiting their mobility and reducing radiation damage (Stark et al., 1996). The rapid freezing process developed by Jacques Dubochet led to his sharing of the 2017 Nobel Prize in Chemistry (Dubochet et al., 1982; "The Nobel Prize in Chemistry 2017,") Once the sample is rapidly frozen, it is kept at low temperatures using liquid nitrogen to preserve the sample's vitrified state and subsequently transferred into a transmission electron microscope as described in Chapter three.

After data collection, cryo-EM images can be analysed using a range of data analysis techniques. EM image analysis has come a long way since its initial developments in the 20th century by several academic institutions, including the MRC Laboratory of Molecular Biology. The main concept of single particle cryo-EM analysis involves picking particles from cryo-EM micrograph images, averaging these to increase the signal to noise ratio (SNR), and detemining the different orientations of particles in the images that are then used to reconstruct the final structure (Crowther et al., 1971; DeRosier and Moore, 1970; van Heel and Hollenberg, 1980). After particle picking from the images collected, several rounds of 2D classification steps are conducted to allow for enhancement of the image contrast, which results in identification of features within the particles and helps in improving their alignment and particle picking orientation for 3D reconstruction (Fig. 5.13) (De Rosier and Klug, 1968).

In order to obtain a 3D reconstruction, the Euler angles of each projection of the samples need to be calculated (Fig. 5.13). There are several techniques by which these angles can be determined, including random conical tilt (in which paired images of the particles at high tilt (45-60°) and no tilt are obtained), and projection matching (in which the angles of each projection are deduced from a starting model and back projection algorithms are used to obtain the 3D reconstruction) and ab-initio methods such as angular reconstitution (Radermacher, 1988; van Heel, 1984). This process is then iteratively repeated until the resolution stops improving. The resolution is tracked using Fourier shell correlation (FSC) (Harauz and Heel, 1986; Saxton and Baumeister, 1982). To calculate the FSC, particles are equally and randomly divided into two subsets and two independent 3D reconstructions are calculated from each. The correlation between these two 3D reconstructions is calculated for every Fourier shell and an overall resolution is determined at a fixed set cut off (usually 0.143 FSC) (Rosenthal and Henderson, 2003; van Heel and Schatz, 2005). The higher the resolution, the more structural information of the protein that is available. At lower resolution (above 10 Å), only the overall shape of the complex can be seen. Intermediate resolutions (4-10 Å) allows for subdomain observation and α-helices identification, and at high resolution (<4 Å), β-sheets as well as individual residues can be identified (Chen et al., 2015; Orlova and Saibil, 2011).

Because all biological macromolecules have a level of heterogeneity, 3D classification is an important step that allows for sorting of data into homogenous subsets (Scheres, 2012). 3D classification also identifies smaller differences between particles that cannot be sorted at the 2D classification step (Fig. 5.13). Different image analysis packages are available for the analysis of cryo-EM data such as EMAN, XMIPP, FREEALIGN and more recently RELION and CISTEM (Grant et al., 2018; Grigorieff, 2007; Ludtke et al., 1999; Marabini et al., 1996; Scheres, 2012). New software packages include maximum-likelihood procedures in which each particle contributes to structurally homogenous subsets of particles based on weighted probabilities (Scheres, 2012; Sigworth et al., 2010). Maximum-likelihood methods, along with other methods such as cross-correlation approaches, are very powerful as they allow different conformations of the macromolecule to be separated within a sample and for a 3D reconstruction to be obtained (Bai et al., 2015).

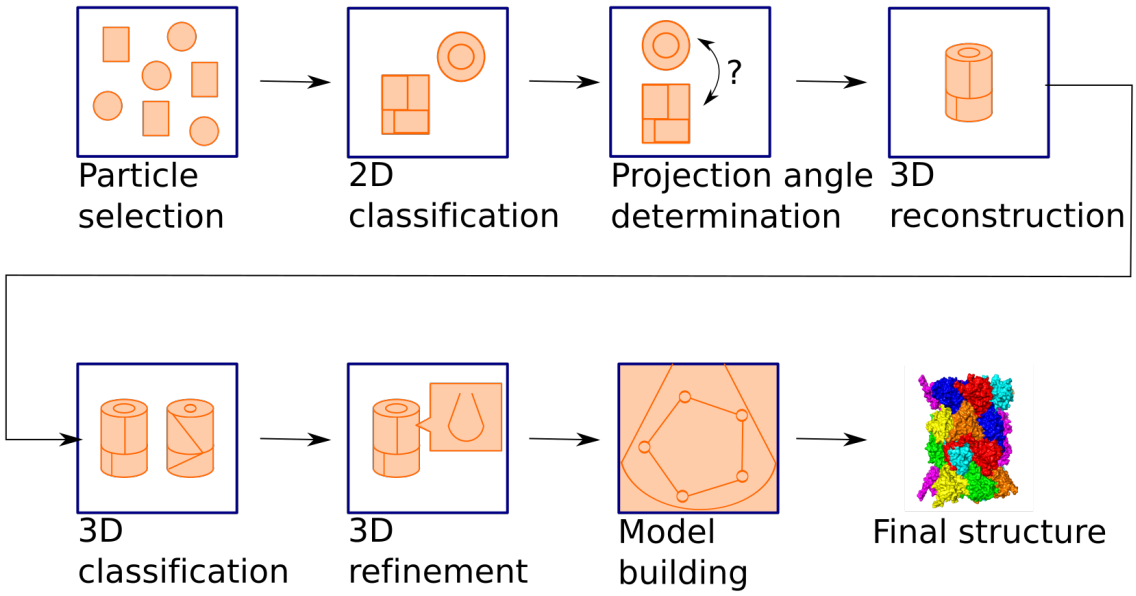


Fig. 5.13: Image processing workflow of PI31-20S sample. Image processing started with picking particles from the micrograph, which are then used for 2D classification. Particles from the best 2D classes are used for 3D reconstruction. Different protein populations in the sample are separated using 3D classification. After 3D classification, 3D classes can be further independently refined to improve detail. The final 3D maps can then be used in protein model building.

Despite the many advantages of cryo-EM, there are also some limitations. The radiation exposure can cause damage to the sample, however, using low electron dose minimises

this damage (Baker and Rubinstein, 2010). A large number of particles are often required due to low signal to noise ratio in cryo-EM images to improve resolution of the final 3D reconstruction. Although the use of direct electron detectors (DDDs) has increased the signal to noise ratio in images, and provided a partial solution to this problem, they still suffer from a less than perfect detective quantum efficiency (DQE) across all frequencies. DDDs detect electrons directly with a fast readout, allowing images to be fractionated into sub-frames and correct for motion-induced blurring (Faruqi and McMullan, 2011; Li et al., 2013; McMullan et al., 2014). Conformational heterogeneity of the sample may also be difficult to separate, resulting in low resolution 3D-reconstructions and the sample may have orientation preferences hindering a correct 3D reconstruction from being obtained (Tan et al., 2017).

5.2.6.2 Initial image processing on the PI31-20S sample

Initial PI31-20S samples were prepared with only one round of PI31-20S SEC as highlighted in Fig. 5.6. The samples were obtained from co-expression of tagged 20S proteasomes with PI31 in insect cells and the same samples were used in the activity assays shown in Figure 5.8. Cryo-grids of the PI31-20S complex (sample PI31-20S from Fig. 5.10) were prepared and data was collected on a Titan Krios electron microscope using counting mode and dose fractionation of a Falcon III camera at the MRC-LMB, with a representative micrograph shown in Fig. 5.14A. The total number of images collected were 1840 micrographs. Datasets of PI31-20S were collected at high magnification of 96K, pixel size of 0.81, using automated data collection (EPU, FEI). The total dose the sample was exposed to was 43.93 electrons/ $Å^2$ and beam-induced drift was corrected using whole-subframe alignment (Li et al., 2013). Contrast transfer function (CTF) parameters of the resulting motion corrected frames were then determined using CISTEM software (Grant et al., 2018).

The contrast transfer function is a mathematical description of imperfections in an electron microscope that limit the information content in the formed image. Contrast, namely how well the sample can be distinguished from the background, is not uniformly transferred to the image but rather varies as a function of the phase of the scattered waves. This variability in contrast transfer arises primarily due to spherical and other aberrations in the

microscope optics, and also depends on the applied defocus (Wade, 1992). The purpose of CTF correction is to restore as much of the lost information as possible by collecting images at a variety of defocus values. This enables spatial frequencies with zero or near zero contrast at a particular defocus value can be filled in with non-zero contrast values from an image acquired at a different defocus. High resolution structures are unable to be obtained without CTF correction as the negative phase contrast results in compromise of the high resolution features of the structure (Thon, 1966; van Heel et al., 2000). After motion correction and CTF estimation, I closely inspected the images and discarded any that had uncorrected drift, high contamination, low contrast, or crystalline ice. This resulted in a final selection of 649 images. Data processing was then continued using CISTEM software (Grant et al., 2018).

Initially, CISTEM software was used for particle picking. However, results showed that some particles that were clearly visible on the micrograph had not been chosen and some particles that had contamination or noise had been picked. SHPIRE-crYOLO (Cryolo) software was therefore used instead to pick particles more accurately (Wagner et al., 2019). Cryolo is a novel particle picking software that uses deep learning techniques to pick particles. This is a different approach from CISTEM, which uses a combination of standard noise models, reduced representation of the reference and statistical analysis to identify artifacts (Sigworth, 2004). The advantages of Cryolo include high precision and speed (Wagner et al., 2019). The program is trained on a subset of images (in my case 30 images were used for training), where particles were originally hand-picked. After training, particles from the rest of the images are then automatically picked. The results showed an improvement in particle quality, as indicated by less noise and fewer contaminants, and a good density of particles per micrograph were picked. This resulted in 35546 of particles picked (Fig.5.14A). After close inspection of the autopicking, it was evident that most particles had been selected, without an overload of false-positives. The resultant particles were then subjected to 2D classification using the CISTEM program (Fig. 5.14B) (Grant et al., 2018).

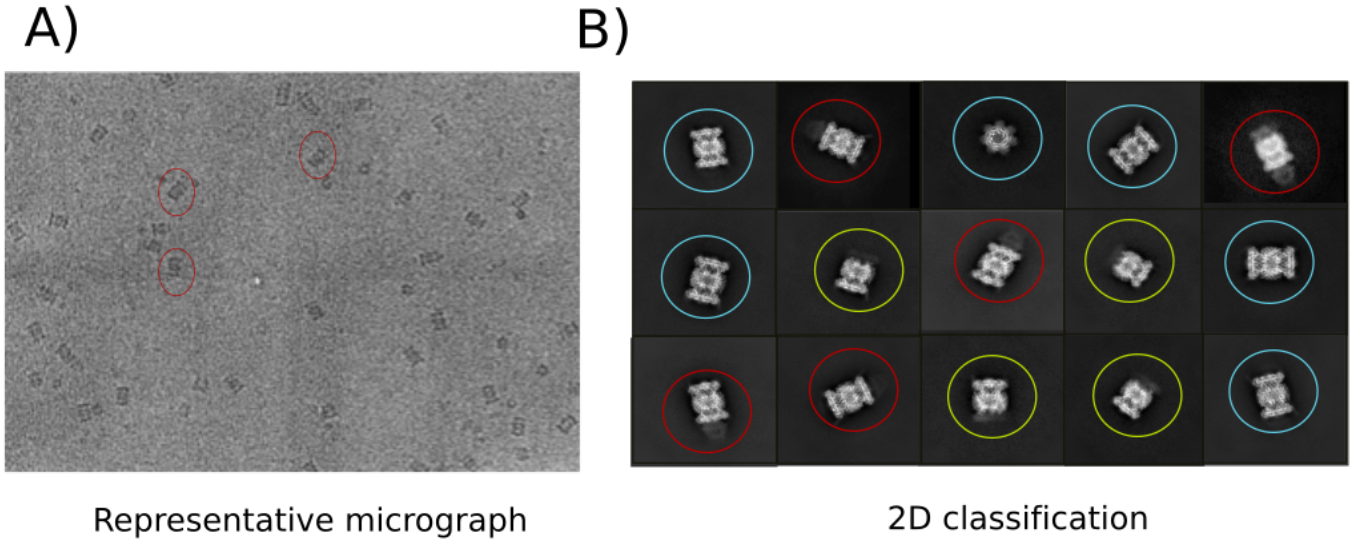


Fig. 5.14: Cryo-EM data collection and 2D classification of initial PI31-20S sample. Cryo-grids of PI31-20S samples were prepared and data was collected on FEI Krios electron microscope. A) A representative micrograph of the sample where examples of picked particles are shown in red circles. Particles were picked and processed using CISTEM software. B) 2D classification results from the processing. 20S proteasomes (blue), capped 20S proteasomes (red) and half proteasomes (green) were seen.

Several rounds of 2D classification, of 50 classes, were conducted, to remove particle picking false positives and bad particles, or those that fail to classify into classes with clear, well defined molecular details. The final round of 2D classification, of 17695 particles, showed mostly 20S proteasomes, some extra 'cap-like' structures on some classes and half-proteasomes which had not been seen in the lab before (Fig. 5.14B). It was clear at this stage that I had to improve the PI31-sample and be more selective with the fractions used to make cryo-EM grids, in order to reduce the number of empty 20S proteasomes seen.

5.2.6.3 Image processing with an improved PI31-20S sample

To improve the PI31-20S sample, two additional rounds of SEC were conducted, instead of just one as was previously done (as shown in Fig. 5.6). Cryo-grids of the PI31-20S complex (sample PI31-20S from Fig. 5.10) were prepared and data was collected as before on a Titan Krios electron microscope using counting mode and dose fractionation of a Falcon III camera at the MRC-LMB with 0.81 pixel size and 96K magnification (Fig. 5.15). The total number of images collected was 1861. The same image processing

workflow as described in Section 5.2.6.2 was used resulting in a total of 1660 images after CTF correction.

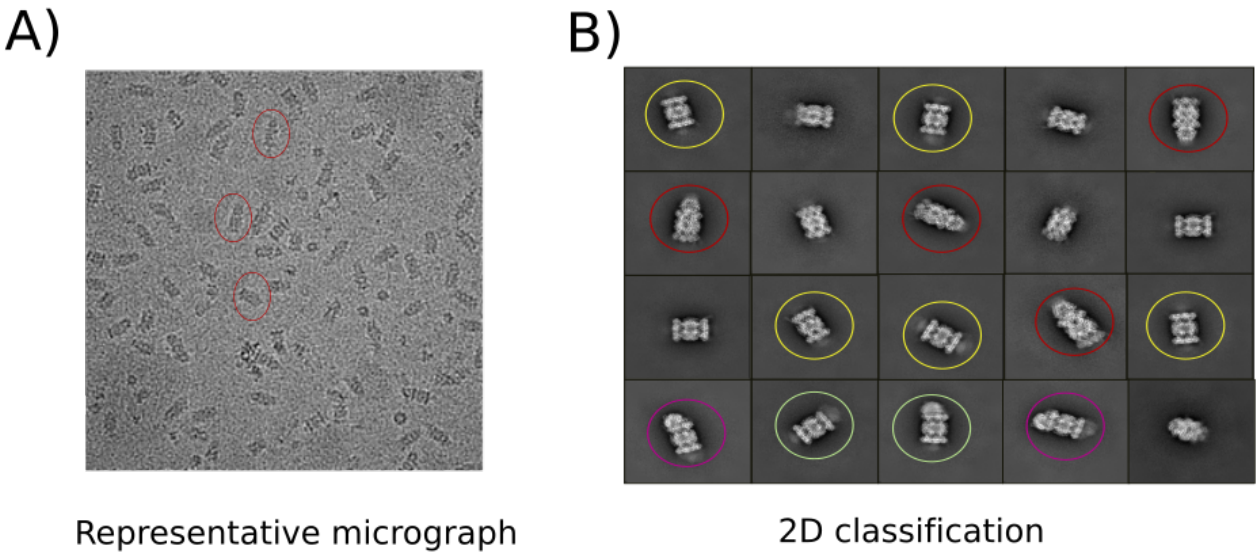


Fig. 5.15: Cryo-EM data collection and 2D classification of optimised PI31-20S sample. Cryo-grids of PI31-20S samples were prepared and data was collected on FEI Krios electron microscope. A) A representative micrograph of the sample; particles were picked (examples shown in red circles) and processed using CISTEM software. B) 2D classification results from the processing. Several different capped complexes were seen (shown in different colours).

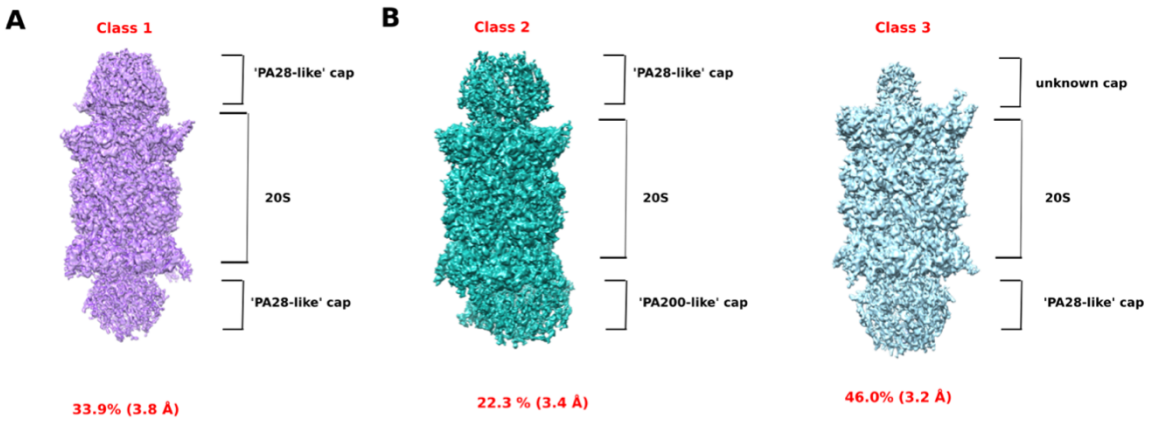


Fig. 5.16: **3D classification of the PI31-20S sample**. 3D classification from the merged PI31-20S sample datasets resulted in identification of 3 classes from 38, 583 particles with different caps bound to the 20S proteasome. Class 1 and Class 2 had caps that resembled PA28 and PA200 regulators respectively. In class 3, there was an unidentifiable cap observed. Classes 1, 2, 3 consisted of 33.9%, 22.3, and 46.0 % respectively with estimated resolutions of 3.8 Å, 3.4 Å, and 3.2 Å respectively. The representations of structures were created using UCSF Chimera (Pettersen et al., 2004). Englarged maps available in the appendix.

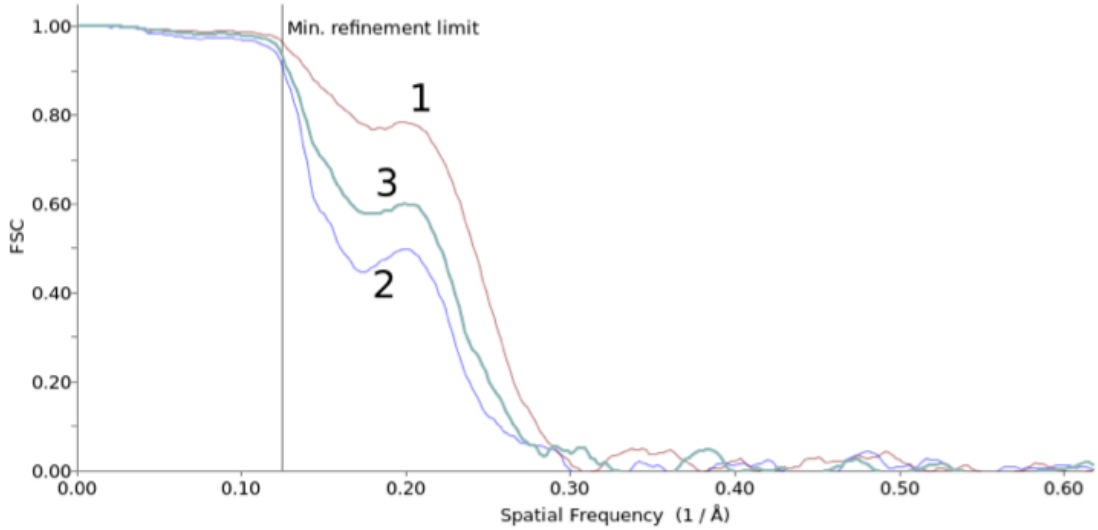


Fig. 5.17: Fourier shell correlation plots of three maps. Fourier shell correlation plots of all three classes (numbered 1, 2, 3 on the plot) are shown, which was used to estimate resolution of the three classes from 3D classification presented.

Several rounds of 2D classification, as described above with 50 classes, were conducted. However, this time the final round of 2D classification, of 50 classes and 38,583 particles, showed highly detailed molecular densities (Fig. 5.15B). Some classes had small densities attached to the 20S, whereas bigger caps were also identified. After I observed that further 2D classification was not reducing noisy or contaminating particles, I performed 3D classification on the resultant particles of 38,583 particles using the 20S proteasome from Rego et al. (2019) as a starting reference.

The obtained resolution of the 3D classification was insufficient to visualize finer details in the caps and to understand their resemblance to other caps. A second dataset of the optimised PI31-20S sample was therefore collected to increase the number of particles. The same procedure was used to process particles to the 2D classification stage. Briefly, a total of 1816 images were collected and 1510 images with good CTF were retained. Particles were then extracted and combined with the previous data set, doubling the number of particles to a total of 104,090. Five rounds of 2D classification were performed to ensure the particles from the merged datasets agreed with the datasets before merging, and to get rid of unwanted noise. 3D classification of the merged data set of 67,144 particles from the final 2D classification round was then performed. After 2D

classification, an initial 3D auto-refine step was conducted with the published 20S proteasome from Rego et al. (2019) as a reference and no symmetry was imposed. The coordinates of the centre of the cap were then obtained using UCSF Chimera and a mask was created to perform focussed classification with C2 symmetry. Focussed classification is a process that involves masking of a region of interest so that structural variability can be ignored in other regions (Scheres, 2016). This procedure is especially useful for structurally heterogeneous samples and can be used to separate complexes that have either conformational differences or are different complexes altogether (Scheres, 2016).

3D classification resulted in three classes with three different caps (Fig. 5.16). The estimated resolution of the caps had improved from over 4 Å to 3.2-3.4 Å as calculated from their respective fourier shell correlation plots (Fig. 5.17) with most classes consisting of the hybrid class three described (Figs. 5.16).

5.2.6.4 Interpretation of 3D models

The final three classes showed three different sets of caps. Class one had a striking resemblance to the PA28 structure but with one cap better defined than the other and is referred to as 'PA28-like'. The second class was a hybrid consisting of two different caps, one resembling PA200 and the other PA28 (Fig. 5.16). This class is referred to as 'PA200-like'. The third class was also a hybrid consisting of an unidentified small cap at one end of the 20S proteasome and a PA28-like structure on the other end, and is referred to as 'PI31-like' hybrid (Fig. 5.16). The estimated resolutions for each class were in the range of 3.2-3.4 Å (Fig. 5.16 and 5.17).

Due to its similarity in shape, the PA28-20S model by Metcalfe et al. (2019) was docked onto the 'PA28-like' map (Fig. 5.18). As described in Chapter one, PA28 is a proteasome regulator with a heteroheptomeric structure that contains a central pore (Xie et al., 2019; Huber and Groll, 2017). Although the biological significance of this pore remains to be elucidated, as described previously, it may allow access of 20S proteasome substrates that will be proteolytically cleaved (Stadtmueller and Hill, 2011). As suspected, the cap structures in class one bore striking similarity to the published PA28 map. This was assessed by the good secondary structure fitting between the published PA28 model and

the class one map, however, local resolution does not allow for fitting of all side chains which will need to be improved by further classification (Fig. 5.18). In published PA28αβ structures, this central pore is not occluded, however, in the class one structure obtained from 3D classification there is extra density observed that occludes access to the pore of the PA28 cap (shown in pink in Fig. 5.18). Upon closer inspection, using UCSF Chimera software (Pettersen et al., 2004), the extra density observed at the pore of the PA28-like cap resembles a ring structure. This makes the pore of the PA28-like class much more constricted and narrower compared to the PA28 model, however, the density in the cap could not be defined to the residue resolution. It therefore, remains unclear what this density could be. The consequences of this constriction could mean regulation of the PA28 cap in that substrates destined for degradation cannot enter, or that PA28 itself is structurally constricted and impairs 20S activity. Better defined maps will help to distinguish between these two different possibilities. Further, there was extra density observed within the β chamber of class 1 (Fig. 5.18), however, its density could not be further resolved to identify the residues or co-factors involved. The consequences of this extra density could affect 20S proteolytic activity.

Similarly, the PA200 model was also docked onto the 'PA200-like' structure (class 2) (Fig. 5.19). As before, the PA200 fitted well onto the map, as shown by the fitting of secondary structure elements of the published PA200 and PA28 models (Fig. 5.19). There is also extra density seen in the PA28-like cap in this class (shown in pink in Fig. 5.19). The fact that both PA28-like and PA200-like activators are found bound to the 20S proteasome when PI31 is co-expressed with the 20S proteasome is interesting as it could mean that PI31 is having a regulatory role on these activators or acting as chaperones to allow for PA28 or PA200 binding. This phenomenon is not observed when purifying PA28 or PA200 with the proteasome complex and other activators are not seen.

The third class, referred to as 'PI31-like', consisted of an unknown cap on one end of the 20S proteasome, and the presence of a PA28-like cap on the other end is also intriguing (Fig. 5.20). The unknown cap exhibited low resolution and therefore its origins cannot be determined. However, this may be PI31 itself as the structure is roughly a similar size to the dimer form of PI31. It is important to note that the global B factor applied in the post-processing step is inaccurate for areas of low resolution and therefore may have led to a

noisy appearance of the novel cap. Further processing should take this into account and calculate local B-factor. In more refined maps, it would be useful to compare thte Phyre predicted structure with the novel gap and use high pass filtering to help recover some domain or back-bone information. The new cap has not previously been seen with the 20S proteasome and its functions could be regulating the 20S proteasome activity. The fact that there are several complexes observed is in agreement with the biophysical data in which lots of molecular masses are identified in the PI31-20S sample.

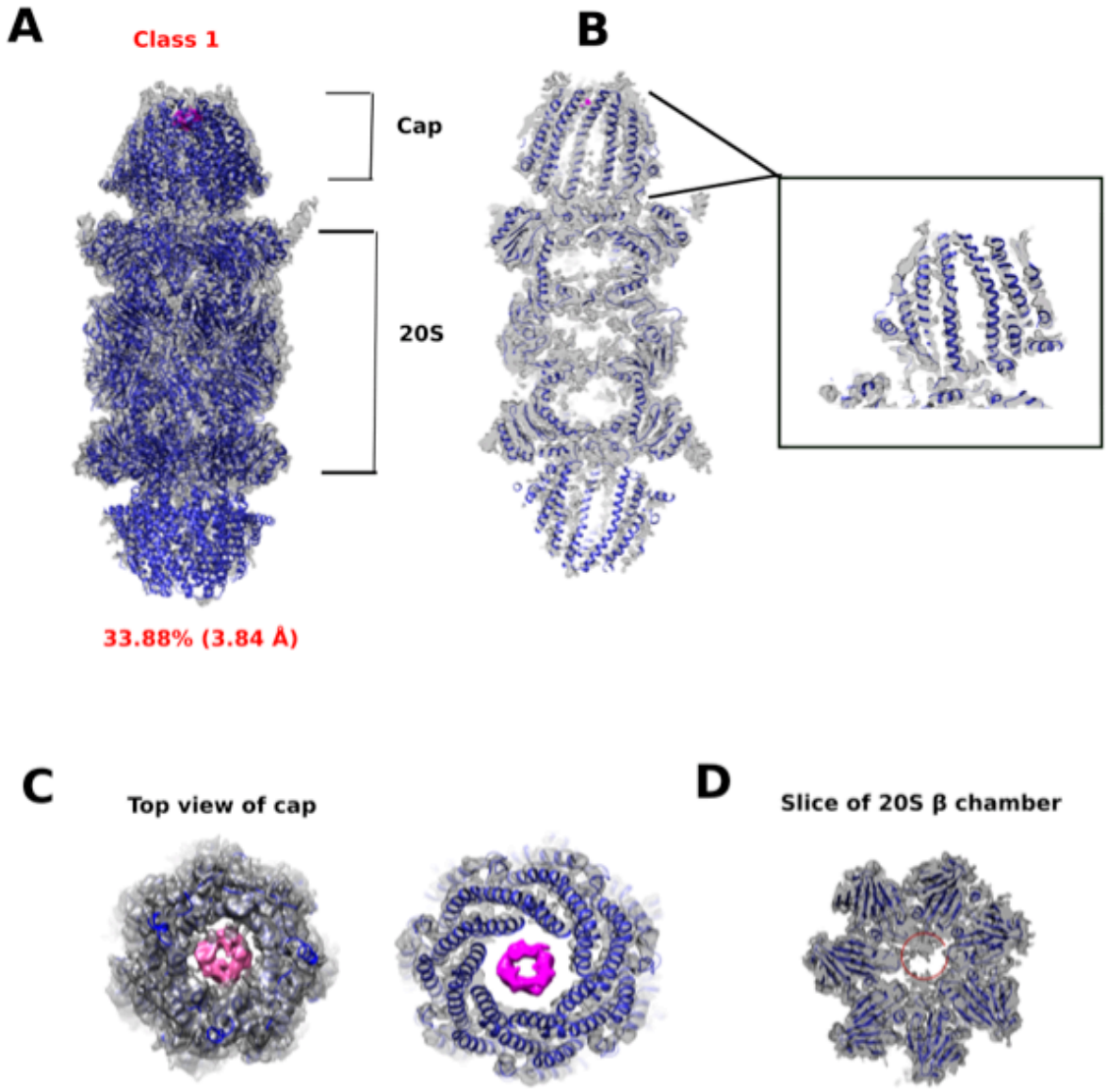


Fig. 5.18: Structural analysis of Class 1 particles. Analysis of Class 1 obtained from the 3D classification of PI31-20S sample is shown A and B) the map of class 1 (grey) and the PA28-20S model docked onto the model (blue). Extra density at the top of the cap is shown in pink. C and D) An extra density (pink) present in the class 1 map (viewed from the top of the cap-like structure) makes the opening of the cap much more constricted than the PA28 cap (blue). D) A slice of the β subunits of the 20S chamber with the same subunits of the PA28-20S model docked on top. There are clear fragments of extra density observed in the PI31-20S class 1 map compared to that of the PA28-20S (red circle). The extra density in the 20S pore is contoured to the same level as the rest of the map, such that any differences observed are real differences and not due to visualization effects. The representations of structures were created using UCSF Chimera (Pettersen et al., 2004). An enlarged map of class one is provided in the appendix.

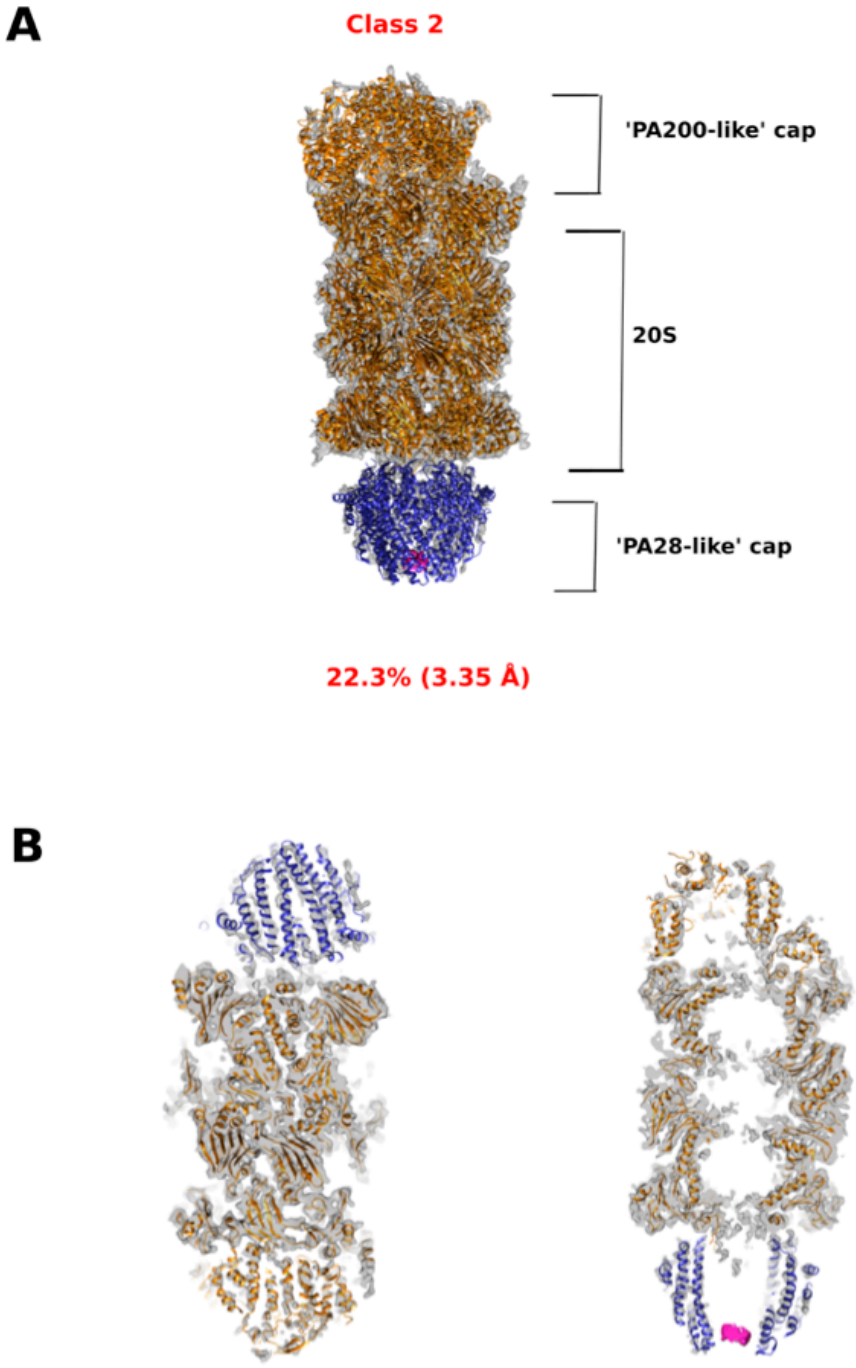


Fig. 5.19: Structural analysis of Class 2 particles. A) Class 2 obtained from the 3D classification of PI31-20S sample reconstruction (grey) with the P200 model (PDB: 6REY) (orange) docked onto one side of the map, whereas the PA28 map docked onto the other side (PDB: 6MUX) (blue). An extra density is seen to occlude the PA28 pore (shown in pink). B) Slices of the class 3 map (grey) with the PA28 and PA200-20S models docked. There is good fitting between the map and models as can be seen in the secondary structure fitting. The extra density in the 20S pore is contoured to the same level as the rest of the map, such that any differences observed are real differences and not due to visualization effects. The representations of structures were created using UCSF Chimera (Pettersen et al., 2004). An enlarged map of class two is provided in the appendix.

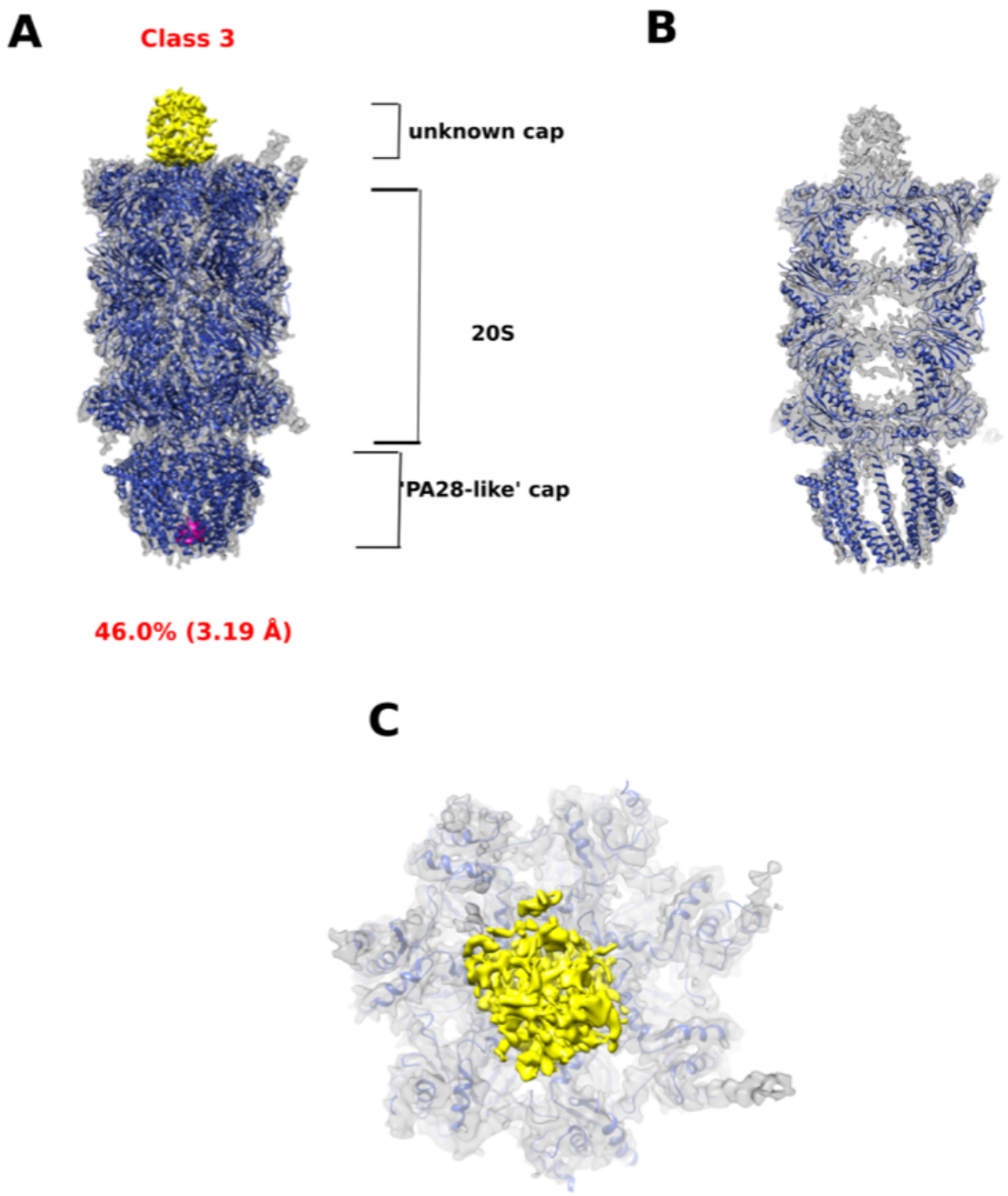


Fig. 5.20: Structural analysis of Class 3 particles. Class 3 particles obtained from the 3D classification of PI31-20S sample map (grey) with the PA28 map docked on top (PDB: 6MUX) (blue). An extra density seen in the PA28-like cap as shown in pink. An unknown cap-like structure (yellow) is present in one of the ends. B) A slice of the class 3 map (grey) with the PA28 and PA200-20S models docked (blue) with good secondary structure fitting between the grey map and the PA28-20S structure. The extra density in the 20S pore is contoured to the same level as the rest of the map, such that any differences observed are real differences and not due to visualization effects. C) Top view of class three showing the cap-like density occupies a central position on the 20S proteasome. The representations of structures were created using UCSF Chimera (Pettersen et al., 2004). An enlarged map of class three is provided in the appendix.

5.3 Discussion

From the current literature available, PI31 remains an intriguing protein with many interesting roles proposed by researchers in the field. The fact that PI31 has been linked to proteins in a variety of different pathways highlights the significance of the protein and its far-reaching effects. The conserved nature of PI31 and its detrimental effects if deleted in Drosophila also shows a potentially evolutionary role in cells that is worthy of further investigation. However, based on current data, it is clear that PI31's most intriguing role is in relation to the proteasome. PI31 has been shown to inhibit 20S proteasome activity and has been implicated in the 26S proteasome assembly (Bader et al., 2011; Chu-Ping et al., 1992). The fact that some studies have shown potential antagonistic roles of PI31 compared to known 20S activators, such as PA28, suggests a possible regulatory role of PI31 with other known proteasome-interacting proteins or complexes. PI31's role in the immunoproteasome, in regulating antigen presentation, is not only interesting from a molecular point of view, but also of therapeutic importance. My results therefore shed light into a relatively unknown protein which may have important implications in the proteasome field.

From my results, it was evident that there is an interaction between PI31 with the 20S proteasome. This was clear from the initial experiment in which tagged PI31 was co-expressed with untagged 20S, allowing the pull-down of the tagged PI31 co-eluted 20S proteasome (Fig. 5.2). The fact that the PI31-20S complex remained intact despite several rounds of purification and concentration steps highlighted the stability and robustness of this complex. However, interestingly, the binding experiment in which PI31 and 20S were individually purified, incubated and analysed by size-exclusion chromatography showed that the PI31-20S is not able to assemble *in vitro,* despite using high concentrations of PI31 and ensuring 20S proteasome saturation in the *in vitro* binding experiment (Fig. 5.9). This raises several questions and suggests that the PI31-20S complex may need other proteins or co-factors that are found in insect cells during overexpression and therefore would only be formed *in vivo*. An alternative explanation of why the PI31-20S complex does not assemble *in vitro* could be due to PI31 acting as an adapter or chaperone protein

and binds to not yet fully formed or active 20S proteasomes. Once proteasomes are fully formed or bind to activating caps, such as the PA28, PA200 or 19S particles, then PI31 is no longer able to bind. Identification of many different populations of PI31-20S cryo-EM samples, such as PA200 and PA28, suggests that this explanation could hold true as these complexes may be depending on PI31 to aid in 20S binding. Further, proteolytic activity assays showed that the PI31-20S sample had lower proteolytic activity compared to 20S alone samples. However, when PI31 is titrated to 20S proteasomes that were separately purified, PI31 had no significant effect on proteolytic activity levels. This stark difference between proteolytic activity assays is consistent with the fact that PI31 and 20S can only be purified together when co-expressed *in vivo*, suggesting that PI31 may be not only affecting proteolytic activity by directly inhibiting 20S proteasomes itself through binding, but also affecting the distribution of other activators that bind the 20S, such as PA200 and PA28 seen in my structural analysis. It is this global reorganisation that could be resulting in the slightly reduced activity when adding PI31 to the sample. It would be interesting to conduct kinetic studies, such as surface plasmon resonance, of 20S and PI31 to understand how tightly these two proteins bind and to what extent other factors, such as PA28, PA200, affect this binding.

The most efficient PI31-20S purification strategy involved co-expressing tagged 20S proteasome with untagged PI31 and 20S chaperones in insect cells (Fig. 5.6). This resulted in the highest PI31-20S yield, with no endogenous insect 20S contamination. By pulling down on the 20S proteasome instead, I solved the problem of having 20S from insect cells contaminating 19S-20S recombinant complexes. Western blots against PI31 allowed for verification of PI31 proteins within samples, as well as the elution fractions of PI31-bound 20S compared to free 20S in a size-exclusion chromatogram. Attempts at optimising this purification strategy showed the importance of using higher concentrations of glycerol as this improved the sample's ability to concentrate and thus increased the yield of PI31-20S obtained (Fig. 5.7). However, changing concentrations of EDTA had no effect on PI31-20S yield despite the fact that Mg^{2+}/Ca^{2+} exposure had previously been shown by the da Fonseca lab to impact 26S proteasome purifications.

In addition, despite attempts to co-express flag-tagged 20S proteasome with Twin-Strep tagged PI31 and pull down on each component, the yield obtained from flag affinity purifications was too low and was therefore not a purification strategy that was pursued

further (Fig. 5.6). Further, additional experiments investigating PI31's oligomeric state showed a typical monomer-dimer equilibrium as reported by others (Fig. 5.3) (Kirk et al., 2008). Interestingly, PI31 bands with lower molecular weight than the full-length protein was also observed by SDS-PAGE during purifications. These were confirmed by MS analysis that showed the loss of the N-terminal region. However, increasing the concentration of protease inhibitors within the lysis buffer reduced the cleavage seen in later experiments, suggesting that they were a result of protease cleavage.

Molecular masses obtained by both SEC-MALS and ISCAT showed the PI31-20S sample to be in the region of 748 kDa to approximately 1.1 MDa, suggesting the presence of other complexes bound to 20S proteasomes in the presence of PI31 (Fig. 5.12). The range of molecular masses of the PI31-20S seen in the biophysical data corresponds to the range of structural species identified (Fig. 5.12 and Fig. 5.13). Of these structural components in the PI31-20S sample, there is striking similarity between PA28 and PA200 in two out of three structures obtained. By docking the published structures on PA28 and PA200, it is evident that there is good fitting between the secondary structure components of PA28 and PA200 models and the obtained structures from the PI31-20S sample. However, in the case of PA28, there is a clear difference between this complex obtained with the 20S and the conventional PA28-20S complex as there is extra density observed (Fig. 5.17 and Fig. 5.18). The extra density observed in the middle of the 'PA28-like' structure as well as in the middle of the 20S chamber is intriguing and requires further investigation (Fig. 5.17). It is worth noting that the structure of insect PA28 is unknown and thus the extra density observed may be corresponding to PA28 from insect cells. However, it may also be linked to PI31 itself.

As well as PA28 and PA200-like structures, there are also smaller cap complexes seen. These could correspond to PI31 itself, but its low resolution and incomplete structural information on PI31 does not allow a conclusive assignment. These small caps are also seen in a hybrid complex with the PA28-like caps on the other side of the 20S proteasome (Fig. 5.19).

In both the case of PA28 and PA200, it is interesting to see both these 20S proteasome activators purifying with PI31 preparations, which has not been reported before. This is particularly striking as the 19S RP is not seen in these purifications. It is therefore possible

that PI31 may be acting as an adaptor that promotes the 20S proteasome binding to alternative caps, like PA28 or PA200, which are not needed for the 19S RP. PI31 may be keeping the proteasome in a 'dormant state' as is seen in the lower activity of PI31-bound 20S, unless activators, such as PA28 or PA200, bind. To investigate this further, PI31 could be titrated against different concentrations of 26S proteasomes to show if there is any effect on proteolytic activity. This would indicate whether PI31 is able to bind to 26S proteasomes and if PI31 has any effect on the 26S proteasome. Interestingly, in the presence of PI31, the 20S proteasome prefers to associate with PA200 and PA28 rather than the 19S, while in the absence of PI31 the 20S proteasome strongly and preferably binds to the 19S. This is a clear indication that PI31 has an adaptor or chaperone-like function. As introduced previously, PI31 has been implicated in spermatogenesis and the immune response; both PA200 and PA28 have also been associated with spermatogenesis and inflammatory responses respectively. This provides further clues into which particular pathways PI31 may be regulating. The association between PI31 and the 20S proteasome may be transient and my results presented in this chapter provide first insights into this interaction.

Chapter 6 Concluding remarks

In summary, I have investigated three proteins, UBR4, P97 and PI31, and their relationship to the 26S and 20S proteasome complexes using an integrative structural biochemistry approach. My experiments involving UBR4 showed that UBR4 co-eluted with proteasome sub-complexes when purifying endogenous human 26S proteasomes from HEK293F cells. Further biochemical separation techniques, such as ion exchange chromatography, calmodulin affinity purification and glycerol gradient centrifugation, showed that UBR4 co-eluted specifically with the 19S regulatory particle and that this interaction was strong since any experimental attempt to break this interaction was unsuccessful. SEC-MALS analysis of UBR4-19S samples showed molecular weights in the region of approximately 1.4- 1.8 MDa, suggesting a sub-complex of proteasome subunits with UBR4. Preliminary image processing of negatively stained UBR4-19S samples showed a heterogeneous sample consistent with what was seen on SDS-PAGE gels after affinity chromatography. Further, 2D and 3D classification showed globular proteins, however, their low resolution prevented further characterisation. Ubiquitin detection methods such as ubiquitination western blots and ubiquitin pull down assays of UBR4-19S samples showed that these samples are found in ubiquitinated states, however, the functional relevance of these findings remain intriguing. In parallel, CRISPR methods to tag endogenous UBR4 was successful and provided means to directly purify UBR4 from HEK293F cells directly. Using tagged UBR4 cell lines followed by affinity purification methods, proteasome subunits including 20S proteasomes were observed in negative stain EM images and MS analysis of UBR4 samples. This provided further support for a direct interaction of UBR4 with proteasome sub-complexes. Comparison of 20S proteasome proteolytic activity of these samples against known concentrations of 20S proteasome samples showed that tagged UBR4 samples contained traces of proteolytic activity. However, the very low yields prevented further structural or biochemical characterisation. As discussed previously in Chapter 3, further experiments involving overexpressed UBR4, potentially in insect cells, would shed more light into the nature of the interaction of UBR4 with proteasome subunits.

My investigation into P97 and its potential interaction with proteasomes involved co-expression of P97 with constitutive or immuno - 20S proteasome in insect cells. Despite successfully showing both 20S and P97 co-eluted from insect cells, the interaction is either weak or transient or that it involves other cofactors in order to stabilise the interaction. Once P97 was re-cloned using a codon-optimised gene sequence the P97 expression yield improved, however, this did not affect the overall yield of the P97-20S complex. As discussed in Chapter 4, identification of the cofactors required for p97 to form a stable complex with the proteasome is of paramount importance.

Finally, my investigation into PI31 showed that it interacted with recombinant human 20S proteasome. I showed that this interaction required co-expression as *in vitro* reconstitution of individually purified partners and further binding assays did not show a PI31-20S interaction. This was consistent with proteolytic activity assays in which proteolytic activity of all three ß subunits were inhibited in the PI31-20S sample compared to the control 20S proteasome sample, but were not significantly affected when PI31 was titrated into purified 20S proteasome. In general, this suggests that external cofactors or *in vivo* binding of PI31 binding to 20S is essential for the PI31-20S interaction. Biophysical analyses using SEC-MALS and ISCAT of the PI31-20S sample showed a broad range in the estimated molecular weight (748 kDa to approximately 1 MDa). Such an observation was consistent with the cryo-EM analysis that showed that the sample was heterogeneous and contained 20S proteasomes apparently bound to different regulators. Cryo-EM data and initial 2D classification immediately showed the presence of PA28-like proteasomes, PA200-like proteasomes as well as a proteasome species with extra density within the 20S proteasome chamber and PA28-like cap that could likely correspond to PI3I. Interestingly, this is the first time all three regulators have been co-purified suggesting that PI31 may have an important role in the regulation of proteasome activators. Reconstruction from 3D classes showed extra density in the 20S ß chamber particularly in PA28-like structures. Hybrid proteasomes of an unknown cap and PA28-like as well as PA200-like were also seen. However, further classifications are needed to confirm the extent of these populations. My ongoing research is aimed at improving the overall resolution of these three cryo-EM maps to better understand the atomic level molecular details of the proteasome-ancillary factors interaction.

Therefore, the biochemical and cryo-EM analysis presented in this dissertation have paved the way for the full characterisation of PI31 with the 20S proteasome and opened doors for potential studies of its role on the regulation of other proteasome activators.

In summary, I believe that my PhD contributes to the proteasome field and sheds light into the ancillary proteins the proteasome is thought to interact with.

1) Along with Dr. Kisonaite, I have optimised the purification of endogenous, human 26S proteasomes in HEK293F cells.

2) I have investigated the biophysical properties of UBR4 protein, set up a CRISPR system in which endogenously tagged UBR4 can be purified from HEK293F cells and conducted experiments which provide support for UBR4 binding to the 26S proteasome.

3) I have investigated the P97 protein, showed through binding experiments that P97 co-elutes with the 20S proteasome, suggesting other cofactors may play a role in stabilising this interaction due to the low yields obtained from P97-20S co-expression in insect cells.

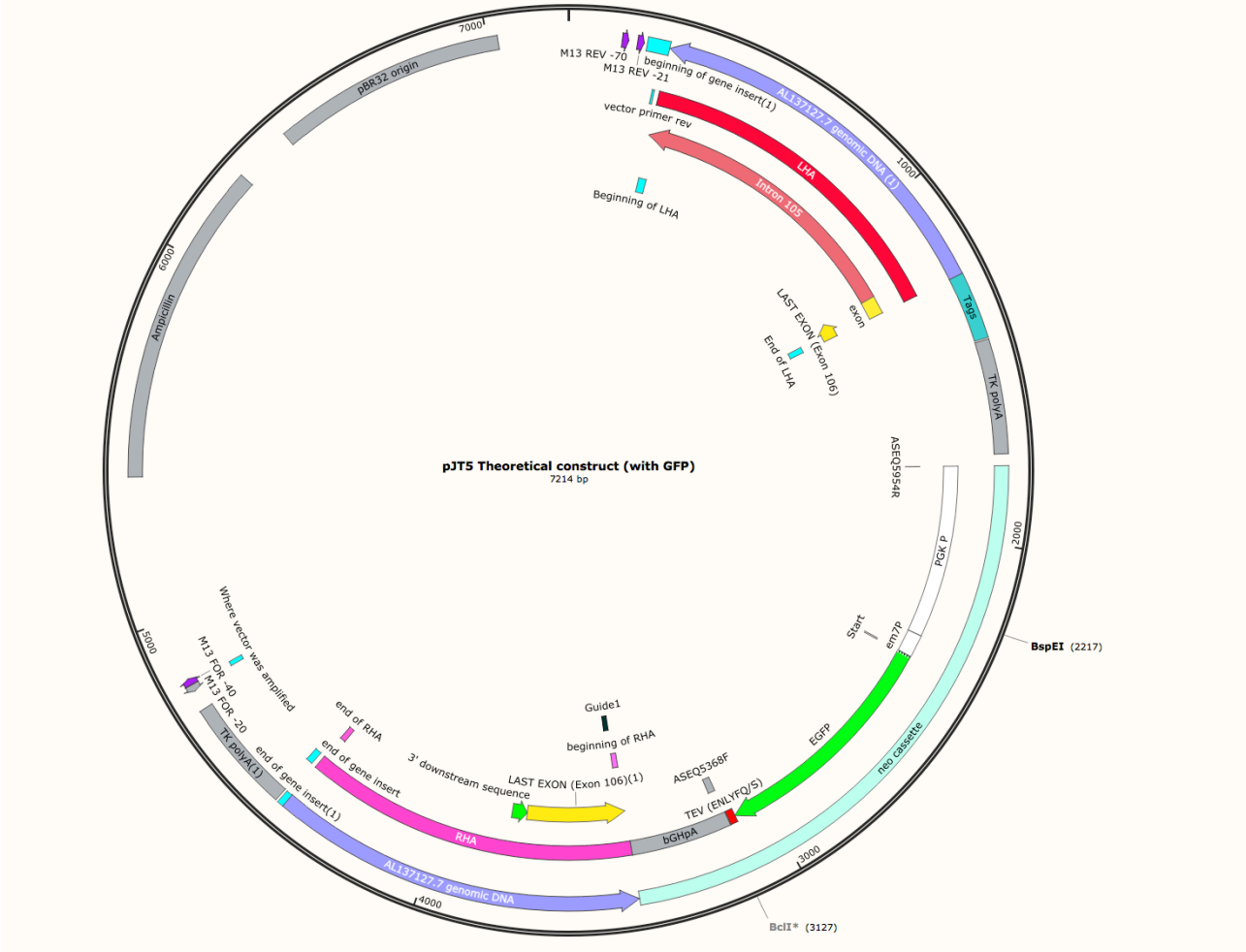
4) Finally, I have investigated the PI31 protein, showed that it inhibits the 20S proteasome proteolytic activity, and generated 3D reconstructions of proteasome complexes with PA28-like, PA200-like and a novel, cap-like density. The structures show novel extra density occluding entry into the PA28-like regulator and in the ß chamber of the 20S proteasome, showing other forms of regulatory components that bind the 20S proteasome.

Looking forward, my work has paved the way for these structures to be resolved to higher resolution. There is still a lot of unknowns in the proteasome field especially in the wide interaction network of ancillary proteins. Further studies are necessary, as discussed in Chapters three-five, to fully unravel the intricacies of these interactions. Determining the high resolution structures of UBR4, PI31 with and without the 20S as

well as identifying P97 cofactors that bind to the proteasome will help to fill in gaps in these proteins' roles as ancillary proteins of the proteasome.

Appendix

CRISPR Donor plasmid vector map and sequence



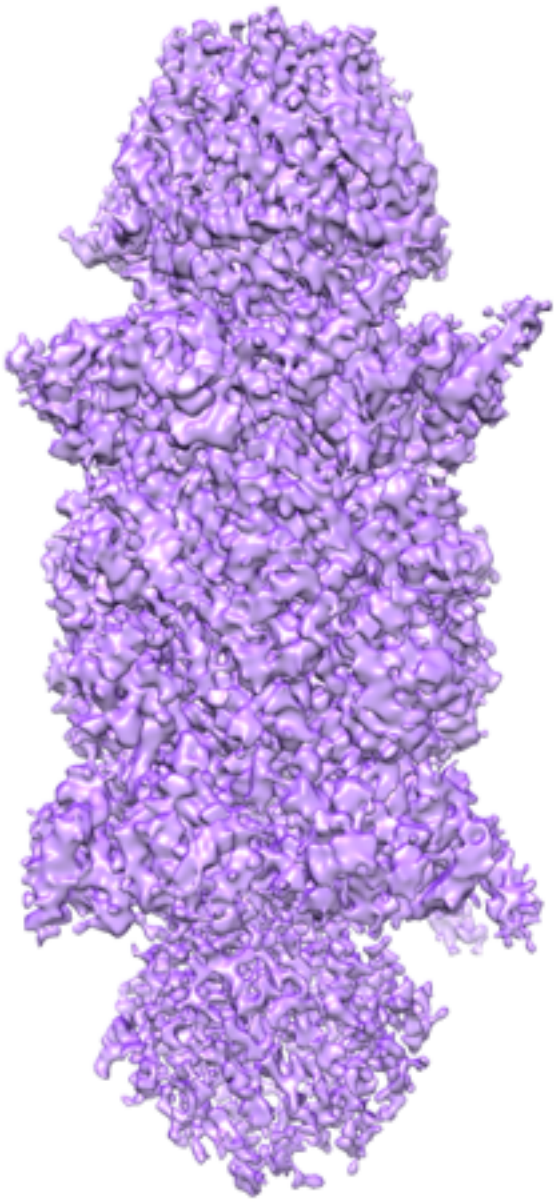
cctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactggaaagcgggcagtgagcgcaacgcaattaatgtgagttagctca
ctcattaggcaccccaggctttacactttatgcttccggctcgtatgttgtgtggaattgtgagcggataacaatttcacacaggaaacagctatgaccatgattac
gaattcatcggcggccgcagcggccatgtttaaacaccggtcgtacccgcggtctagatccCTGTTAGGAGAACGCCTCTGCTAACA
TTTTCTCTATCTTGTTATCCTCTGGGAATGAGACCCACTAAAGGGCTAGAGTGTTGCTCAGTGTG
AATTCCTCTTTCTTCCCAGCTCAGAGCCAGCTGACCGATTTTCCCAAACAGCTCTGTGCCTCCGC
ATTGGCTGCCCCTTGCAGTGTCCTCAGCGGAAGCTCTGTGCTGGGATGGTCTTCGTCGACCTGG
TGCCTGCCTCCGCGTCTGGGCGGGCCTGTGCTAACACTGCTCACCTCTGTGTGTTTGTTAACTCC
ACCACCTCTTTACGGCTCCCACTGGCTCTGTTTCAGCCTAGATGTTTCACCAGCATCATACCTTA
GCACATGCCTCGCGTTCTTCCCAGAAAGTCAGCCCTGGGACCAGGTGGGCTGGTAAAGTCCTCA
GTGCTGGGCCCTGTGCCTGTTGCATGAGCCTGTGTCACTGTGACTGTGGAAGTTGGAGCTAGTG
AGAGGAGGAGGAGGTGGGCAGAGGAAGGAGGGTCCCTTGGGCTCGAGCTCCACATCAGCCCT
GAGCTGCTGGCAAGCACCCCTGGCCTGCAGCCAGGGAGAGCCAGAGCCTTTGCCAGGCAGCGA
GCCTCCGTGCCTTCCACTCCTAGGTCAGGTTTTGACAGCACCACTTTATTACTTTGTTTAAGCTG
TTAGAAAATTTCAGAACATCAGACACACCCTATTCAAATGTGAGAACAAAAACACCTTCCAGC
AGTGCCTGCTTCTGGAACAGGACTCCTCCTGCGGAACAGGGTTTTGCTCAATGACTACCTCCCC
GGTCATGCTGCTGTCCAGTGAGCAGCCGGGCCAGAGCTTGTCGGACCGCAGTGTGTTCTCTGCT

AGATGCGGCAGCTGTGACTCAGGGGGAATGGGGGCTGTATGCCTGCTCCTGACACACCTCTGCC
CGTTGTGCTTTGCAGGTCTTTTATCAGAAATCACCGATCCAGAGAGCTTCCTGAAGGACCTGTT
GAACTCAGTCCCCCTGGAAGTTCTGTTCCAGGGGCCAATGGATTATAAAGACCATGATGGTGAC
TACAAGGATCATGACATTGATTACAAGGACGATGATGACAAATGGTCTCATCCACAGTTCGAA
AAAGGTTCTGCGGGATCAGCTGCCGGATCGGGAGCGGGTTGGAGCCATCCGCAATTCGAGAAA
TAAtatgacggcaataaaaagacagaataaaacgcacgggtgttgggtcgtttgttcataaacgcggggttcggtcccagggctggcactctgtcgatacc
ccaccgagaccccattggggccaatacgcccgcgtttcttccttttccccacccccacccccaagttcgggtgaaggcccagggctcgcagccaacgtcgg
ggcggcaggccctgccatagccactggccccgtgggttagggacggggtcccccatggggaatggtttatggttcgtgggggttattattttgggcgttgcgt
ggggtcaggtccacgacccaagcttggctgcaggtcgatctaccgggtaggggaggcgcttttcccaaggcagtctggagcatgcgctttagcagccccgc
tgggcacttggcgctacacaagtggcctctggcctcgcacacattccacatccaccggtaggcgccaaccggctccgttctttggtggccccttcgcgccacc
ttctactcctcccctagtcaggaagttcccccccgccccgcagctcgcgtcgtgcaggacgtgacaaatggaagtagcacgtctcactagtctcgtgcagatg
gacagcaccgctgagcaatggaagcgggtaggcctttggggcagcggccaatagcagctttgctccttcgctttctgggctcagaggctgggaaggggtgg
gtccgggggcgggctcaggggcgggctcaggggcggggcgggcgcccgaaggtcctccggaggcccggcattctgcacgcttcaaaagcgcacgtct
gccgcgctgttctcctcttcctcatctccgggcctttcgacctgcagcctgttgacaattaatcatcggcatagtatatcggcatagtataatacgacaaggtgag
gaactaaaccATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTG
GACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTAC
GGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCG
TGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACG
ACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGA
CGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGA
GCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTA
CAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAA
GATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCC
CATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGC
AAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATC
ACTCTCGGCATGGACGAGCTGTACAAGGAGAACCTGTACTTCCAGTCGtagctagagctcgctgatcagcctcg
actgtgccttctagttgccagccatctgttgtttgcccctccccgtgccttccttgaccctggaaggtgccactcccactgtcctttcctaataaaatgaggaaatt
gcatcgcattgtctgagtaggtgtcattctattctgggggtggggtggggcaggacagcaaggggaggattgggaagacaatagcaggcatgctgggg
atgcggtgggctctatggcttctgaggcggaaagaaccagctggggctcgactagagcttgcggaacccttaCCACCACACAGCAGCTGC
GGCGGCGAAGACGAAGCTGGCTTGCCTTCCACCCTCTGTTCTCCCTCCTTGTGCATTAAGTTCCC
TCCGCGGGATGCTGCATTGTTACCCCGCCCTCCCCTCTCTCATTTTTCTTGGTGTGGCTTGGGGT
TTTTAGGCTTCCTGTTTTATCTCGTGTGTGTGGTGCACCAGCTATGAGGTTGTCTGTAACCCAAG
CCATCAAAGGGCCTGTACATACCTAGGAGCCATGAGTTGTCCCGGCCAGCTTCATACTTGAGTG
TGCACATCTTGAGAAATAAACAAGTGACTTAACACACATTGAAAAGGAAGGTGTTGCCTGCTG
CTTCAACTCGTCCTCAGTGGCGCAGGCCAAATGGTTTCCCTTGAATGCACAGCCATTTCCGGCG
TCAGCCCGGGAAGACTCATGTCAGACCCGCTGTCTCCAGGCCTGTTGTCAGCGCTAAGTGCAAC
CTGTCTATTCTTGGCCAGCAAGTTGGGGCACCACATTGGATGGCTCTGATTTTTCCCAAAACAA
AGCCCATGCCTTGCTCTTAATTTTTTGGCCTTAAGAGCTGGAACCCAAAGACACGTGGAGCCCA
GAGCTTTCCTAGCCCTCGGCTGAAGGAAGGCTCGCGGAGTAGGGCACAGTCTTGTTGGACCAC
AGAGGCCTGGGGAGCATTTGACATTCTCATGCTGCTGTGAATCGGGAATCCCTTGGTGACTTCT
CTCCAGGGTTGTCATGTTAAGAGAGGGCCTGGGGCAGACTTAGGATCAGGATAAAGAGAGTGG
CTGCTATTGACATTGCCAAGCACCCAGTGTATCTTATTTCTAATCTCTTTGCCCATGAGCCAGAG
CCGCTTGCCCAAGGGCTCCCAGCCAGGGGCTGTGGAGTTGTCAGGTTTGTGCCTCCGAAGCTGT
GTCCTGACCACTGTGGCTGCCGCTCCTGGAGGCCTCAGACCTGCTCCCCTACACACGTGGGCAG
CAGGACTCCCGAGGGCCTCTCtaccggaaggaacccgcgctatgacggcaataaaaagacagaataaaacgcacgggtgttgggtcgt
ttgttcataaacgcggggttcggtcccagggctggcactctgtcgataccccaccgagaccccattggggccaatacgcccgcgtttcttccttttccccaccc
cacccccaagttcgggtgaaggcccagggctcgcagccaacgtcggggcggcaggccctgccatagccactggccccgtgggttagggacggggtcc
cccatggggaatggtttatggttcgtgggggttattattttgggcgttgcgtggggtcaggtccacgacccaagcttggctgcaggtcgacggatcgatccccg
ggaattcactggccgtcgttttacaacgtcgtgactgggaaaaccctggcgttacccaacttaatcgccttgcagcacatcccccctttcgccagctggcgtaata
gcgaagaggcccgcaccgatcgcccttcccaacagttgcgcagcctgaatggcgaatggcgcctgatgcggtattttctccttacgcatctgtgcggtatttca
caccgcatatggtgcactctcagtacaatctgctctgatgccgcatagttaagccagccccgacacccgccaacacccgctgacgcgccctgacgggcttgt
ctgctcccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggttttcaccgtcatcaccgaaacgcgcgagacgaaagggc
ctcgtgatacgcctatttttataggttaatgtcatgataataatggtttcttagacgtcaggtggcacttttcggggaaatgtgcgcggaaccccctatttgtttatttttc
taaatacattcaaatatgtatccgctcatgagacaataaccctgataaatgcttcaataatattgaaaaaggaagagtatgagtattcaacatttccgtgtcgcccctt
attcccttttttgcggcattttgccttcctgtttttgctcacccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcga
actggatctcaacagcggtaagatccttgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtggcgcggtattatcccgta
ttgacgccgggcaagagcaactcggtcgccgcatacactattctcagaatgacttggttgagtactcaccagtcacagaaaagcatcttacggatggcatgac
agtaagagaattatgcagtgctgccataaccatgagtgataacactgcggccaacttacttctgacaacgatcggaggaccgaaggagctaaccgcttttttgc
acaacatgggggatcatgtaactcgccttgatcgttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaatg
gcaacaacgttgcgcaaactattaactggcgaactacttactctagcttcccggcaacaattaatagactggatggaggcggataaagttgcaggaccacttct

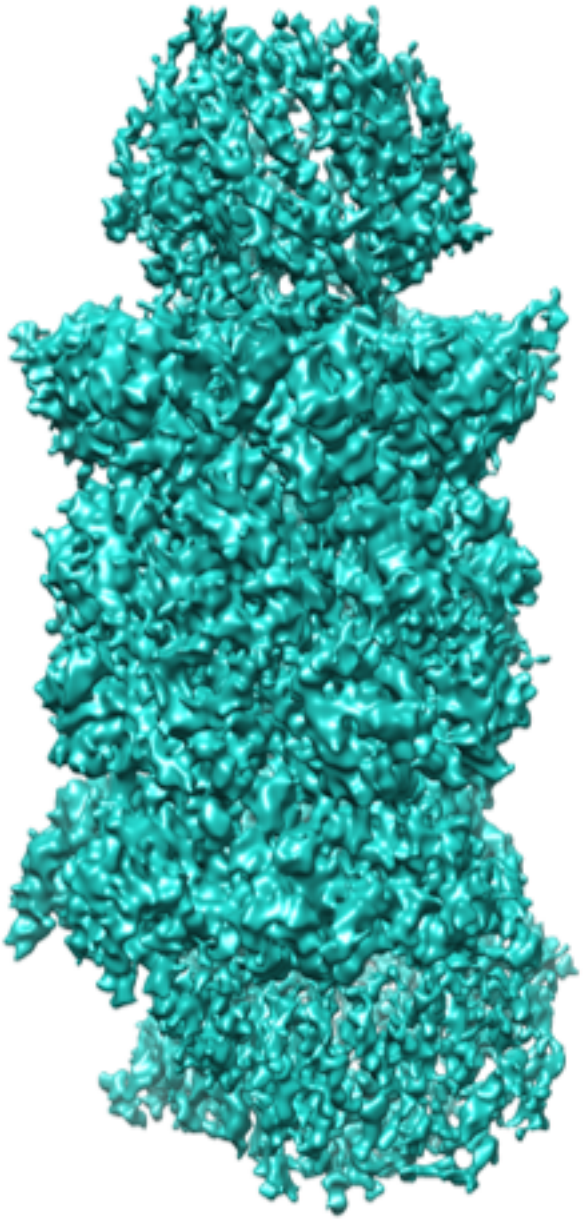
gcgctcggcccttccggctggctggtttattgctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggtaagccc
tcccgtatcgtagttatctacacgacggggagtcaggcaactatggatgaacgaaatagacagatcgctgagataggtgcctcactgattaagcattggtaact
gtcagaccaagtttactcatatatactttagattgatttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttgataatctcatgaccaaaatcccttaa
cgtgagttttcgttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatcctttttttctgcgcgtaatctgctgcttgcaaacaaaaaaacc
accgctaccagcggtggtttgtttgccggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagataccaaatactgtccttctagt
gtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggcgataagtcgtg
tcttaccgggttggactcaagacgatagttaccggataaggcgcagcggtcgggctgaacgggggttcgtgcacacagcccagcttggagcgaacgacc
tacaccgaactgagatacctacagcgtgagctatgagaaagcgccacgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcgga
acaggagagcgcacgagggagcttccagggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctc
gtcaggggggcggagcctatggaaaaacgccagcaacgcggcctttttacggttcctggccttttgctggccttttgctcacatgttctttcctgcgttatcccct
gattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgaccgagcgcagcgagtcagtgagcgaggaagcggaag
agcgcccaatacgcaaaccg

Fig. 5.16 Enlarged maps of the PI31-20S samples

Class one:



Class two:



Class three:

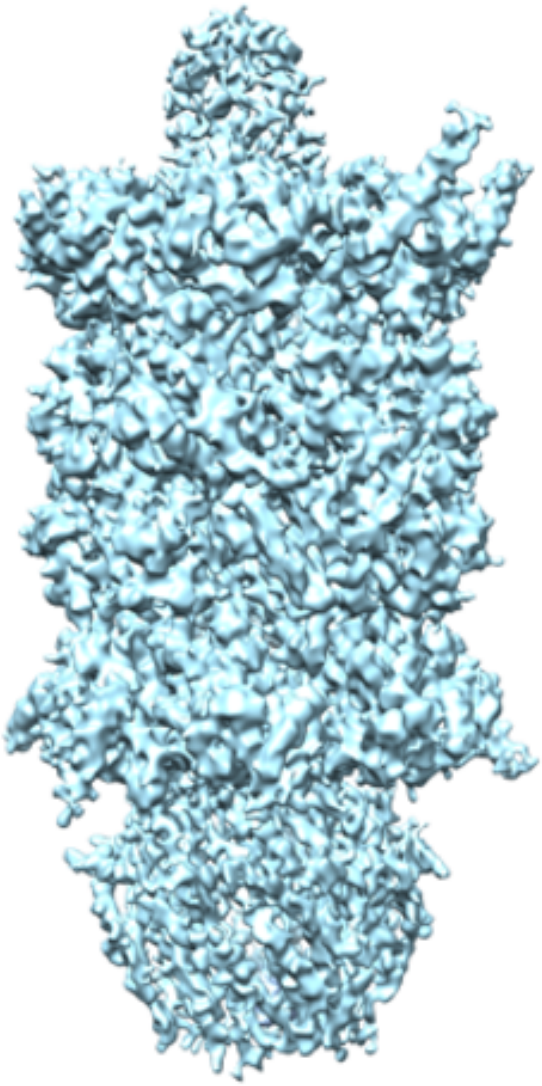
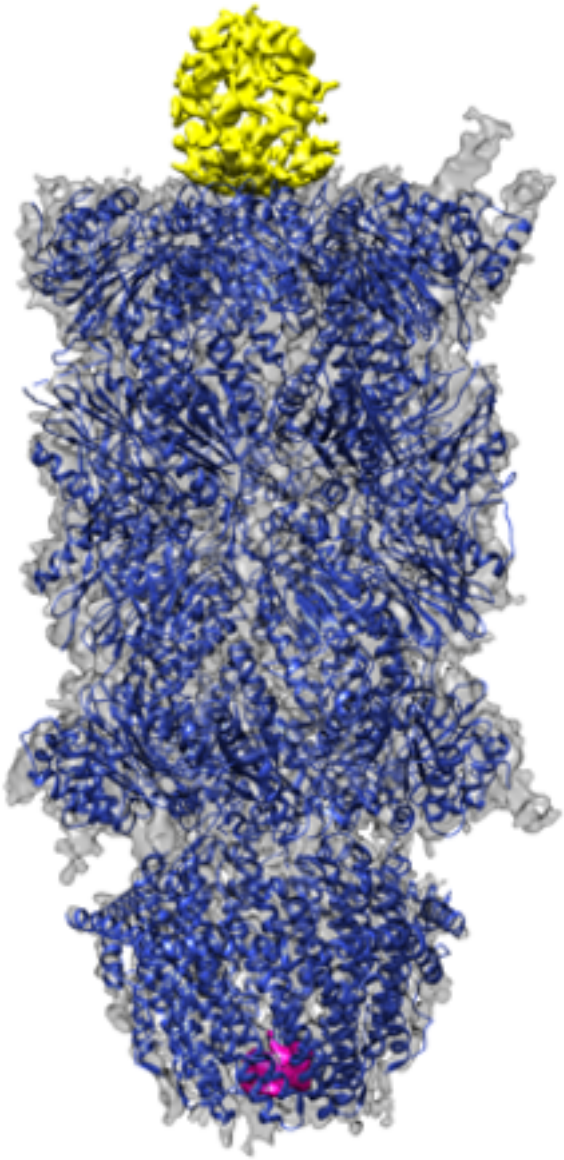


Fig. 5.20 -Englarged map of PI31-20S class three with docked 20S and PA28



Reference List

Adams, J., Palombella, V.J., Sausville, E.A., Johnson, J., Destree, A., Lazarus, D.D., Maas, J., Pien, C.S., Prakash, S., Elliott, P.J., 1999. Proteasome inhibitors: a novel class of potent and effective antitumor agents. Cancer Res. 59, 2615–2622.

Agerberth, B., Lee, J.Y., Bergman, T., Carlquist, M., Boman, H.G., Mutt, V., Jörnvall, H., 1991. Amino acid sequence of PR-39. Isolation from pig intestine of a new member of the family of proline-arginine-rich antibacterial peptides. Eur. J. Biochem. 202, 849–854.

Aguilera, M., Oliveros, M., Martínez-Padrón, M., Barbas, J.A., Ferrús, A., 2000. Ariadne-1: a vital Drosophila gene is required in development and defines a new conserved family of ring-finger proteins. Genetics 155, 1231–1244.

Aguileta, M.A., Korac, J., Durcan, T.M., Trempe, J.-F., Haber, M., Gehring, K., Elsasser, S., Waidmann, O., Fon, E.A., Husnjak, K., 2015. The E3 Ubiquitin Ligase Parkin Is Recruited to the 26 S Proteasome via the Proteasomal Ubiquitin Receptor Rpn13. J. Biol. Chem. 290, 7492–7505.

Aki, M., Shimbara, N., Takashina, M., Akiyama, K., Kagawa, S., Tamura, T., Tanahashi, N., Yoshimura, T., Tanaka, K., Ichihara, A., 1994. Interferon-gamma induces different subunit organizations and functional diversity of proteasomes. J. Biochem. (Tokyo) 115, 257–269.

Akutsu, M., Dikic, I., Bremm, A., 2016. Ubiquitin chain diversity at a glance. J. Cell Sci. 129, 875–880.

Albert, S., Wietrzynski, W., Lee, C.-W., Schaffer, M., Beck, F., Schuller, J.M., Salomé, P.A., Plitzko, J.M., Baumeister, W., Engel, B.D., 2020. Direct visualization of degradation microcompartments at the ER membrane. Proc. Natl. Acad. Sci. 117, 1069–1080.

Arita, M., Wakita, T., Shimizu, H., 2012. Valosin-Containing Protein (VCP/p97) Is Required for Poliovirus Replication and Is Involved in Cellular Protein Secretion Pathway in Poliovirus Infection. J. Virol. 86, 5541–5553.

Arumughan, A., Roske, Y., Barth, C., Forero, L.L., Bravo-Rodriguez, K., Redel, A., Kostova, S., McShane, E., Opitz, R., Faelber, K., Rau, K., Mielke, T., Daumke, O., Selbach, M., Sanchez-Garcia, E., Rocks, O., Panáková, D., Heinemann, U.,

Wanker, E.E., 2016. Quantitative interaction mapping reveals an extended UBX domain in ASPL that disrupts functional p97 hexamers. Nat. Commun. 7, 1–13.

Asher, G., Tsvetkov, P., Kahana, C., Shaul, Y., 2005. A mechanism of ubiquitin-independent proteasomal degradation of the tumor suppressors p53 and p73. Genes Dev. 19, 316–

Ashford, T.P., Porter, K.R., 1962. Cytoplasmic components in hepatic cell lysosomes. J. Cell Biol. 12, 198–202.

Badadani, M., Nalbandian, A., Watts, G.D., Vesa, J., Kitazawa, M., Su, H., Tanaja, J., Dec, E., Wallace, D.C., Mukherjee, J., Caiozzo, V., Warman, M., Kimonis, V.E., 2010. VCP associated inclusion body myopathy and paget disease of bone knock-in mouse model exhibits tissue pathology typical of human disease. PloS One 5: e13183.

Bader, M., Benjamin, S., Wapinski, O.L., Smith, D.M., Goldberg, A.L., Steller, H., 2011. A conserved F-box–regulatory complex controls proteasome activity in Drosophila. Cell 145, 371–382.

Bai, C., Sen, P., Hofmann, K., Ma, L., Goebl, M., Harper, J.W., Elledge, S.J., 1996. SKP1 connects cell cycle regulators to the ubiquitin proteolysis machinery through a novel motif, the F-box. Cell 86, 263–274.

Bai, M., Zhao, X., Sahara, K., Ohte, Y., Hirano, Y., Kaneko, T., Yashiroda, H., Murata, S., 2014. Assembly mechanisms of specialized core particles of the proteasome. Biomolecules 4, 662–677.

Bai, X., McMullan, G., Scheres, S.H.W., 2015. How cryo-EM is revolutionizing structural biology. Trends Biochem. Sci. 40, 49–57.

Bailly, V., Prakash, S., Prakash, L., 1997. Domains required for dimerization of yeast Rad6 ubiquitin-conjugating enzyme and Rad18 DNA binding protein. Mol. Cell. Biol. 17, 4536–4543.

Baker, L.A., Rubinstein, J.L., 2010. Radiation damage in electron cryomicroscopy. Methods Enzymol. 481, 371–388.

Ballatore, C., Lee, V.M.-Y., Trojanowski, J.Q., 2007. Tau-mediated neurodegeneration in Alzheimer's disease and related disorders. Nat. Rev. Neurosci. 8, 663–672.

Banerjee, S., Bartesaghi, A., Merk, A., Rao, P., Bulfer, S.L., Yan, Y., Green, N., Mroczkowski, B., Neitz, R.J., Wipf, P., Falconieri, V., Deshaies, R.J., Milne, J.L.S., Huryn, D., Arkin, M., Subramaniam, S., 2016. 2.3 Å resolution cryo-EM

structure of human p97 and mechanism of allosteric inhibition. Science 351, 871–875.

Barlow, P.N., Luisi, B., Milner, A., Elliott, M., Everett, R., 1994. Structure of the C3HC4 domain by 1H-nuclear magnetic resonance spectroscopy. A new structural class of zinc-finger. J. Mol. Biol. 237, 201–211.

Bartel, B., Wünning, I., Varshavsky, A., 1990. The recognition component of the N-end rule pathway. EMBO J. 9, 3179–3189.

Barthelme, D., Sauer, R.T., 2012. Identification of the Cdc48•20S Proteasome as an Ancient AAA+ Proteolytic Machine 337, 843-846.

Basler, M., Kirk, C.J., Groettrup, M., 2013. The immunoproteasome in antigen processing and other immunological functions. Curr. Opin. Immunol. 25, 74–80.

Baumeister, W., Walz, J., Zühl, F., Seemüller, E., 1998. The Proteasome: Paradigm of a Self-Compartmentalizing Protease. Cell 92, 367–380.

Beck, F., Unverdorben, P., Bohn, S., Schweitzer, A., Pfeifer, G., Sakata, E., Nickell, S., Plitzko, J.M., Villa, E., Baumeister, W., Förster, F., 2012. Near-atomic resolution structural model of the yeast 26S proteasome. Proc. Natl. Acad. Sci. 109, 14870–14875.

Beck, M., Baumeister, W., 2016. Cryo-Electron Tomography: Can it Reveal the Molecular Sociology of Cells in Atomic Detail? Trends Cell Biol. 26, 825–837.

Belzil, C., Asada, N., Ishiguro, K., Nakaya, T., Parsons, K., Pendolino, V., Neumayer, G., Mapelli, M., Nakatani, Y., Sanada, K., Nguyen, M.D., 2014. p600 regulates spindle orientation in apical neural progenitors and contributes to neurogenesis in the developing neocortex. Biol. Open 3, 475–485.

Benvegnu, S., Mateo, MI., Palomer, E., Jurado-Arjona, J., Dotti, C.G., 2017. Aging Triggers cytoplasmic depletion and nuclear translocation of the E3 ligase mahogunin: a function for ubiquitin in neuronal survival. Mol. Cell. 66, 358-372.

Berndsen, C.E., Wolberger, C., 2014. New insights into ubiquitin E3 ligase mechanism. Nat. Struct. Mol. Biol. 21, 301–307.

Bertolaet, B.L., Clarke, D.J., Wolff, M., Watson, M.H., Henze, M., Divita, G., Reed, S.I., 2001. UBA domains of DNA damage-inducible proteins interact with ubiquitin. Nat. Struct. Biol. 8, 417–422.

Besche, H.C., Haas, W., Gygi, S.P., Goldberg, A.L., 2009. Isolation of Mammalian 26S Proteasomes and p97/VCP Complexes Using the Ubiquitin-like Domain from

HHR23B Reveals Novel Proteasome-Associated Proteins. Biochemistry 48, 2538-2549.

Besche, H.C., Sha, Z., Kukushkin, N.V., Peth, A., Hock, E.-M., Kim, W., Gygi, S., Gutierrez, J.A., Liao, H., Dick, L., Goldberg, A.L., 2014. Autoubiquitination of the 26S Proteasome on Rpn13 Regulates Breakdown of Ubiquitin Conjugates. EMBO J. 33, 1159–1176.

Bhattacharyya, S., Yu, H., Mim, C., Matouschek, A., 2014. Regulated protein turnover: snapshots of the proteasome in action. Nat. Rev. Mol. Cell Biol. 15, 122–133.

Bibo-Verdugo, B., Jiang, Z., Caffrey, C.R., O'Donoghue, A.J., 2017. Targeting proteasomes in infectious organisms to combat disease. FEBS J. 284, 1503–1517.

Bieniossek, C., Imasaki, T., Takagi, Y., Berger, I., 2012. MultiBac: expanding the research toolbox for multiprotein complexes. Trends Biochem. Sci. 37, 49–57.

Bilen, J., Bonini, N.M., 2005. Drosophila as a model for human neurodegenerative disease. Annu. Rev. Genet. 39, 153–171.

Blickwedehl, J., Agarwal, M., Seong, C., Pandita, R.K., Melendy, T., Sung, P., Pandita, T.K., Bangia, N., 2008. Role for proteasome activator PA200 and postglutamyl proteasome activity in genomic stability. Proc. Natl. Acad. Sci. U. S. A. 105, 16165–16170.

Bodnar, N., Rapoport, T., 2017. Toward an understanding of the Cdc48/p97 ATPase. F1000Research 6, 1318.

Boes, B., Hengel, H., Ruppert, T., Multhaup, G., Koszinowski, U.H., Kloetzel, P.M., 1994. Interferon gamma stimulation modulates the proteolytic activity and cleavage site preference of 20S mouse proteasomes. J. Exp. Med. 179, 901–909.

Bogyo, M., Wang, E.W., 2002. Proteasome inhibitors: complex tools for a complex enzyme. Curr. Top. Microbiol. Immunol. 268, 185–208.

Boise, L.H., Kaufman, J.L., Bahlis, N.J., Lonial, S., Lee, K.P., 2014. The Tao of myeloma. Blood 124, 1873–1879.

Booth, D.S., Avila-Sakar, A., Cheng, Y., 2011. Visualizing proteins and macromolecular complexes by negative stain EM: from grid preparation to image acquisition. J. Vis. Exp. JoVE.

Bordallo, J., Plemper, R.K., Finger, A., Wolf, D.H., 1998. Der3p/Hrd1p is required for endoplasmic reticulum-associated degradation of misfolded lumenal and integral membrane proteins. Mol. Biol. Cell 9, 209–222.

Borden, K.L., Boddy, M.N., Lally, J., O'Reilly, N.J., Martin, S., Howe, K., Solomon, E., Freemont, P.S., 1995. The solution structure of the RING finger domain from the acute promyelocytic leukaemia proto-oncoprotein PML. EMBO J. 14, 1532–1541.

Botelho-Machado, C., Cabral, F.J., Soares, C.S., Moreira, E.B.C., Morais, E.R., Magalhães, L.G., Gomes, M.S., Guerra-Sá, R., Rosa, J.C., Ruller, R., Ward, R.J., Rodrigues, V., 2010. Characterization and mRNA expression analysis of PI31, an endogenous proteasome inhibitor from Schistosoma mansoni. Parasitol. Res. 107, 1163–1171.

Bowerman, B., Kurz, T., 2006. Degrade to create: developmental requirements for ubiquitin-mediated proteolysis during early C. elegans embryogenesis. Dev. Camb. Engl. 133, 773–784.

Bozzola, J.J., Russell, L.D., 1999. Electron Microscopy: Principles and Techniques for Biologists. Jones & Bartlett Learning.

Brodsky, J.L., 2012. Cleaning Up: ER-Associated Degradation to the Rescue. Cell 151, 1163–1167.

Brundin, P., Melki, R., Kopito, R., 2010. Prion-like transmission of protein aggregates in neurodegenerative diseases. Nat. Rev. Mol. Cell Biol. 11, 301–307.

Brzovic, P.S., Keeffe, J.R., Nishikawa, H., Miyamoto, K., Fox, D., Fukuda, M., Ohta, T., Klevit, R., 2003. Binding and recognition in the assembly of an active BRCA1/BARD1 ubiquitin-ligase complex. Proc. Natl. Acad. Sci. U. S. A. 100, 5646–5651.

Buchberger, A., Howard, M.J., Proctor, M., Bycroft, M., 2001. The UBX domain: a widespread ubiquitin-like module. J. Mol. Biol. 307, 17–24.

Buchberger, A., Schindelin, H., Hänzelmann, P., 2015. Control of p97 function by cofactor binding. FEBS Lett. 589, 2578–2589.

Cagney, G., Uetz, P., Fields, S., 2001. Two-hybrid analysis of the Saccharomyces cerevisiae 26S proteasome. Physiol. Genomics 7, 27–34.

Callaghan, M.J., Russell, A.J., Woollatt, E., Sutherland, G.R., Sutherland, R.L., Watts, C.K., 1998. Identification of a human HECT family protein with homology to the Drosophila tumor suppressor gene hyperplastic discs. Oncogene 17, 3479–3491.

Cascio, P., Goldberg, A.L., 2005. Preparation of hybrid (19S-20S-PA28) proteasome complexes and analysis of peptides generated during protein degradation. Methods Enzymol. 398, 336–352.

Cavo, M., 2006. Proteasome inhibitor bortezomib for the treatment of multiple myeloma. Leukemia 20, 1341–1352.

Cavo, M., Tosi, P., Zamagni, E., Cellini, C., Tacchetti, P., Patriarca, F., Di Raimondo, F., Volpe, E., Ronconi, S., Cangini, D., Narni, F., Carubelli, A., Masini, L., Catalano, L., Fiacchini, M., de Vivo, A., Gozzetti, A., Lazzaro, A., Tura, S., Baccarani, M., 2007. Prospective, Randomized Study of Single Compared With Double Autologous Stem-Cell Transplantation for Multiple Myeloma: Bologna 96 Clinical Study. J. Clin. Oncol. 25, 2434–2441.

Chauhan, D., Bianchi, G., Anderson, K.C., 2008. Targeting the UPS as therapy in multiple myeloma. BMC Biochem. 9, S1.

Chauhan, D., Catley, L., Li, G., Podar, K., Hideshima, T., Velankar, M., Mitsiades, C., Mitsiades, N., Yasui, H., Letai, A., Ovaa, H., Berkers, C., Nicholson, B., Chao, T.-H., Neuteboom, S.T.C., Richardson, P., Palladino, M.A., Anderson, K.C., 2005. A novel orally active proteasome inhibitor induces apoptosis in multiple myeloma cells with mechanisms distinct from Bortezomib. Cancer Cell 8, 407–419.

Chauhan, D., Hideshima, T., Anderson, K.C., 2006. A novel proteasome inhibitor NPI-0052 as an anticancer therapy. Br. J. Cancer 95, 961–965.

Chen, B., Kaledhonkar, S., Sun, M., Shen, B., Lu, Z., Barnard, D., Lu, T.-M., Gonzalez, R.L., Frank, J., 2015. Structural dynamics of ribosome subunit association studied by mixing-spraying time-resolved cryogenic electron microscopy. Struct. Lond. Engl. 1993 23, 1097–1105.

Chen, L., Shinde, U., Ortolan, T.G., Madura, K., 2001. Ubiquitin-associated (UBA) domains in Rad23 bind ubiquitin and promote inhibition of multi-ubiquitin chain assembly. EMBO Rep. 2, 933–938.

Chen, X., Randles, L., Shi, K., Tarasov, S.G., Aihara, H., Walters, K.J., 2016. Structures of Rpn1 T1:Rad23 and hRpn13:hPLIC2 Reveal Distinct Binding Mechanisms between Substrate Receptors and Shuttle Factors of the Proteasome. Struct. Lond. Engl. 1993 24, 1257–1270.

Chin, D.T., Kuehl, L., Rechsteiner, M., 1982. Conjugation of ubiquitin to denatured hemoglobin is proportional to the rate of hemoglobin degradation in HeLa cells. Proc. Natl. Acad. Sci. U. S. A. 79, 5857–5861.

Cho-Park, P.F., Steller, H., 2013. Proteasome regulation by ADP-ribosylation. Cell 153, 614–627.

Chu-Ping, M., Slaughter, C.A., DeMartino, G.N., 1992. Purification and characterization of a protein inhibitor of the 20S proteasome (macropain). Biochim. Biophys. Acta 1119, 303–311.

Ciechanover, A., 2005. Proteolysis: from the lysosome to ubiquitin and the proteasome. Nat. Rev. Mol. Cell Biol. 6, 79–87.

Ciechanover, A., Elias, S., Heller, H., Hershko, A., 1982. "Covalent affinity" purification of ubiquitin-activating enzyme. J. Biol. Chem. 257, 2537–2542.

Ciechanover, A., Heller, H., Elias, S., Haas, A.L., Hershko, A., 1980. ATP-dependent conjugation of reticulocyte proteins with the polypeptide required for protein degradation. Proc. Natl. Acad. Sci. 77, 1365–1368.

Ciechanover, A., Heller, H., Katz-Etzion, R., Hershko, A., 1981. Activation of the heat-stable polypeptide of the ATP-dependent proteolytic system. Proc. Natl. Acad. Sci. U. S. A. 78, 761–765.

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

Clark, S.L., 1957. Cellular differentiation in the kidneys of newborn mice studied with the electron microscope. J. Biophys. Biochem. Cytol. 3, 349–362.

Collins, G.A., Goldberg, A.L., 2017. The Logic of the 26S Proteasome. Cell 169, 792–806.

Conedera, S., Apaydin, H., Li, Y., Yoshino, H., Ikeda, A., Matsushima, T., Funayama, M., Nishioka, K., Hattori, N., 2016. FBXO7 mutations in Parkinson's disease and multiple system atrophy. Neurobiol. Aging 40, 192.e1-192.e5.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., Zhang, F., 2013. Multiplex genome engineering using CRISPR/Cas systems. Science 339, 819–823.

Cope, G.A., Deshaies, R.J., 2003. COP9 signalosome: a multifunctional regulator of SCF and other cullin-based ubiquitin ligases. Cell 114, 663–671.

Crawford, L.J.A., Walker, B., Ovaa, H., Chauhan, D., Anderson, K.C., Morris, T.C.M., Irvine, A.E., 2006. Comparative selectivity and specificity of the proteasome inhibitors BzLLLCOCHO, PS-341, and MG-132. Cancer Res. 66, 6379–6386.

Crosas, B., Hanna, J., Kirkpatrick, D.S., Zhang, D.P., Tone, Y., Hathaway, N.A., Buecker, C., Leggett, D.S., Schmidt, M., King, R.W., Gygi, S.P., Finley, D., 2006. Ubiquitin

chains are remodeled at the proteasome by opposing ubiquitin ligase and deubiquitinating activities. Cell 127, 1401–1413.

Crowther, R.A., Huxley, H.E., Klug, A., 1971. Procedures for three-dimensional reconstruction of spherical viruses by Fourier synthesis from electron micrographs. Philos. Trans. R. Soc. Lond. B Biol. Sci. 261, 221–230.

Cusack, J.C., 2003. Rationale for the treatment of solid tumors with the proteasome inhibitor bortezomib. Cancer Treat. Rev. 29 Suppl 1, 21–31.

da Fonseca, P.C.A., He, J., Morris, E.P., 2012. Molecular model of the human 26S proteasome. Mol. Cell 46, 54–66.

Dächsel, J.C., Lücking, C.B., Deeg, S., Schultz, E., Lalowski, M., Casademunt, E., Corti, O., Hampe, C., Patenge, N., Vaupel, K., Yamamoto, A., Dichgans, M., Brice, A., Wanker, E.E., Kahle, P.J., Gasser, T., 2005. Parkin interacts with the proteasome subunit α4. FEBS Lett. 579, 3913–3919.

Dahlmann, B., Kuehn, L., Reinauer, H., 1995. Studies on the activation by ATP of the 26S proteasome complex from rat skeletal muscle. The Biochemical Journal 209: 195-202.

Dai, R.M., Chen, E., Longo, D.L., Gorbea, C.M., Li, C.C., 1998. Involvement of valosin-containing protein, an ATPase Co-purified with IkappaBalpha and 26 S proteasome, in ubiquitin-proteasome-mediated degradation of IkappaBalpha. J. Biol. Chem. 273, 3562–3573.

David, Y., Ternette, N., Edelmann, M.J., Ziv, T., Gayer, B., Sertchook, R., Dadon, Y., Kessler, B.M., Navon, A., 2011. E3 ligases determine ubiquitination site and conjugate type by enforcing specificity on E2 enzymes. J. Biol. Chem. 286, 44104–44115.

Davies, J.M., Brunger, A.T., Weis, W.I., 2008. Improved structures of full-length p97, an AAA ATPase: implications for mechanisms of nucleotide-dependent conformational change. Struct. Lond. Engl. 1993 16, 715–726.

Davy, A., Bello, P., Thierry-Mieg, N., Vaglio, P., Hitti, J., Doucette-Stamm, L., Thierry-Mieg, D., Reboul, J., Boulton, S., Walhout, A.J.M., Coux, O., Vidal, M., 2001. A protein–protein interaction map of the Caenorhabditis elegans 26S proteasome. EMBO Rep. 2, 821–828.

de Duve, C., Pressman, B.C., Gianetto, R., Wattiaux, R., Appelmans, F., 1955. Tissue fractionation studies. 6. Intracellular distribution patterns of enzymes in rat-liver tissue. Biochem. J. 60, 604–617.

de Duve, C., Wattiaux, R., 1966. Functions of Lysosomes. Annu. Rev. Physiol. 28, 435–492.

de la Peña, A.H., Goodall, E.A., Gates, S.N., Lander, G.C., Martin, A., 2018. Substrate-engaged 26S proteasome structures reveal mechanisms for ATP-hydrolysis-driven translocation. Science 362, 725.

De Rosier, D.J., Klug, A., 1968. Reconstruction of Three Dimensional Structures from Electron Micrographs. Nature 217, 130–134.

DeMasi, J., Huh, K.-W., Nakatani, Y., Münger, K., Howley, P.M., 2005. Bovine papillomavirus E7 transformation function correlates with cellular p600 protein binding. Proc. Natl. Acad. Sci. 102, 11486–11491.

DeRosier, D.J., Moore, P.B., 1970. Reconstruction of three-dimensional images from electron micrographs of structures with helical symmetry. J. Mol. Biol. 52, 355–369.

Deshaies, R.J., 2014. Proteotoxic crisis, the ubiquitin-proteasome system, and cancer therapy. BMC Biol. 12, 94.

Deshaies, R.J., Joazeiro, C.A.P., 2009. RING Domain E3 Ubiquitin Ligases. Annu. Rev. Biochem. 78, 399–434.

Deter, R.L., de Duve, C., 1967. Influence of Glucagon, an inducer of cellular autophagy, on some physical properties of rat liver lysosomes. J. Cell Biol. 33, 437–449.

Di Fonzo, A., Dekker, M., Montagna, P., Baruzzi, A., Yonova-Doing, E., Guedes, L., Szczerbinska, A., Zhao, T., Dubbel-Hulsman, L., Wouters, C., de Graaff, E., Oyen, W., Simons, E., Breedveld, G.J., Oostra, B.A., Horstink, M.W., Bonifati, V., 2008. FBXO7 mutations cause autosomal recessive, early-onset parkinsonian-pyramidal syndrome. Neurology 72, 240–245.

Díaz-Hernández, M., Martín-Aparicio, E., Avila, J., Hernández, F., Lucas, J.J., 2004. Enhanced induction of the immunoproteasome by interferon gamma in neurons expressing mutant Huntingtin. Neurotox. Res. 6, 463–468.

Dick, T.P., Nussbaum, A.K., Deeg, M., Heinemeyer, W., Groll, M., Schirle, M., Keilholz, W., Stevanović, S., Wolf, D.H., Huber, R., Rammensee, H.G., Schild, H., 1998. Contribution of proteasomal beta-subunits to the cleavage of peptide substrates analyzed with yeast mutants. J. Biol. Chem. 273, 25637–25646.

Dominguez, C., Bonvin, A.M.J.J., Winkler, G.S., van Schaik, F.M.A., Timmers, H.T.M., Boelens, R., 2004. Structural model of the UbcH5B/CNOT4 complex revealed by

combining NMR, mutagenesis, and docking approaches. Struct. Lond. Engl. 1993 12, 633–644.

Driscoll, J., Brown, M.G., Finley, D., Monaco, J.J., 1993. MHC-linked LMP gene products specifically alter peptidase activities of the proteasome. Nature 365, 262–264.

Driscoll, J., Goldberg, A.L., 1990. The proteasome (multicatalytic protease) is a component of the 1500-kDa proteolytic complex which degrades ubiquitin-conjugated proteins. J. Biol. Chem. 265, 4789–4792.

Dubochet, J., Lepault, J., Freeman, R., Berriman, J.A., Homo, J.-C., 1982. Electron microscopy of frozen water and aqueous solutions. J. Microsc. 128, 219–237.

Duda, D.M., Olszewski, J.L., Schuermann, J.P., Kurinov, I., Miller, D.J., Nourse, A., Alpi, A.F., Schulman, B.A., 2013. Structure of HHARI, a RING-IBR-RING ubiquitin ligase: autoinhibition of an Ariadne-family E3 and insights into ligation mechanism. Struct. Lond. Engl. 1993 21, 1030–1041.

Egelman, E.H., 2016. The Current Revolution in Cryo-EM. Biophys. J. 110, 1008–1012.

Ehlinger, A., Walters, K.J., 2013. Structural insights into proteasome activation by the 19S regulatory particle. Biochemistry 52, 3618–3628.

Eisele, M.R., Reed, R.G., Rudack, T., Schweitzer, A., Beck, F., Nagy, I., Pfeifer, G., Plitzko, J.M., Baumeister, W., Tomko, R.J., Sakata, E., 2018. Expanded Coverage of the 26S Proteasome Conformational Landscape Reveals Mechanisms of Peptidase Gating. Cell Rep. 24, 1301-1315.

Eisenhaber, B., Chumak, N., Eisenhaber, F., Hauser, M.-T., 2007. The ring between ring fingers (RBR) protein family. Genome Biol. 8, 209.

Erales, J., Coffino, P., 2012. Ubiquitin-independent proteasomal degradation. Biochimica et Biophysica Acta-Molecular Cell Research 1843, 216-221.

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

Etlinger, J.D., Goldberg, A.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.

Eytan, E., Ganoth, D., Armon, T., Hershko, A., 1989. ATP-dependent incorporation of 20S protease into the 26S complex that degrades proteins conjugated to ubiquitin. Proc. Natl. Acad. Sci. U. S. A. 86, 7751–7755.

Faruqi, A.R., McMullan, G., 2011. Electronic detectors for electron microscopy. Q. Rev. Biophys. 44, 357–390.

Fehlker, M., Wendler, P., Lehmann, A., Enenkel, C., 2003. Blm3 is part of nascent proteasomes and is involved in a late stage of nuclear proteasome assembly. EMBO Rep. 4, 959–963.

Fernandez-Moran, H., 1960. Low-temperature preparation techniques for electron microscopy of biological specimens based on rapid freezing with liquid helium II. Ann. N. Y. Acad. Sci. 85, 689–713.

Fernández-Sáiz, V., Buchberger, A., 2010. Imbalances in p97 co-factor interactions in human proteinopathy. EMBO Rep. 11, 479–485.

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

Finley, D., Varshavsky, A., 1985. The ubiquitin system: functions and mechanisms. Trends Biochem. Sci. 10, 343–347.

Fitzgerald, D.J., Berger, P., Schaffitzel, C., Yamada, K., Richmond, T.J., Berger, I., 2006. Protein complex expression by using multigene baculoviral vectors. Nat. Methods 3, 1021–1032.

Fleig, L., Bergbold, N., Sahasrabudhe, P., Geiger, B., Kaltak, L., Lemberg, M.K., 2012. Ubiquitin-dependent intramembrane rhomboid protease promotes ERAD of membrane proteins. Mol. Cell 47, 558–569.

Förster, A., Masters, E.I., Whitby, F.G., Robinson, H., Hill, C.P., 2005. The 1.9 A structure of a proteasome-11S activator complex and implications for proteasome-PAN/PA700 interactions. Mol. Cell 18, 589–599.

Förster, A., Whitby, F.G., Hill, C.P., 2003. The pore of activated 20S proteasomes has an ordered 7-fold symmetric conformation. EMBO J. 22, 4356–4364.

Fort, P., Kajava, A.V., Delsuc, F., Coux, O., 2015. Evolution of proteasome regulators in eukaryotes. Genome Biol. Evol. 7, 1363–1379.

Franke, N.E., Niewerth, D., Assaraf, Y.G., van Meerloo, J., Vojtekova, K., van Zantwijk, C.H., Zweegman, S., Chan, E.T., Kirk, C.J., Geerke, D.P., Schimmer, A.D., Kaspers, G.J.L., Jansen, G., Cloos, J., 2012. Impaired bortezomib binding to mutant β5 subunit of the proteasome is the underlying basis for bortezomib resistance in leukemia cells. Leukemia 26, 757–768.

Freemont, P.S., Hanson, I.M., Trowsdale, J., 1991. A novel cysteine-rich sequence motif. Cell 64, 483–484.

Frentzel, S., Pesold-Hurt, B., Seelig, A., Kloetzel, P.-M., 1994. 20 S proteasomes are assembled via distinct precursor complexes processing of LMP2 and LMP7 proproteins takes place in 13–16 S preproteasome complexes. J. Mol. Biol. 236, 975–981.

Fujita, K., Nakamura, Y., Oka, T., Ito, H., Tamura, T., Tagawa, K., Sasabe, T., Katsuta, A., Motoki, K., Shiwaku, H., Sone, M., Yoshida, C., Katsuno, M., Eishi, Y., Murata, M., Taylor, J.P., Wanker, E.E., Kono, K., Tashiro, S., Sobue, G., La Spada, A.R., Okazawa, H., 2013. A functional deficiency of TERA/VCP/p97 contributes to impaired DNA repair in multiple polyglutamine diseases. Nat. Commun. 4, 1816.

Furniss, J.J., Grey, H., Wang, Z., Nomoto, M., Jackson, L., Tada, Y., Spoel, S.H., 2018. Proteasome-associated HECT-type ubiquitin ligase activity is required for plant immunity. PLOS Pathog. 14, e1007447.

Gaczynska, M., Osmulski, P.A., Gao, Y., Post, M.J., Simons, M., 2003. Proline- and arginine-rich peptides constitute a novel class of allosteric inhibitors of proteasome activity. Biochemistry 42, 8663–8670.

Gaczynska, M., Rock, K.L., Goldberg, A.L., 1993. Gamma-interferon and expression of MHC genes regulate peptide hydrolysis by proteasomes. Nature 365, 264–267.

Gandolfi, S., Laubach, J.P., Hideshima, T., Chauhan, D., Anderson, K.C., Richardson, P.G., 2017. The proteasome and proteasome inhibitors in multiple myeloma. Cancer Metastasis Rev. 36, 561–584.

Ganoth, D., Leshinsky, E., Eytan, E., Hershko, A., 1988. A multicomponent system that degrades proteins conjugated to ubiquitin. Resolution of factors and evidence for ATP-dependent complex formation. J. Biol. Chem. 263, 12412–12419.

García-Nafría, J., Watson, J.F., Greger, I.H., 2016. IVA cloning: A single-tube universal cloning system exploiting bacterial In Vivo Assembly. Sci. Rep. 6, 1–12.

Garman, E.F., Weik, M., 2017. X-ray radiation damage to biological macromolecules: further insights. J. Synchrotron Radiat. 24, 1–6.

Gavin, A.-C., Bösche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J.M., Michon, A.-M., Cruciat, C.-M., Remor, M., Höfert, C., Schelder, M., Brajenovic, M., Ruffner, H., Merino, A., Klein, K., Hudak, M., Dickson, D., Rudi, T., Gnau, V., Bauch, A., Bastuck, S., Huhse, B., Leutwein, C., Heurtier, M.-A., Copley, R.R., Edelmann, A., Querfurth, E., Rybin, V., Drewes, G., Raida, M., Bouwmeester, T., Bork, P., Seraphin, B., Kuster, B., Neubauer, G., Superti-Furga,

G., 2002. Functional organization of the yeast proteome by systematic analysis of protein complexes. Nature 415, 141–147.

Glickman, M.H., Raveh, D., 2005. Proteasome plasticity. FEBS Lett. 579, 3214–3223.

Glickman, M.H., Rubin, D.M., Coux, O., Wefes, I., Pfeifer, G., Cjeka, Z., Baumeister, W., Fried, V.A., 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, 615–623.

Goldberg, A.L., 2007. Functions of the proteasome: from protein degradation and immune surveillance to cancer therapy. Biochem. Soc. Trans. 35, 12–17.

Goldenberg, S.J., Cascio, T.C., Shumway, S.D., Garbutt, K.C., Liu, J., Xiong, Y., Zheng, N., 2004. Structure of the Cand1-Cul1-Roc1 complex reveals regulatory mechanisms for the assembly of the multisubunit cullin-dependent ubiquitin ligases. Cell 119, 517–528.

Goldknopf, I.L., Busch, H., 1977. Isopeptide linkage between nonhistone and histone 2A polypeptides of chromosomal conjugate-protein A24. Proc. Natl. Acad. Sci. U. S. A. 74, 864–868.

Goldstein, G., Scheid, M., Hammerling, U., Schlesinger, D.H., Niall, H.D., Boyse, E.A., 1975. Isolation of a polypeptide that has lymphocyte-differentiating properties and is probably represented universally in living cells. Proc. Natl. Acad. Sci. U. S. A. 72, 11–15.

Gonzalez, M.A., Feely, S.M., Speziani, F., Strickland, A.V., Danzi, M., Bacon, C., Lee, Y., Chou, T.-F., Blanton, S.H., Weihl, C.C., Zuchner, S., Shy, M.E., 2014. A novel mutation in VCP causes Charcot-Marie-Tooth Type 2 disease. Brain J. Neurol. 137, 2897–2902.

Gorbea, C., Goellner, G.M., Teter, K., Holmes, R.K., Rechsteiner, M., 2004. Characterization of Mammalian Ecm29, a 26 S Proteasome-associated Protein That Localizes to the Nucleus and Membrane Vesicles. J. Biol. Chem. 279, 54849–54861.

Gorbea, C., Pratt, G., Ustrell, V., Bell, R., Sahasrabudhe, S., Hughes, R.E., Rechsteiner, M., 2010. A Protein Interaction Network for Ecm29 Links the 26 S Proteasome to Molecular Motors and Endosomal Components. J. Biol. Chem. 285, 31616–31633.

Grant, T., Rohou, A., Grigorieff, N., 2018. cisTEM, user-friendly software for single-particle image processing. eLife 7, e35383.

Grice, G.L., Nathan, J., 2016. The recognition of ubiquitinated proteins by the proteasome.

Grigorieff, N., 2007. FREALIGN: high-resolution refinement of single particle structures. J. Struct. Biol. 157, 117–125.

Groettrup, M., Kirk, C.J., Basler, M., 2010. Proteasomes in immune cells: more than peptide producers? Nat. Rev. Immunol. 10, 73–78.

Groll, M., Bajorek, M., Köhler, A., Moroder, L., Rubin, D.M., Huber, R., Glickman, M.H., Finley, D., 2000. A gated channel into the proteasome core particle. Nat. Struct. Biol. 7, 1062–1067.

Groll, M., Ditzel, L., Löwe, J., Stock, D., Bochtler, M., Bartunik, H.D., Huber, R., 1997. Structure of 20S proteasome from yeast at 2.4 A resolution. Nature 386, 463–471.

Groll, M., Heinemeyer, W., Jäger, S., Ullrich, T., Bochtler, M., Wolf, D.H., Huber, R., 1999. The catalytic sites of 20S proteasomes and their role in subunit maturation: A mutational and crystallographic study. Proc. Natl. Acad. Sci. 96, 10976–10983.

Guerriero, C.J., Brodsky, J.L., 2012. The delicate balance between secreted protein folding and endoplasmic reticulum-associated degradation in human physiology. Physiol. Rev. 92, 537–576.

Haas, A.L., Warms, J.V., Hershko, A., Rose, I.A., 1982. Ubiquitin-activating enzyme. Mechanism and role in protein-ubiquitin conjugation. J. Biol. Chem. 257, 2543–2548.

Hamazaki, J., Iemura, S.-I., Natsume, T., Yashiroda, H., Tanaka, K., Murata, S., 2006. A novel proteasome interacting protein recruits the deubiquitinating enzyme UCH37 to 26S proteasomes. EMBO J. 25, 4524–4536.

Hanson, P.I., Whiteheart, S.W., 2005. AAA+ proteins: have engine, will work. Nat. Rev. Mol. Cell Biol. 6, 519–529.

Hänzelmann, P., Buchberger, A., Schindelin, H., 2011. Hierarchical binding of cofactors to the AAA ATPase p97. Struct. Lond. Engl. 1993 19, 833–843.

Hänzelmann, P., Schindelin, H., 2017. The Interplay of Cofactor Interactions and Post-translational Modifications in the Regulation of the AAA+ ATPase p97. Front. Mol. Biosci. 4, 21.

Harauz, G., Heel, M.G. van, 1986. Exact filters for general geometry three dimensional reconstruction. Optik, 73, 146-156.

Harris, J.R., Scheffler, D., 2002. Routine preparation of air-dried negatively stained and unstained specimens on holey carbon support films: a review of applications. Micron Oxf. Engl. 1993 33, 461–480.

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

Haselbach, D., Schrader, J., Lambrecht, F., Henneberg, F., Chari, A., Stark, H., 2017. Long-range allosteric regulation of the human 26S proteasome by 20S proteasome-targeting cancer drugs. Nat. Commun. 8, 1–8.

Henderson, R., Unwin, P.N.T., 1975. Three-dimensional model of purple membrane obtained by electron microscopy. Nature 257, 28–32.

Hershko, A., Ciechanover, A., Heller, H., Haas, A.L., Rose, I.A., 1980. Proposed role of ATP in protein breakdown: conjugation of protein with multiple chains of the polypeptide of ATP-dependent proteolysis. Proc. Natl. Acad. Sci. 77, 1783–1786.

Hershko, A., Eytan, E., Ciechanover, A., Haas, A.L., 1982. Immunochemical analysis of the turnover of ubiquitin-protein conjugates in intact cells. Relationship to the breakdown of abnormal proteins. J. Biol. Chem. 257, 13964–13970.

Hershko, A., Heller, H., Elias, S., Ciechanover, A., 1983. Components of ubiquitin-protein ligase system. Resolution, affinity purification, and role in protein breakdown. J. Biol. Chem. 258, 8206–8214.

Hershko, A., Heller, H., Ganoth, D., Ciehanover, A., 1978. Mode of degradation of abnormal globin chains in rabbit reticulocytes. Protein Turnover and Lysosome Function. Academic Press, pp. 149–169.

Hershko, A., Leshinsky, E., Ganoth, D., Heller, H., 1984. ATP-dependent degradation of ubiquitin-protein conjugates. Proc. Natl. Acad. Sci. 81, 1619–1623.

Hideshima, T., Richardson, P., Chauhan, D., Palombella, V.J., Elliott, P.J., Adams, J., Anderson, K.C., 2001. The proteasome inhibitor PS-341 inhibits growth, induces apoptosis, and overcomes drug resistance in human multiple myeloma cells. Cancer Res. 61, 3071–3076.

Hipp, M.S., Park, S.-H., Hartl, F.U., 2014. Proteostasis impairment in protein-misfolding and -aggregation diseases. Trends Cell Biol. 24, 506–514.

Hipp, M.S., Patel, C.N., Bersuker, K., Riley, B.E., Kaiser, S.E., Shaler, T.A., Brandeis, M., Kopito, R.R., 2012. Indirect inhibition of 26S proteasome activity in a cellular model of Huntington's disease. J. Cell Biol. 196, 573–587.

Hirano, Y., Hayashi, H., Iemura, S.-I., Hendil, K.B., Niwa, S.-I., Kishimoto, T., Kasahara, M., Natsume, T., Tanaka, K., Murata, S., 2006. Cooperation of multiple chaperones required for the assembly of mammalian 20S proteasomes. Mol. Cell 24, 977–984.

Hirano, Y., Hendil, K.B., Yashiroda, H., Iemura, S., Nagane, R., Hioki, Y., Natsume, T., Tanaka, K., Murata, S., 2005. A heterodimeric complex that promotes the assembly of mammalian 20S proteasomes. Nature 437, 1381–1385.

Hjerpe, R., Aillet, F., Lopitz-Otsoa, F., Lang, V., England, P., Rodriguez, M.S., 2009. Efficient protection and isolation of ubiquitylated proteins using tandem ubiquitin-binding entities. EMBO Rep. 10, 1250–1258.

Ho, Y., Gruhler, A., Heilbut, A., Bader, G.D., Moore, L., Adams, S.-L., Millar, A., Taylor, P., Bennett, K., Boutilier, K., Yang, L., Wolting, C., Donaldson, I., Schandorff, S., Shewnarane, J., Vo, M., Taggart, J., Goudreault, M., Muskat, B., Alfarano, C., Dewar, D., Lin, Z., Michalickova, K., Willems, A.R., Sassi, H., Nielsen, P.A., Rasmussen, K.J., Andersen, J.R., Johansen, L.E., Hansen, L.H., Jespersen, H., Podtelejnikov, A., Nielsen, E., Crawford, J., Poulsen, V., Sørensen, B.D., Matthiesen, J., Hendrickson, R.C., Gleeson, F., Pawson, T., Moran, M.F., Durocher, D., Mann, M., Hogue, C.W.V., Figeys, D., Tyers, M., 2002. Systematic identification of protein complexes in Saccharomyces cerevisiae by mass spectrometry. Nature 415, 180–183.

Hoffman, L., Pratt, G., Rechsteiner, M., 1992. Multiple forms of the 20 S multicatalytic and the 26 S ubiquitin/ATP-dependent proteases from rabbit reticulocyte lysate. J. Biol. Chem. 267: 22362-22368.

Hough, R., Pratt, G., Rechsteiner, M., 1986. Ubiquitin-lysozyme conjugates. Identification and characterization of an ATP-dependent protease from rabbit reticulocyte lysates. J. Biol. Chem. 1986 261: 2400-2408.

Hsiao, S.J., Smith, S., 2008. Tankyrase function at telomeres, spindle poles, and beyond. Biochimie 90, 83–92.

Hsu, M.-T., Guo, C.-L., Liou, A.Y., Chang, T.-Y., Ng, M.-C., Florea, B.I., Overkleeft, H.S., Wu, Y.-L., Liao, J.-C., Cheng, P.-L., 2015. Stage-Dependent Axon Transport of Proteasomes Contributes to Axon Development. Dev. Cell 35, 418–431.

Huang, X., Luan, B., Wu, J., Shi, Y., 2016. An atomic structure of the human 26S proteasome. Nat. Struct. Mol. Biol. 23, 778–785.

Huber, E.M., Groll, M., 2017. The Mammalian Proteasome Activator PA28 Forms an Asymmetric α4β3 Complex. Struct. Lond. Engl. 1993 25, 1473-1480.e3.

Huh, K.-W., DeMasi, J., Ogawa, H., Nakatani, Y., Howley, P.M., Münger, K., 2005. Association of the human papillomavirus type 16 E7 oncoprotein with the 600-kDa

retinoblastoma protein-associated factor, p600. Proc. Natl. Acad. Sci. U. S. A. 102, 11492–11497.

Huibregtse, J.M., Scheffner, M., Beaudenon, S., Howley, P.M., 1995. A family of proteins structurally and functionally related to the E6-AP ubiquitin-protein ligase. Proc. Natl. Acad. Sci. U. S. A. 92, 5249.

Huryn, D.M., Kornfilt, D.J.P., Wipf, P., 2019. p97: An Emerging Target for Cancer, Neurodegenerative Diseases, and Viral Infections. J. Med. Chem.

Hwang, J., Qi, L., 2018. Quality Control in the Endoplasmic Reticulum: Crosstalk between ERAD and UPR pathways. Trends Biochem. Sci. 43, 593–605.

Inobe, T., Fishbain, S., Prakash, S., Matouschek, A., 2011. Defining the geometry of the two- component proteasome degron. Nat. Chem. Biol. 7: 161-167.

Inobe, T., Matouschek, A., 2014. Paradigms of protein degradation by the proteasome. Curr. Opin. Struct. Biol. 24, 156–164.

Im, E., Chung, KC., 2016. Precise assembly and regulation of 26S proteasome and correlation between proteasome dysfunction and neurodegenerative diseases. BMB Reports 49, 459.

Irvine, G.B., El-Agnaf, O.M., Shankar, G.M., Walsh, D.M., 2008. Protein aggregation in the brain: the molecular basis for Alzheimer's and Parkinson's diseases. Mol. Med. Camb. Mass 14, 451–464.

Isaacson, R.L., Pye, V.E., Simpson, P., Meyer, H.H., Zhang, X., Freemont, P.S., Matthews, S., 2007. Detailed Structural Insights into the p97-Npl4-Ufd1 Interface. J. Biol. Chem. 282, 21361–21369.

Isakov, E., Stanhill, A., 2011. Stalled proteasomes are directly relieved by P97 recruitment. J. Biol. Chem. 286, 30274–30283.

Islam, M.T., Ogura, T., Esaki, M., 2020. The Cdc48-20S proteasome degrades a class of endogenous proteins in a ubiquitin-independent manner. Biochem. Biophys. Res. Commun.

Itakura, E., Zavodszky, E., Shao, S., Wohlever, M.L., Keenan, R.J., Hegde, R.S., 2016. Ubiquilins Chaperone and Triage Mitochondrial Membrane Proteins for Degradation. Mol. Cell 63, 21–33.

Jacobson, A., Macfadden, A., wu, Z., Peng, J., Liu, C.-W., 2014. Autoregulation of the 26S Proteasome by in situ Ubiquitination. Mol. Biol. Cell 25, 1824-1835.

Jäger, S., Strayle, J., Heinemeyer, W., Wolf, D.H., 2001. Cic1, an adaptor protein specifically linking the 26S proteasome to its substrate, the SCF component Cdc4. EMBO J. 20, 4423–4431.

Jarvis, D.L., 2009. Chapter 14 Baculovirus–Insect Cell Expression Systems, in: Burgess, R.R., Deutscher, M.P. (Eds.), Methods in Enzymology, Guide to Protein Purification, 2nd Edition. Academic Press, pp. 191–222.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., Charpentier, E., 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337, 816–821.

Kabashi, E., Agar, J.N., Taylor, D.M., Minotti, S., Durham, H.D., 2004. Focal dysfunction of the proteasome: a pathogenic factor in a mouse model of amyotrophic lateral sclerosis. J. Neurochem. 89, 1325–1335.

Kamadurai, H.B., Qiu, Y., Deng, A., Harrison, J.S., MacDonald, C., Actis, M., Rodrigues, P., Miller, D.J., Souphron, J., Lewis, S.M., Kurinov, I., Fujii, N., Hammel, M., Piper, R., Kuhlman, B., Schulman, B.A., 2013. Mechanism of ubiquitin ligation and lysine prioritization by a HECT E3. eLife 2, e00828.

Kammerl, I.E., Meiners, S., 2016. Proteasome function shapes innate and adaptive immune responses. Am. J. Physiol. Lung Cell. Mol. Physiol. 311, L328-336.

Kamura, T., Hara, T., Matsumoto, M., Ishida, N., Okumura, F., Hatakeyama, S., Yoshida, M., Nakayama, K., Nakayama, K.I., 2004. Cytoplasmic ubiquitin ligase KPC regulates proteolysis of p27(Kip1) at G1 phase. Nat. Cell Biol. 6, 1229–1235.

Kamura, T., Koepp, D.M., Conrad, M.N., Skowyra, D., Moreland, R.J., Iliopoulos, O., Lane, W.S., Kaelin, W.G., Elledge, S.J., Conaway, R.C., Harper, J.W., Conaway, J.W., 1999. Rbx1, a component of the VHL tumor suppressor complex and SCF ubiquitin ligase. Science 284, 657–661.

Kaplun, L., Tzirkin, R., Bakhrat, A., Shabek, N., Ivantsiv, Y., Raveh, D., 2005. The DNA Damage-Inducible UbL-UbA Protein Ddi1 Participates in Mec1-Mediated Degradation of Ho Endonuclease. Mol. Cell. Biol. 25, 5355–5362.

Kaur, G., Subramanian, S., 2015. The UBR-box and its relationship to binuclear RING-like treble clef zinc fingers. Biol. Direct 10, 36.

Keck, S., Nitsch, R., Grune, T., Ullrich, O., 2003. Proteasome inhibition by paired helical filament-tau in brains of patients with Alzheimer's disease. J. Neurochem. 85, 115–122.

Kern, A., Behl, C., 2019. Special Issue on "Proteostasis and Autophagy." Cells 8.

Khare, S., Nagle, A.S., Biggart, A., Lai, Y.H., Liang, F., Davis, L.C., Barnes, S.W., Mathison, C.J.N., Myburgh, E., Gao, M.-Y., Gillespie, J.R., Liu, X., Tan, J.L., Stinson, M., Rivera, I.C., Ballard, J., Yeh, V., Groessl, T., Federe, G., Koh, H.X.Y., Venable, J.D., Bursulaya, B., Shapiro, M., Mishra, P.K., Spraggon, G., Brock, A., Mottram, J.C., Buckner, F.S., Rao, S.P.S., Wen, B.G., Walker, J.R., Tuntland, T., Molteni, V., Glynne, R.J., Supek, F., 2016. Proteasome inhibition for treatment of leishmaniasis, Chagas disease and sleeping sickness. Nature 537, 229–233.

Khor, B., Bredemeyer, A.L., Huang, C.-Y., Turnbull, I.R., Evans, R., Maggi, L.B., White, J.M., Walker, L.M., Carnes, K., Hess, R.A., Sleckman, B.P., 2006. Proteasome activator PA200 is required for normal spermatogenesis. Mol. Cell. Biol. 26, 2999–3007.

Khoshouei, M., Radjainia, M., Baumeister, W., Danev, R., 2017. Cryo-EM structure of haemoglobin at 3.2 Å determined with the Volta phase plate. Nat. Commun. 8.

Kim, I., Mi, K., Rao, H., 2004. Multiple Interactions of Rad23 Suggest a Mechanism for Ubiquitylated Substrate Delivery Important in Proteolysis. Mol. Biol. Cell 15, 3357–3365.

Kimonis, V.E., Fulchiero, E., Vesa, J., Watts, G., 2008. VCP disease associated with myopathy, Paget disease of bone and frontotemporal dementia: review of a unique disorder. Biochim. Biophys. Acta 1782, 744–748.

Kirk, R., Laman, H., Knowles, P.P., Murray-Rust, J., Lomonosov, M., Meziane, E.K., McDonald, N.Q., 2008. Structure of a Conserved Dimerization Domain within the F-box Protein Fbxo7 and the PI31 Proteasome Inhibitor. J. Biol. Chem. 283, 22325–22335.

Kish-Trier, E., Hill, C.P., 2013. Structural biology of the proteasome. Annu. Rev. Biophys. 42, 29–49.

Kisselev, A.F., Goldberg, A.L., 2001. Proteasome inhibitors: from research tools to drug candidates. Chem. Biol. 8, 739–758.

Kleiger, G., Mayor, T., 2014. Perilous journey: a tour of the ubiquitin–proteasome system. Trends Cell Biol. 24, 352–359.

Klenk, H.-D., 2002. Post-translational modifications in insect cells, in: Vlak, J.M., de Gooijer, C.D., Tramper, J., Miltenburger, H.G. (Eds.), Insect Cell Culture: Fundamental and Applied Aspects, Current Applications of Cell Culture Engineering. Springer Netherlands, Dordrecht, pp. 139–144.

Kloetzel, P.-M., 2004. The proteasome and MHC class I antigen processing. Biochim. Biophys. Acta 1695, 225–233.

Knowlton, J.R., Johnston, S.C., Whitby, F.G., Realini, C., Zhang, Z., Rechsteiner, M., Hill, C.P., 1997. Structure of the proteasome activator REGα (PA28α). Nature 390, 639–643.

Kofler, M., Motzny, K., Beyermann, M., Freund, C., 2005. Novel interaction partners of the CD2BP2-GYF domain. J. Biol. Chem. 280, 33397–33402.

Koizumi, S., Irie, T., Hirayama, S., Sakurai, Y., Yashiroda, H., Naguro, I., Ichijo, H., Hamazaki, J., Murata, S., 2016. The aspartyl protease DDI2 activates Nrf1 to compensate for proteasome dysfunction. eLife. e18357.

Koleva, R.I., Conley, B.A., Romero, D., Riley, K.S., Marto, J.A., Lux, A., Vary, C.P.H., 2006. Endoglin structure and function: Determinants of endoglin phosphorylation by transforming growth factor-beta receptors. J. Biol. Chem. 281, 25110–25123.

Komander, D., Rape, M., 2012. The ubiquitin code. Annu. Rev. Biochem. 81, 203–229.

Koning, R.I., Koster, A.J., Sharp, T.H., 2018. Advances in cryo-electron tomography for biology and medicine. Ann. Anat. - Anat. Anz. 217, 82–96.

Krause, S., Kuckelkorn, U., Dörner, T., Burmester, G.-R., Feist, E., Kloetzel, P.-M., 2006. Immunoproteasome subunit LMP2 expression is deregulated in Sjogren's syndrome but not in other autoimmune disorders. Ann. Rheum. Dis. 65, 1021–1027.

Kreko-Pierce, T., Eaton, B.A., 2017. The Drosophila LC8 homolog cut up specifies the axonal transport of proteasomes. J. Cell Sci. 130, 3388–3398.

Krishnan, K.M., Williamson, K.C., 2018. The proteasome as a target to combat malaria: hits and misses. Transl. Res. J. Lab. Clin. Med. 198, 40–47.

Kühlbrandt, W., 2014. Cryo-EM enters a new era. eLife 3, e03678.

Kunjappu, M.J., Hochstrasser, M., 2014. Assembly of the 20S proteasome. Biochim. Biophys. Acta 1843, 2–12.

Kurland, C.G., 1991. Codon bias and gene expression. FEBS Lett. 285, 165–169.

Kwon, Y.T., Reiss, Y., Fried, V.A., Hershko, A., Yoon, J.K., Gonda, D.K., Sangan, P., Copeland, N.G., Jenkins, N.A., Varshavsky, A., 1998. The mouse and human genes encoding the recognition component of the N-end rule pathway. Proc. Natl. Acad. Sci. U. S. A. 95, 7898–7903.

Kwon, Y.T., Xia, Z., An, J.Y., Tasaki, T., Davydov, I.V., Seo, J.W., Sheng, J., Xie, Y., Varshavsky, A., 2003. Female Lethality and Apoptosis of Spermatocytes in Mice

Lacking the UBR2 Ubiquitin Ligase of the N-End Rule Pathway. Mol. Cell. Biol. 23, 8255–8271.

Kwon, Y.T., Xia, Z., Davydov, I.V., Lecker, S.H., Varshavsky, A., 2001. Construction and Analysis of Mouse Strains Lacking the Ubiquitin Ligase UBR1 (E3α) of the N-End Rule Pathway. Mol. Cell. Biol. 21, 8007–8021.

Lam, Y.A., Pickart, C.M., Alban, A., Landon, M., Jamieson, C., Ramage, R., Mayer, R.J., Layfield, R., 2000. Inhibition of the ubiquitin-proteasome system in Alzheimer's disease. Proc. Natl. Acad. Sci. U. S. A. 97, 9902–9906.

Lander, G.C., Estrin, E., Matyskiela, M.E., Bashore, C., Nogales, E., Martin, A., 2012. Complete subunit architecture of the proteasome regulatory particle. Nature 482, 186–191.

Lasker, K., Förster, F., Bohn, S., Walzthoeni, T., Villa, E., Unverdorben, P., Beck, F., Aebersold, R., Sali, A., Baumeister, W., 2012. Molecular architecture of the 26S proteasome holocomplex determined by an integrative approach. Proc. Natl. Acad. Sci. 109, 1380–1387.

Lee, D.H., Goldberg, A.L., 1998. Proteasome inhibitors: valuable new tools for cell biologists. Trends Cell Biol. 8, 397–403.

Leestemaker, Y., de Jong, A., Witting, K.F., Penning, R., Schuurman, K., Rodenko, B., Zaal, E.A., van de Kooij, B., Laufer, S., Heck, A.J.R., Borst, J., Scheper, W., Berkers, C.R., Ovaa, H., 2017. Proteasome Activation by Small Molecules. Cell Chem. Biol. 24, 725-736.

Leggett, D.S., Hanna, J., Borodovsky, A., Crosas, B., Schmidt, M., Baker, R.T., Walz, T., Ploegh, H., Finley, D., 2002. Multiple associated proteins regulate proteasome structure and function. Mol. Cell 10, 495–507.

Lehrbach, N.J., Ruvkun, G., 2016. Proteasome dysfunction triggers activation of SKN-1A/Nrf1 by the aspartic protease DDI-1. eLife 5, e17721.

Lev, S., 2012. Nonvesicular lipid transfer from the endoplasmic reticulum. Cold Spring Harb. Perspect. Biol. 4, a013300.

Li, H., O'Donoghue, A.J., van der Linden, W.A., Xie, S.C., Yoo, E., Foe, I.T., Tilley, L., Craik, C.S., da Fonseca, P.C.A., Bogyo, M., 2016. Structure- and function-based design of Plasmodium-selective proteasome inhibitors. Nature 530, 233–236.

Li, H., Ponder, E.L., Verdoes, M., Asbjornsdottir, K.H., Deu, E., Edgington, L.E., Lee, J.T., Kirk, C.J., Demo, S.D., Williamson, K.C., Bogyo, M., 2012. Validation of the

proteasome as a therapeutic target in Plasmodium using an epoxyketone inhibitor with parasite-specific toxicity. Chem. Biol. 19, 1535–1545.

Li, H., van der Linden, W.A., Verdoes, M., Florea, B.I., McAllister, F.E., Govindaswamy, K., Elias, J.E., Bhanot, P., Overkleeft, H.S., Bogyo, M., 2014. Assessing Subunit Dependency of the Plasmodium Proteasome Using Small Molecule Inhibitors and Active Site Probes. ACS Chem. Biol. 9, 1869–1876.

Li, W., Bengtson, M.H., Ulbrich, A., Matsuda, A., Reddy, V.A., Orth, A., Chanda, S.K., Batalov, S., Joazeiro, C.A.P., 2008. Genome-wide and functional annotation of human E3 ubiquitin ligases identifies MULAN, a mitochondrial E3 that regulates the organelle's dynamics and signaling. PloS One 3, e1487.

Li, X., Kusmierczyk, A.R., Wong, P., Emili, A., Hochstrasser, M., 2007. beta-Subunit appendages promote 20S proteasome assembly by overcoming an Ump1-dependent checkpoint. EMBO J. 26, 2339–2349.

Li, X., Mooney, P., Zheng, S., Booth, C.R., Braunfeld, M.B., Gubbens, S., Agard, D.A., Cheng, Y., 2013. Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. Nat. Methods 10, 584–590.

Li, X., Thompson, D., Kumar, B., DeMartino, G.N., 2014. Molecular and cellular roles of PI31 (PSMF1) protein in regulation of proteasome function. J. Biol. Chem. 289, 17392–17405.

Li, X.-J., Li, S., 2011. Proteasomal dysfunction in aging and Huntington disease. Neurobiol. Dis., Autophagy and protein degradation in neurological diseases 43, 4–8.

Li, Z.-H., Wang, Y., Xu, M., Jiang, T., 2017. Crystal structures of the UBX domain of human UBXD7 and its complex with p97 ATPase. Biochem. Biophys. Res. Commun. 486, 94–100.

Liang, R.-Y., Chen, L., Ko, B.-T., Shen, Y.-H., Li, Y.-T., Chen, B.-R., Lin, K.-T., Madura, K., Chuang, S.-M., 2014. Rad23 interaction with the proteasome is regulated by phosphorylation of its ubiquitin-like (UbL) domain. J. Mol. Biol. 426, 4049–4060.

Linares, L.K., Hengstermann, A., Ciechanover, A., Müller, S., Scheffner, M., 2003. HdmX stimulates Hdm2-mediated ubiquitination and degradation of p53. Proc. Natl. Acad. Sci. 100, 12009–12014.

Lindsten, K., de Vrij, F.M.S., Verhoef, L.G.G.C., Fischer, D.F., van Leeuwen, F.W., Hol, E.M., Masucci, M.G., Dantuma, N.P., 2002. Mutant ubiquitin found in

neurodegenerative disorders is a ubiquitin fusion degradation substrate that blocks proteasomal degradation. J. Cell Biol. 157, 417–427.

Ling, H.H., Beaulé, C., Chiang, C.-K., Tian, R., Figeys, D., Cheng, H.-Y.M., 2014. Time-of-Day- and Light-Dependent Expression of Ubiquitin Protein Ligase E3 Component N-Recognin 4 (UBR4) in the Suprachiasmatic Nucleus Circadian Clock. PLoS ONE 9.

Liu, C.-W., Li, X., Thompson, D., Wooding, K., Chang, T., Tang, Z., Yu, H., Thomas, P.J., DeMartino, G.N., 2006. ATP binding and ATP hydrolysis play distinct roles in the function of 26S proteasome. Mol. Cell 24, 39–50.

Liu, K., Jones, S., Minis, A., Rodriguez, J., Molina, H., Steller, H., 2019. PI31 Is an Adaptor Protein for Proteasome Transport in Axons and Required for Synaptic Development. Dev. Cell 50, 509-524.

Liu, Y., Ye, Y., 2012. Roles of p97-associated deubiquitinases in protein quality control at the endoplasmic reticulum. Curr. Protein Pept. Sci. 13, 436–446.

Livneh, I., Cohen-Kaplan, V., Cohen-Rosenzweig, C., Avni, N., Ciechanover, A., 2016. The life cycle of the 26S proteasome: from birth, through regulation and function, and onto its death. Cell Res. 26, 869–885.

Lorick, K.L., Jensen, J.P., Fang, S., Ong, A.M., Hatakeyama, S., Weissman, A.M., 1999. RING fingers mediate ubiquitin-conjugating enzyme (E2)-dependent ubiquitination. Proc. Natl. Acad. Sci. U. S. A. 96, 11364–11369.

Löwe, J., Stock, D., Jap, B., Zwickl, P., Baumeister, W., Huber, R., 1995. Crystal structure of the 20S proteasome from the archaeon T. acidophilum at 3.4 A resolution. Science 268, 533–539.

Ludtke, S.J., Baldwin, P.R., Chiu, W., 1999. EMAN: semiautomated software for high-resolution single-particle reconstructions. J. Struct. Biol. 128, 82–97.

Lundgren, J., Masson, P., Realini, C.A., Young, P., 2003. Use of RNA Interference and Complementation To Study the Function of the Drosophila and Human 26S Proteasome Subunit S13. Mol. Cell. Biol. 23, 5320–5330.

Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., Church, G.M., 2013. RNA-guided human genome engineering via Cas9. Science 339, 823–

Manasanch, E.E., Orlowski, R.Z., 2017. Proteasome inhibitors in cancer therapy. Nat. Rev. Clin. Oncol. 14, 417–433.

Mao, I., Liu, J., Li, X., Luo, H., 2008. REGγ, a proteasome activator and beyond? Cell. Mol. Life Sci. 65, 3971.

Marabini, R., Masegosa, I.M., San Martín, M.C., Marco, S., Fernández, J.J., de la Fraga, L.G., Vaquerizo, C., Carazo, J.M., 1996. Xmipp: An Image Processing Package for Electron Microscopy. J. Struct. Biol. 116, 237–240.

Masson, P., Lundin, D., Söderbom, F., Young, P., 2009. Characterization of a REG/PA28 Proteasome Activator Homolog in Dictyostelium discoideum Indicates that the Ubiquitin- and ATP-Independent REGγ Proteasome Is an Ancient Nuclear Protease. Eukaryot. Cell 8, 844–851.

Matouschek, A., Finley, D., 2012. An Ancient Portal to Proteolysis. Science 337, 813–814.

McCutchen-Maloney, S.L., Matsuda, K., Shimbara, N., Binns, D.D., Tanaka, K., Slaughter, C.A., DeMartino, G.N., 2000. cDNA Cloning, Expression, and Functional Characterization of PI31, a Proline-rich Inhibitor of the Proteasome. J. Biol. Chem. 275, 18557–18565.

McIntosh, R., Nicastro, D., Mastronarde, D., 2005. New views of cells in 3D: an introduction to electron tomography. Trends Cell Biol. 15, 43–51.

McMullan, G., Faruqi, A.R., Clare, D., Henderson, R., 2014. Comparison of optimal performance at 300keV of three direct electron detectors for use in low dose electron microscopy. Ultramicroscopy 147, 156–163.

McNaught, K.S., Jenner, P., 2001. Proteasomal function is impaired in substantia nigra in Parkinson's disease. Neurosci. Lett. 297, 191–194.

Mendez, J.H., Mehrani, A., Randolph, P., Stagg, S., 2019. Throughput and resolution with a next-generation direct electron detector. IUCrJ 6, 1007-1013.

Metzger, M.B., Pruneda, J.N., Klevit, R.E., Weissman, A.M., 2014. RING-type E3 ligases: master manipulators of E2 ubiquitin-conjugating enzymes and ubiquitination. Biochim. Biophys. Acta 1843, 47–60.

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

Meyer, H., Weihl, C.C., 2014. The VCP/p97 system at a glance: connecting cellular function to disease pathogenesis. J. Cell Sci. 127, 3877–3883.

Miller, Z., Ao, L., Kim, K.B., Lee, W., 2013. Inhibitors of the immunoproteasome: current status and future directions. Curr. Pharm. Des. 19, 4140–4151.

Minis, A., Rodriguez, J., Levin, A., Liu, K., Govek, E.-E., Hatten, M.E., Steller, H., 2019. The proteasome regulator PI31 is required for protein homeostasis, synapse

maintenance and neuronal survival in mice, 2019. Proc. Nat. Acad. Sci. 116, 24639-24650.

Mitra, K., Gangopadhaya, P.K., Das, S.K., 2003. Parkinsonism plus syndrome--a review. Neurol. India 51, 183–188.

Morozov, A.V., Karpov, V.L., 2019. Proteasomes and Several Aspects of Their Heterogeneity Relevant to Cancer. Front. Oncol. 9, 761.

Morris, E.P., da Fonseca, P.C.A., 2017. High-resolution cryo-EM proteasome structures in drug development. Acta Crystallogr. Sect. Struct. Biol. 73, 522–533.

Mortimore, G.E., Pösö, A.R., 1987. Intracellular Protein Catabolism and its Control During Nutrient Deprivation and Supply. Annu. Rev. Nutr. 7, 539–568.

Mortimore, G.E., Schworer, C.M., 1977. Induction of autophagy by amino-acid deprivation in perfused rat liver. Nature 270, 174–176.

Mott, J.D., Pramanik, B.C., Moomaw, C.R., Afendis, S.J., DeMartino, G.N., Slaughter, C.A., 1994. PA28, an activator of the 20 S proteasome, is composed of two nonidentical but homologous subunits. J. Biol. Chem. 269, 31466–31471.

Murata, K., Wolf, M., 2018. Cryo-electron microscopy for structural analysis of dynamic biological macromolecules. Biochim. Biophys. Acta Gen. Subj. 1862, 324–334.

Murata, S., Sasaki, K., Kishimoto, T., Niwa, S.-I., Hayashi, H., Takahama, Y., Tanaka, K., 2007. Regulation of CD8+ T cell development by thymus-specific proteasomes. Science 316, 1349–1353.

Murata, S., Yashiroda, H., Tanaka, K., 2009. Molecular mechanisms of proteasome assembly. Nat. Rev. Mol. Cell Biol. 10, 104–115.

Myeku, N., Clelland, C.L., Emrani, S., Kukushkin, N.V., Yu, W.H., Goldberg, A.L., Duff, K.E., 2016. Tau-driven 26S proteasome impairment and cognitive dysfunction can be prevented early in disease by activating cAMP-PKA signaling. Nat. Med. 22, 46–53.

Nakatani, Y., Konishi, H., Vassilev, A., Kurooka, H., Ishiguro, K., Sawada, J., Ikura, T., Korsmeyer, S.J., Qin, J., Herlitz, A.M., 2005. p600, a unique protein required for membrane morphogenesis and cell survival. Proc. Natl. Acad. Sci. 102, 15093–15098.

Nandi, D., Tahiliani, P., Kumar, A., Chandu, D., 2006. The ubiquitin-proteasome system. J. Biosci. 31, 137–155.

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

Naujokat, C., Hoffmann, S., 2002. Role and function of the 26S proteasome in proliferation and apoptosis. Lab. Investig. J. Tech. Methods Pathol. 82, 965–980.

Nie, J., Wu, M., Wang, J., Xing, G., He, F., Zhang, L., 2010. REGgamma proteasome mediates degradation of the ubiquitin ligase Smurf1. FEBS Lett. 584, 3021–3027.

Niwa, H., Ewens, C.A., Tsang, C., Yeung, H.O., Zhang, X., Freemont, P.S., 2012. The Role of the N-Domain in the ATPase Activity of the Mammalian AAA ATPase p97/VCP. J. Biol. Chem. 287, 8561–8570.

Nonaka, T., Arai, T., Buratti, E., Baralle, F.E., Akiyama, H., Hasegawa, M., 2009. Phosphorylated and ubiquitinated TDP-43 pathological inclusions in ALS and FTLD-U are recapitulated in SH-SY5Y cells. FEBS Lett. 583, 394–400.

Novikoff, A.B., 1959. The Proximal Tubule Cell in Experimental Hydronephrosis. J. Biophys. Biochem. Cytol. 6, 136–138.

Novikoff, A.B., Essner, E., Goldfischer, S., Heus, M., 1962. Nucleosidephosphatase activities of cytomembranes. Symp. Intl. Soc. Cell Bio., 1, 1490192.

Nowicka, U., Zhang, D., Walker, O., Krutauz, D., Castañeda, C.A., Chaturvedi, A., Chen, T.Y., Reis, N., Glickman, M.H., Fushman, D., 2015. DNA-damage-inducible 1 protein (Ddi1) contains an uncharacteristic ubiquitin-like domain that binds ubiquitin. Struct. Lond. Engl. 23, 542–557.

Oh, S., Hong, H.S., Hwang, E., Sim, H.J., Lee, W., Shin, S.J., Mook-Jung, I., 2005. Amyloid peptide attenuates the proteasome activity in neuronal cells. Mech. Ageing Dev. 126, 1292–1299.

Ohara, O., Nagase, T., Ishikawa, K., Nakajima, D., Ohira, M., Seki, N., Nomura, N., 1997. Construction and characterization of human brain cDNA libraries suitable for analysis of cDNA clones encoding relatively large proteins. DNA Res. Int. J. Rapid Publ. Rep. Genes Genomes 4, 53–59.

Ohta, T., Michel, J.J., Schottelius, A.J., Xiong, Y., 1999. ROC1, a Homolog of APC11, Represents a Family of Cullin Partners with an Associated Ubiquitin Ligase Activity. Mol. Cell 3, 535–541.

Orlova, E.V., Saibil, H.R., 2011. Structural analysis of macromolecular assemblies by electron microscopy. Chem. Rev. 111, 7710–7748.

Otero, M.G., Alloatti, M., Cromberg, L., Almenar, A., Encalada, S., Pozo Devoto, V., Bruno, L., Goldstein, L., Falzone, T., 2014. Fast axonal transport of the proteasome complex depends on membrane interaction and molecular motor function. J. Cell Sci. 127, 1537-1549.

Ott, D.E., Coren, L.V., Sowder, R.C., Adams, J., Schubert, U., 2003. Retroviruses Have Differing Requirements for Proteasome Function in the Budding Process. J. Virol. 77, 3384–3393.

Paisán-Ruiz, C., Guevara, R., Federoff, M., Hanagasi, H., Sina, F., Elahi, E., Schneider, S.A., Schwingenschuh, P., Bajaj, N., Emre, M., Singleton, A.B., Hardy, J., Bhatia, K.P., Brandner, S., Lees, A.J., Houlden, H., 2010. Early-onset L-dopa-responsive parkinsonism with pyramidal signs due to ATP13A2, PLA2G6, FBXO7 and spatacsin mutations. Mov. Disord. Off. J. Mov. Disord. Soc. 25, 1791–1800.

Park, J.E., Miller, Z., Jun, Y., Lee, W., Kim, K.B., 2018. Next-generation proteasome inhibitors for cancer therapy. Transl. Res. J. Lab. Clin. Med. 198, 1–16.

Park, Y., Hwang, Y.-P., Lee, J.-S., Seo, S.-H., Yoon, S.K., Yoon, J.-B., 2005. Proteasomal ATPase-associated factor 1 negatively regulates proteasome activity by interacting with proteasomal ATPases. Mol. Cell. Biol. 25, 3842–3853.

Peters, J.M., Cejka, Z., Harris, J.R., Kleinschmidt, J.A., Baumeister, W., 1993. Structural features of the 26 S proteasome complex. J. Mol. Biol. 234, 932–937.

Peters, J.M., Walsh, M.J., Franke, W.W., 1990. An abundant and ubiquitous homo-oligomeric ring-shaped ATPase particle related to the putative vesicle fusion proteins Sec18p and NSF. EMBO J. 9, 1757–1767.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E., 2004. UCSF Chimera--a visualization system for exploratory research and analysis. J. Comput. Chem. 25, 1605–1612.

Pick, E., Berman, T., 2013. Formation of alternative proteasomes: Same lady, different cap? FEBS Lett. 587, 389-393.

Potuschak, T., Stary, S., Schlögelhofer, P., Becker, F., Nejinskaia, V., Bachmair, A., 1998. PRT1 of Arabidopsis thaliana encodes a component of the plant N-end rule pathway. Proc. Natl. Acad. Sci. 95, 7904–7908.

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

Raasi, S., Wolf, D.H., 2007. Ubiquitin receptors and ERAD: a network of pathways to the proteasome. Semin. Cell Dev. Biol. 18, 780–791.

Rabinovitz, M., Fisher, J.M., 1964. Characteristics of the inhibition of hemoglobin synthesis in rabbit reticulocytes by threo-α-amino-β-chlorobutyric acid. Biochim. Biophys. Acta BBA - Spec. Sect. Nucleic Acids Relat. Subj. 91, 313–322.

Rabl, J., Smith, D.M., Yu, Y., Chang, S.-C., Goldberg, A.L., Cheng, Y., 2008. Mechanism of gate opening in the 20S proteasome by the proteasomal ATPases. Mol. Cell 30, 360–368.

Radermacher, M., 1988. Three-Dimensional reconstruction of single particles from random and nonrandom tilt series. J. Electron Microsc. Tech. 9, 359–394.

Radhakrishnan, S.K., Lee, C.S., Young, P., Beskow, A., Chan, J.Y., Deshaies, R.J., 2010. Transcription factor Nrf1 mediates the proteasome recovery pathway after proteasome inhibition in mammalian cells. Mol. Cell 38, 17–28.

Ramos, P.C., Höckendorff, J., Johnson, E.S., Varshavsky, A., Dohmen, R.J., 1998. Ump1p is required for proper maturation of the 20S proteasome and becomes its substrate upon completion of the assembly. Cell 92, 489–499.

Ran, F.A., Hsu, P.D., Lin, C.-Y., Gootenberg, J.S., Konermann, S., Trevino, A.E., Scott, D.A., Inoue, A., Matoba, S., Zhang, Y., Zhang, F., 2013. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. Cell 154, 1380–1389.

Ray, M.K., Fagan, S.P., Brunicardi, F.C., 2000. The Cre–loxP System: A Versatile Tool for Targeting Genes in a Cell- and Stage-Specific Manner. Cell Transplant. 9, 805–815.

Rechsteiner, M., Hill, C.P., 2005. Mobilizing the proteolytic machine: cell biological roles of proteasome activators and inhibitors. Trends Cell Biol. 15, 27–33.

Rêgo, A.T., Fonseca, P.C.A. da, 2019. Characterization of Fully Recombinant Human 20S and 20S-PA200 Proteasome Complexes. Mol. Cell 76, 138-147.

Reijden, B.A.V.D., Erpelinck-Verschueren, C.A.J., Löwenberg, B., Jansen, J.H., 1999. TRIADs: A new class of proteins with a novel cysteine-rich signature. Protein Sci. 8, 1557–1561.

Reyes-Turcu, F.E., Venti, K.H., Wilkinson, KD., 2009. Regulation and cellular roles of ubiquitin-specific deubiquitinating enzymes. Annual Review of Biochemistry 78, 363- 397.

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

Riley, B.E., Lougheed, J.C., Callaway, K., Velasquez, M., Brecht, E., Nguyen, L., Shaler, T., Walker, D., Yang, Y., Regnstrom, K., Diep, L., Zhang, Z., Chiou, S., Bova, M., Artis, D.R., Yao, N., Baker, J., Yednock, T., Johnston, J.A., 2013. Structure and function of Parkin E3 ubiquitin ligase reveals aspects of RING and HECT ligases. Nat. Commun. 4, 1982.

Ritz, D., Vuk, M., Kirchner, P., Bug, M., Schütz, S., Hayer, A., Bremer, S., Lusk, C., Baloh, R.H., Lee, H., Glatter, T., Gstaiger, M., Aebersold, R., Weihl, C.C., Meyer, H., 2011. Endolysosomal sorting of ubiquitylated caveolin-1 is regulated by VCP and UBXD1 and impaired by VCP disease mutations. Nat. Cell Biol. 13, 1116–1123.

Rock, K.L., Goldberg, A.L., 1999. Degradation of cell proteins and the generation of MHC class I-presented peptides. Annu. Rev. Immunol. 17, 739–779.

Rock, K.L., Gramm, C., Rothstein, L., Clark, K., Stein, R., Dick, L., Hwang, D., Goldberg, A.L., 1994. Inhibitors of the proteasome block the degradation of most cell proteins and the generation of peptides presented on MHC class I molecules. Cell 78, 761–771.

Rosenthal, P.B., Henderson, R., 2003. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. J. Mol. Biol. 333, 721–745.

Ross, C.A., Pickart, C.M., 2004. The ubiquitin-proteasome pathway in Parkinson's disease and other neurodegenerative diseases. Trends Cell Biol. 14, 703–711.

Ross, C.A., Poirier, M.A., 2004. Protein aggregation and neurodegenerative disease. Nat. Med. 10 Suppl, S10-17.

Rotin, D., Kumar, S., 2009. Physiological functions of the HECT family of ubiquitin ligases. Nat. Rev. Mol. Cell Biol. 10, 398–409.

Rouiller, I., DeLaBarre, B., May, A.P., Weis, W.I., Brunger, A.T., Milligan, R.A., Wilson-Kubalek, E.M., 2002. Conformational changes of the multifunction p97 AAA ATPase during its ATPase cycle. Nat. Struct. Biol. 9, 950–957.

Rousseau, A., Bertolotti, A., 2018. Regulation of proteasome assembly and activity in health disease. Nat. Rev. Mol. Cell. Biol. 19, 697-712.

Ruprecht, J., Nield, J., 2001. Determining the structure of biological macromolecules by transmission electron microscopy, single particle analysis and 3D reconstruction. Prog. Biophys. Mol. Biol. 75, 121–164.

Ruschak, A.M., Slassi, M., Kay, L.E., Schimmer, A.D., 2011. Novel proteasome inhibitors to overcome bortezomib resistance. J. Natl. Cancer Inst. 103, 1007–1017.

Sadre-Bazzaz, K., Whitby, F.G., Robinson, H., Formosa, T., Hill, C.P., 2010. Structure of a Blm10 complex reveals common mechanisms for proteasome binding and gate opening. Mol. Cell 37, 728–735.

Saeki, Y., 2017. Ubiquitin recognition by the proteasome. J. Biochem. (Tokyo) 161, 113–124.

Sahin, E., Roberts, C.J., 2012. Size-exclusion chromatography with multi-angle light scattering for elucidating protein aggregation mechanisms. Methods Mol. Biol. Clifton NJ 899, 403–423.

Sakata, E., Bohn, S., Mihalache, O., Kiss, P., Beck, F., Nagy, I., Nickell, S., Tanaka, K., Saeki, Y., Förster, F., Baumeister, W., 2012. Localization of the proteasomal ubiquitin receptors Rpn10 and Rpn13 by electron cryomicroscopy. Proc. Natl. Acad. Sci. 109, 1479–1484.

Sakata, E., Yamaguchi, Y., Kurimoto, E., Kikuchi, J., Yokoyama, S., Yamada, S., Kawahara, H., Yokosawa, H., Hattori, N., Mizuno, Y., Tanaka, K., Kato, K., 2003. Parkin binds the Rpn10 subunit of 26S proteasomes through its ubiquitin-like domain. EMBO Rep. 4, 301–306.

Saxton, W.O., Baumeister, W., 1982. The correlation averaging of a regularly arranged bacterial cell envelope protein. J. Microsc. 127, 127–138.

Scanlon, T.C., Gottlieb, B., Durcan, T.M., Fon, E.A., Beitel, L.K., Trifiro, M.A., 2009. Isolation of human proteasomes and putative proteasome-interacting proteins using a novel affinity chromatography method. Exp. Cell Res. 315, 176–189.

Scheffner, M., Staub, O., 2007. HECT E3s and human disease. BMC Biochem. 8, S6.

Scheres, S.H.W., 2012. RELION: implementation of a Bayesian approach to cryo-EM structure determination. J. Struct. Biol. 180, 519–530.

Schmidt, M., Hanna, J., Elsasser, S., Finley, D., 2005. Proteasome-associated proteins: regulation of a proteolytic machine. Biol. Chem. 386, 725–737.

Schmidt, M., Schmidtke, G., Kloetzel, P.M., 1997. Structure and structure formation of the 20S proteasome. Mol. Biol. Rep. 24, 103–112.

Schoenheimer, R., 1946. The dynamic state of body constituents. Dyn. State Body Const.

Schrader, J., Henneberg, F., Mata, R.A., Tittmann, K., Schneider, T.R., Stark, H., Bourenkov, G., Chari, A., 2016. The inhibition mechanism of human 20S proteasomes enables next-generation inhibitor design. Science 353, 594–598.

Schweitzer, A., Aufderheide, A., Rudack, T., Beck, F., Pfeifer, G., Plitzko, J.M., Sakata, E., Schulten, K., Förster, F., Baumeister, W., 2016. Structure of the human 26S proteasome at a resolution of 3.9 Å. Proc. Natl. Acad. Sci. 113, 7816–7821.

Seemüller, E., Lupas, A., Stock, D., Löwe, J., Huber, R., Baumeister, W., 1995. Proteasome from Thermoplasma acidophilum: a threonine protease. Science 268, 579–582.

Segal, H.I., Winkler, J.R., Miyagi, M.P., 1974. Relationship between Degradation Rates of Proteins in Vivo and Their Susceptibility to Lysosomal Proteases. J. Biol. Chem. 249, 6364–6365.

Seglen, P.O., Gordon, P.B., 1982. 3-Methyladenine: Specific inhibitor of autophagic/lysosomal protein degradation in isolated rat hepatocytes. Proc. Natl. Acad. Sci. U. S. A. 79, 1889–1892.

Seifert, U., Bialy, L.P., Ebstein, F., Bech-Otschir, D., Voigt, A., Schröter, F., Prozorovski, T., Lange, N., Steffen, J., Rieger, M., Kuckelkorn, U., Aktas, O., Kloetzel, P.-M., Krüger, E., 2010. Immunoproteasomes preserve protein homeostasis upon interferon-induced oxidative stress. Cell 142, 613–624.

Seol, J.H., Feldman, R.M., Zachariae, W., Shevchenko, A., Correll, C.C., Lyapina, S., Chi, Y., Galova, M., Claypool, J., Sandmeyer, S., Nasmyth, K., Deshaies, R.J., Shevchenko, A., Deshaies, R.J., 1999. Cdc53/cullin and the essential Hrt1 RING-H2 subunit of SCF define a ubiquitin ligase module that activates the E2 enzyme Cdc34. Genes Dev. 13, 1614–1626.

Shang, J., Huang, X., Du, Z., 2015. The FP domains of PI31 and Fbxo7 have the same protein fold but very different modes of protein–protein interaction. J. Biomol. Struct. Dyn. 33, 1528–1538.

Sharp, D., Lanni, F., Beard, J., 1950. The egg white inhibitor of influenza virus hemagglutination II. Electron microscopy of the inhibitor. The Journal of Biological Chemistry 185, 681–688.

Shastri, N., Serwold, T., Paz, P., 1998. Reading within the lines: naturally processed peptides displayed by MHC class I molecules. Curr. Opin. Immunol. 10, 137–144.

Sherva, R., Baldwin, C.T., Inzelberg, R., Vardarajan, B., Cupples, L.A., Lunetta, K., Bowirrat, A., Naj, A., Pericak-Vance, M., Friedland, R.P., Farrer, L.A., 2011.

Identification of Novel Candidate Genes for Alzheimer Disease by Autozygosity Mapping Using Genome Wide SNP Data From an Israeli-Arab Community. J. Alzheimers Dis. JAD 23.

Shi, Yuan, Chen, X., Elsasser, S., Stocks, B.B., Tian, G., Lee, B.-H., Shi, Yanhong, Zhang, N., de Poot, S.A.H., Tuebing, F., Sun, S., Vannoy, J., Tarasov, S.G., Engen, J.R., Finley, D., Walters, K.J., 2016. Rpn1 provides adjacent receptor sites for substrate binding and deubiquitination by the proteasome. Science 351, 9421.

Shimura, H., Hattori, N., Kubo, S. i, Mizuno, Y., Asakawa, S., Minoshima, S., Shimizu, N., Iwai, K., Chiba, T., Tanaka, K., Suzuki, T., 2000. Familial Parkinson disease gene product, parkin, is a ubiquitin-protein ligase. Nat. Genet. 25, 302–305.

Sigworth, F.J., 2004. Classical detection theory and the cryo-EM particle selection problem. J. Struct. Biol., Automated Particle Selection for Cryo-Electron Microscopy 145, 111–122.

Sigworth, F.J., Doerschuk, P.C., Carazo, J.-M., Scheres, S.H.W., 2010. An introduction to maximum-likelihood methods in cryo-EM. Methods Enzymol. 482, 263–294.

Simpson, M.V., 1953. The Release of Labeled Amino Acids from the Proteins of Rat Liver Slices. J. Biol. Chem. 201, 143–154.

Skowyra, D., Craig, K.L., Tyers, M., Elledge, S.J., Harper, J.W., 1997. F-box proteins are receptors that recruit phosphorylated substrates to the SCF ubiquitin-ligase complex. Cell 91, 209–219.

Skowyra, D., Koepp, D.M., Kamura, T., Conrad, M.N., Conaway, R.C., Conaway, J.W., Elledge, S.J., Harper, J.W., 1999. Reconstitution of G1 cyclin ubiquitination with complexes containing SCFGrr1 and Rbx1. Science 284, 662–665.

Sluimer, J., Distel, B., 2018. Regulating the human HECT E3 ligases. Cell. Mol. Life Sci. CMLS 75, 3121–3141.

Smit, J.J., Sixma, T.K., 2014. RBR E3-ligases at work. EMBO Rep. 15, 142–154.

Smith, D.M., Chang, S.-C., Park, S., Finley, D., Cheng, Y., Goldberg, A., 2007. Docking of the Proteasomal ATPases' C-termini in the 20S Proteasomes alpha Ring Opens the Gate for Substrate Entry. Mol. Cell 27, 731–744.

Snyder, D.A., Chen, Y., Denissova, N.G., Acton, T., Aramini, J.M., Ciano, M., Karlin, R., Liu, J., Manor, P., Rajan, P.A., Rossi, P., Swapna, G.V.T., Xiao, R., Rost, B., Hunt, J., Montelione, G.T., 2005. Comparisons of NMR spectral quality and success in crystallization demonstrate that NMR and X-ray crystallography are

complementary methods for small protein structure determination. J. Am. Chem. Soc. 127, 16505–16511.

Song, C., Wang, Q., Li, C.-C.H., 2003. ATPase activity of p97-valosin-containing protein (VCP). D2 mediates the major enzyme activity, and D1 contributes to the heat-induced activity. J. Biol. Chem. 278, 3648–3655.

Stach, L., Freemont, P.S., 2017. The AAA+ ATPase p97, a cellular multitool. Biochem. J. 474, 2953–2976.

Stadtmueller, B.M., Hill, C.P., 2011. Proteasome activators. Mol. Cell 41, 8–19.

Stadtmueller, B.M., Kish-Trier, E., Ferrell, K., Petersen, C.N., Robinson, H., Myszka, D.G., Eckert, D.M., Formosa, T., Hill, C.P., 2012. Structure of a proteasome Pba1-Pba2 complex: implications for proteasome assembly, activation, and biological function. J. Biol. Chem. 287, 37371–37382.

Stark, H., Zemlin, F., Boettcher, C., 1996. Electron radiation damage to protein crystals of bacteriorhodopsin at different temperatures. Ultramicroscopy 63, 75–79.

Stein, A., Ruggiano, A., Carvalho, P., Rapoport, T.A., 2014. Key steps in ERAD of luminal ER proteins reconstituted with purified components. Cell 158, 1375–1388.

Stevenson, J., Huang, E.Y., Olzmann, J.A., 2016. Endoplasmic Reticulum–Associated Degradation and Lipid Homeostasis. Annu. Rev. Nutr. 36, 511–542.

Stowell, J.A.W., Webster, M.W., Kögel, A., Wolf, J., Shelley, K.L., Passmore, L.A., 2016. Reconstitution of Targeted Deadenylation by the Ccr4-Not Complex and the YTH Domain Protein Mmi1. Cell Rep. 17, 1978–1989.

Strehl, B., Seifert, U., Krüger, E., Heink, S., Kuckelkorn, U., Kloetzel, P.-M., 2005. Interferon-γ, the functional plasticity of the ubiquitin–proteasome system, and MHC class I antigen processing. Immunol. Rev. 207, 19–30.

Strehl, B., Textoris-Taube, K., Jäkel, S., Voigt, A., Henklein, P., Steinhoff, U., Kloetzel, P.-M., Kuckelkorn, U., 2008. Antitopes Define Preferential Proteasomal Cleavage Site Usage. J. Biol. Chem. 283, 17891–17897.

Suzuki, R., Kawahara, H., 2016. UBQLN4 recognizes mislocalized transmembrane domain proteins and targets these to proteasomal degradation. EMBO Rep. 17, 842–857.

Suzuki, R., Moriishi, K., Fukuda, K., Shirakura, M., Ishii, K., Shoji, I., Wakita, T., Miyamura, T., Matsuura, Y., Suzuki, T., 2009. Proteasomal Turnover of Hepatitis C Virus Core Protein Is Regulated by Two Distinct Mechanisms: a Ubiquitin-

Dependent Mechanism and a Ubiquitin-Independent but PA28γ-Dependent Mechanism. J. Virol. 83, 2389–2392.

Tai, H.-C., Schuman, E.M., 2008. Ubiquitin, the proteasome and protein degradation in neuronal function and dysfunction. Nat. Rev. Neurosci. 9, 826–838.

Takeda, K., Tonthat, N.K., Glover, T., Xu, W., Koonin, E.V., Yanagida, M., Schumacher, M.A., 2011. Implications for proteasome nuclear localization revealed by the structure of the nuclear proteasome tether protein Cut8. Proc. Natl. Acad. Sci. 108, 16950–16955.

Tan, P., Fuchs, S.Y., Chen, A., Wu, K., Gomez, C., Ronai, Z., Pan, Z.Q., 1999. Recruitment of a ROC1-CUL1 ubiquitin ligase by Skp1 and HOS to catalyze the ubiquitination of I kappa B alpha. Mol. Cell 3, 527–533.

Tan, Y.Z., Baldwin, P.R., Davis, J.H., Williamson, J.R., Potter, C.S., Carragher, B., Lyumkis, D., 2017. Addressing preferred specimen orientation in single-particle cryo-EM through tilting. Nat. Methods 14, 793–796.

Tanaka, K., Waxman, L., Goldberg, A.L., 1983. ATP serves two distinct roles in protein degradation in reticulocytes, one requiring and one independent of ubiquitin. J. Cell Biol. 96, 1580–1585.

Tang, W.K., Odzorig, T., Jin, W., Xia, D., 2019. Structural Basis of p97 Inhibition by the Site-Selective Anticancer Compound CB-5083. Mol. Pharmacol. 95, 286–293.

Tasaki, T., Kim, S.T., Zakrzewska, A., Lee, B.E., Kang, M.J., Yoo, Y.D., Cha-Molstad, H.J., Hwang, J., Soung, N.K., Sung, K.S., Kim, S.-H., Nguyen, M.D., Sun, M., Yi, E.C., Kim, B.Y., Kwon, Y.T., 2013. UBR box N-recognin-4 (UBR4), an N-recognin of the N-end rule pathway, and its role in yolk sac vascular development and autophagy. Proc. Natl. Acad. Sci. 110, 3800–3805.

Tasaki, T., Kwon, Y.T., 2007. The mammalian N-end rule pathway: new insights into its components and physiological roles. Trends Biochem. Sci. 32, 520–528.

Tasaki, T., Mulder, L.C.F., Iwamatsu, A., Lee, M.J., Davydov, I.V., Varshavsky, A., Muesing, M., Kwon, Y.T., 2005. A family of mammalian E3 ubiquitin ligases that contain the UBR box motif and recognize N-degrons. Mol. Cell. Biol. 25, 7120–7136.

Tasaki, T., Zakrzewska, A., Dudgeon, D.D., Jiang, Y., Lazo, J.S., Kwon, Y.T., 2009. The substrate recognition domains of the N-end rule pathway. J. Biol. Chem. 284, 1884–1895.

Taylor, C., Jobin, C., 2005. Ubiquitin protein modification and signal transduction: implications for inflammatory bowel diseases. Inflamm. Bowel Dis. 11, 1097–1107.

Taylor, K.A., Glaeser, R.M., 1974. Electron diffraction of frozen, hydrated protein crystals. Science 186, 1036–1037.

The Nobel Prize in Chemistry 2017. NobelPrize.org. URL https://www.nobelprize.org/prizes/chemistry/2017/press-release/ (accessed 11.03.20).

Thon, F., 1966. Notizen: Zur Defokussierungsabhängigkeit des Phasenkontrastes bei der elektronenmikroskopischen Abbildung. Z. Für Naturforschung A 21, 476–478.

Thorpe, G.H., Kricka, L.J., 1986. Enhanced chemiluminescent reactions catalyzed by horseradish peroxidase. Methods Enzymol. 133, 331–353.

Thrower, J.S., Hoffman, L., Rechsteiner, M., Pickart, C.M., 2000. Recognition of the polyubiquitin proteolytic signal. EMBO J. 19, 94–102.

Tokmakov, A.A., Kurotani, A., Takagi, T., Toyama, M., Shirouzu, M., Fukami, Y., Yokoyama, S., 2012. Multiple Post-translational Modifications Affect Heterologous Protein Synthesis. J. Biol. Chem. 287, 27106–27116.

Tomko, R.J., Hochstrasser, M., 2013. Molecular Architecture and Assembly of the Eukaryotic Proteasome. Annu. Rev. Biochem. 82, 415–445.

Torrecilla, I., Oehler, J., Ramadan, K., 2017. The role of ubiquitin-dependent segregase p97 (VCP or Cdc48) in chromatin dynamics after DNA double strand breaks. Philos. Trans. R. Soc. B Biol. Sci. 372, 20160282.

Trempe, J.-F., Sauvé, V., Grenier, K., Seirafi, M., Tang, M.Y., Ménade, M., Al-Abdul-Wahid, S., Krett, J., Wong, K., Kozlov, G., Nagar, B., Fon, E.A., Gehring, K., 2013. Structure of parkin reveals mechanisms for ubiquitin ligase activation. Science 340, 1451–1455.

Unno, M., Mizushima, T., Morimoto, Y., Tomisugi, Y., Tanaka, K., Yasuoka, N., Tsukihara, T., 2002. The structure of the mammalian 20S proteasome at 2.75 A resolution. Struct. Lond. Engl. 1993 10, 609–618.

Unverdorben, P., Beck, F., Śledź, P., Schweitzer, A., Pfeifer, G., Plitzko, J.M., Baumeister, W., Förster, F., 2014. Deep classification of a large cryo-EM dataset defines the conformational landscape of the 26S proteasome. Proc. Natl. Acad. Sci. 111, 5544–5549.

Ustrell, V., Hoffman, L., Pratt, G., Rechsteiner, M., 2002. PA200, a nuclear proteasome activator involved in DNA repair. EMBO J. 21, 3516–3525.

van Brugggen, E., Wiebenga, E., Gruber, M., 1960. Negative-staining electron microscopy of proteins at pH values below their isolectric points. Its application to hemocyanin. Biochim. Biophys. Acta 42, 171–172.

van den Boom, J., Meyer, H., 2018. VCP/p97-Mediated Unfolding as a Principle in Protein Homeostasis and Signaling. Mol. Cell 69, 182–194.

van Heel, M., 1984. Multivariate statistical classification of noisy images (randomly oriented biological macromolecules). Ultramicroscopy 13, 165–183.

van Heel, M., Gowen, B., Matadeen, R., Orlova, E.V., Finn, R., Pape, T., Cohen, D., Stark, H., Schmidt, R., Schatz, M., Patwardhan, A., 2000. Single-particle electron cryo-microscopy: towards atomic resolution. Q. Rev. Biophys. 33, 307–369.

van Heel, M., Hollenberg, J., 1980. On the Stretching of Distorted Images of Two-Dimensional Crystals, in: Baumeister, W., Vogell, W. (Eds.), Electron Microscopy at Molecular Dimensions, Proceedings in Life Sciences. Springer, Berlin, Heidelberg, pp. 256–260.

van Heel, M., Schatz, M., 2005. Fourier shell correlation threshold criteria. J. Struct. Biol. 151, 250–262.

van Tijn, P., de Vrij, F.M.S., Schuurman, K.G., Dantuma, N.P., Fischer, D.F., van Leeuwen, F.W., Hol, E.M., 2007. Dose-dependent inhibition of proteasome activity by a mutant ubiquitin associated with neurodegenerative disease. J. Cell Sci. 120, 1615–1623.

Varshavsky, A., 2011. The N-end rule pathway and regulation by proteolysis. Protein Sci. Publ. Protein Soc. 20, 1298–1345.

Venkatraman, P., Wetzel, R., Tanaka, M., Nukina, N., Goldberg, A.L., 2004. Eukaryotic proteasomes cannot digest polyglutamine sequences and release them during degradation of polyglutamine-containing proteins. Mol. Cell 14, 95–104.

Verma, R., Aravind, L., Oania, R., McDonald, W.H., Yates, J.R., Koonin, E.V., Deshaies, R.J., 2002. Role of Rpn11 metalloprotease in deubiquitination and degradation by the 26S proteasome. Science 298, 611–615.

Verma, R., Chen, S., Feldman, R., Schieltz, D., Yates, J., Dohmen, J., 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., Cook, W.J., 1987. Structure of ubiquitin refined at 1.8 A resolution. J. Mol. Biol. 194, 531–544.

Vingill, S., Brockelt, D., Lancelin, C., Tatenhorst, L., Dontcheva, G., Preisinger, C., Schwedhelm-Domeyer, N., Joseph, S., Mitkovski, M., Goebbels, S., Nave, K., Schulz, J.B., Marquardt, T., Lingor, P., Stegmüller, J., 2016. Loss of FBXO7 (PARK15) results in reduced proteasome activity and models a parkinsonism-like phenotype in mice. EMBO J. 35, 2008–2025.

Voigt, A., Jäkel, S., Textoris-Taube, K., Keller, C., Drung, I., Szalay, G., Klingel, K., Henklein, P., Stangl, K., Kloetzel, P.M., Kuckelkorn, U., 2010. Generation of in silico predicted coxsackievirus B3-derived MHC class I epitopes by proteasomes. Amino Acids 39, 243–255.

Wade, R.H., 1992. A brief look at imaging and contrast transfer. Ultramicroscopy 46, 145–156.

Wagner, T., Merino, F., Stabrin, M., Moriya, T., Antoni, C., Apelbaum, A., Hagel, P., Sitsel, O., Raisch, T., Prumbaum, D., Quentin, D., Roderer, D., Tacke, S., Siebolds, B., Schubert, E., Shaikh, T.R., Lill, P., Gatsogiannis, C., Raunser, S., 2019. SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM. Commun. Biol. 2, 218.

Wang, H., Wang, L., Erdjument-Bromage, H., Vidal, M., Tempst, P., Jones, R.S., Zhang, Y., 2004. Role of histone H2A ubiquitination in Polycomb silencing. Nature 431, 873–878.

Wang, S., Munro, R.A., Shi, L., Kawamura, I., Okitsu, T., Wada, A., Kim, S.-Y., Jung, K.-H., Brown, L.S., Ladizhansky, V., 2013. Solid-state NMR spectroscopy structure determination of a lipid-embedded heptahelical membrane protein. Nat. Methods 10, 1007–1012.

Waxman, L., Fagan, J.M., Goldberg, A.L., 1987. Demonstration of two distinct high molecular weight proteases in rabbit reticulocytes, one of which degrades ubiquitin conjugates. J. Biol. Chem. 262, 2451–2457.

Weihl, C.C., Dalal, S., Pestronk, A., Hanson, P.I., 2006. Inclusion body myopathy-associated mutations in p97/VCP impair endoplasmic reticulum-associated degradation. Hum. Mol. Genet. 15, 189–199.

Wenzel, D.M., Klevit, R.E., 2012. Following Ariadne's thread: a new perspective on RBR ubiquitin ligases. BMC Biol. 10, 24.

Wenzel, D.M., Lissounov, A., Brzovic, P.S., Klevit, R.E., 2011. UBCH7 reactivity profile reveals parkin and HHARI to be RING/HECT hybrids. Nature 474, 105–108.

Whitby, F.G., Masters, E.I., Kramer, L., Knowlton, J.R., Yao, Y., Wang, C.C., Hill, C.P., 2000. Structural basis for the activation of 20S proteasomes by 11S regulators. Nature 408, 115–120.

Wilkinson, K.D., Urban, M.K., Haas, A.L., 1980. Ubiquitin is the ATP-dependent proteolysis factor I of rabbit reticulocytes. J. Biol. Chem. 255, 7529–7532.

Woodbury, R.G., Brown, J.P., Yeh, M.Y., Hellström, I., Hellström, K.E., 1980. Identification of a cell surface protein, p97, in human melanomas and certain other neoplasms. Proc. Natl. Acad. Sci. U. S. A. 77, 2183–2187.

Woodman, P.G., 2003. p97, a protein coping with multiple identities. J. Cell Sci. 116, 4283–4290.

Xie, S.C., Metcalfe, R.D., Hanssen, E., Yang, T., Gillett, D.L, Leis, A.P., Morton, C.J., Kuiper, M.J., Parker, M.W., Spillman, N.J., Wong, W., Tsu, C., Dick, L.R., Griffin, M.D., Tilley, L., 2019. The structure of the PA28-20S proteasome complex from Plasmodoium falciparum and implications for proteostasis. Nature Mico. 4, 1990- 2000.

Xie, Y., Varshavsky, A., 2000. Physical association of ubiquitin ligases and the 26S proteasome. Proc. Natl. Acad. Sci. 97, 2497–2502.

Xie, Y., Varshavsky, A., 1999. The E2-E3 interaction in the N-end rule pathway: the RING-H2 finger of E3 is required for the synthesis of multiubiquitin chain. EMBO J. 18, 6832–6844.

Yamamoto, S., Tomita, Y., Hoshida, Y., Nagano, H., Dono, K., Umeshita, K., Sakon, M., Ishikawa, O., Ohigashi, H., Nakamori, S., Monden, M., Aozasa, K., 2004. Increased expression of valosin-containing protein (p97) is associated with lymph node metastasis and prognosis of pancreatic ductal adenocarcinoma. Ann. Surg. Oncol. 11, 165–172.

Yamamoto, S., Tomita, Y., Uruno, T., Hoshida, Y., Qiu, Y., Iizuka, N., Nakamichi, I., Miyauchi, A., Aozasa, K., 2005. Increased expression of valosin-containing protein (p97) is correlated with disease recurrence in follicular thyroid cancer. Ann. Surg. Oncol. 12, 925–934.

Yang, B.-J., Han, X.-X., Yin, L.-L., Xing, M.-Q., Xu, Z.-H., Xue, H.-W., 2016. Arabidopsis PROTEASOME REGULATOR1 is required for auxin-mediated

suppression of proteasome activity and regulates auxin signalling. Nat. Commun. 7, 1–9.

Yang, Y., Früh, K., Ahn, K., Peterson, P.A., 1995. In vivo assembly of the proteasomal complexes, implications for antigen processing. J. Biol. Chem. 270, 27687–27694.

Yang, Y., Kitagaki, J., Dai, R.-M., Tsai, Y.C., Lorick, K.L., Ludwig, R.L., Pierre, S.A., Jensen, J.P., Davydov, I.V., Oberoi, P., Li, C.-C.H., Kenten, J.H., Beutler, J.A., Vousden, K.H., Weissman, A.M., 2007. Inhibitors of ubiquitin-activating enzyme (E1), a new class of potential cancer therapeutics. Cancer Res. 67, 9472–9481.

Yao, T., Song, L., Xu, W., DeMartino, G.N., Florens, L., Swanson, S.K., Washburn, M.P., Conaway, R.C., Conaway, J.W., Cohen, R.E., 2006. Proteasome recruitment and activation of the Uch37 deubiquitinating enzyme by Adrm1. Nat. Cell Biol. 8, 994–1002.

Yashiroda, H., Toda, Y., Otsu, S., Takagi, K., Mizushima, T., Murata, S., 2015. N-terminal α7 deletion of the proteasome 20S core particle substitutes for yeast PI31 function. Mol. Cell. Biol. 35, 141–152.

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

Ye, Y., Tang, W.K., Zhang, T., Xia, D., 2017. A Mighty "Protein Extractor" of the Cell: Structure and Function of the p97/CDC48 ATPase. Front. Mol. Biosci. 4.

Young, G., Hundt, N., Cole, D., Fineberg, A., Andrecka, J., Tyler, A., Olerinyova, A., Ansari, A., Marklund, E.G., Collier, M.P., Chandler, S.A., Tkachenko, O., Allen, J., Crispin, M., Billington, N., Takagi, Y., Sellers, J.R., Eichmann, C., Selenko, P., Frey, L., Riek, R., Galpin, M.R., Struwe, W.B., Benesch, J.L.P., Kukura, P., 2018. Quantitative mass imaging of single biological macromolecules. Science 360, 423–427.

Zachariae, W., Schwab, M., Nasmyth, K., Seufert, W., 1998. Control of cyclin ubiquitination by CDK-regulated binding of Hct1 to the anaphase promoting complex. Science 282, 1721–1724.

Zaiss, D.M.W., Standera, S., Holzhütter, H., Kloetzel, P.-M., Sijts, A.J.A.M., 1999. The proteasome inhibitor PI31 competes with PA28 for binding to 20S proteasomes. FEBS Lett. 457, 333–338.

Zaiss, D.M.W., Standera, S., Kloetzel, P.-M., Sijts, A.J.A.M., 2002. PI31 is a modulator of proteasome formation and antigen processing. Proc. Natl. Acad. Sci. 99, 14344–14349.

Zheng, J., Yang, X., Harrell, J.M., Ryzhikov, S., Shim, E.-H., Lykke-Andersen, K., Wei, N., Sun, H., Kobayashi, R., Zhang, H., 2002. CAND1 Binds to Unneddylated CUL1 and Regulates the Formation of SCF Ubiquitin E3 Ligase Complex. Mol. Cell 10, 1519–1526.

Zheng, N., Wang, P., Jeffrey, P.D., Pavletich, N.P., 2000. Structure of a c-Cbl-UbcH7 complex: RING domain function in ubiquitin-protein ligases. Cell 102, 533–539.

Zhou, H., Cao, F., Wang, Z., Yu, Z.-X., Nguyen, H.-P., Evans, J., Li, S.-H., Li, X.-J., 2003. Huntingtin forms toxic NH2-terminal fragment complexes that are promoted by the age-dependent decrease in proteasome activity. J. Cell Biol. 163, 109–118.