The Prokaryotic Origins of the Ubiquitin-Proteasome System

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Adrian Fuchs

aus Stuttgart

Tübingen

2017

Die vorliegende Arbeit wurde unter der Leitung von Herrn Prof. Dr. Andrei Lupas im Zeitraum von Februar 2013 bis Juli 2017 in der Abteilung Protein Evolution am Max-Planck-Institut für Entwicklungsbiologie in Tübingen angefertigt.

Die Betreuung dieser Arbeit in der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard-Karls-Universität Tübingen wurde von Herrn Prof. Dr. Thilo Stehle übernommen.

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	04.12.2017
Dekan:	Prof. Dr. Wolfgang Rosenstiel
1. Berichterstatter:	Prof. Dr. Andrei Lupas
2. Berichterstatter:	Prof. Dr. Thilo Stehle

Summary

Environmental influences demand a constant adaptation of the cell proteome. A key player in this process is the proteasome, a self-compartmentalizing complex that is responsible for the majority of targeted protein degradation via the ubiquitin (Ub) tagging system. This pathway was long believed to have originated in eukaryotes, while prokaryotes were seen to use a multitude of other less characterized systems instead. However, with the increasing number of sequenced genomes and more sophisticated bioinformatic analyses, evolutionary links and common concepts in prokaryotic and eukaryotic systems emerged. In this work, we illuminate these relationships by presenting a characterization of two novel bacterial proteasome homologs, Anbu and BPH, and of the first archaeal (de-) ubiquitination system.

Anbu (ancestral beta subunit) sequences are widely distributed amongst phylogenetically distant organisms and therefore its discovery meant that the broadly accepted concept, according to which the proteasome evolved from its only other known homolog, HslV, needed to be reviewed. In chapter 3.1, we unravel these relationships via cluster analysis and the first biophysical characterization of Anbu. Anbu forms a dodecameric complex with a unique architecture that was only accessible through the combination of X-ray crystallography and small-angle X-ray scattering. While forming continuous helices in crystals and electron microscopy preparations, refinement of sections from the crystal structure against the scattering data revealed a helical open-ring structure in solution, contrasting the ring-shaped structures of proteasome and HslV. Based on this primordial architecture and exhaustive sequence comparisons, we propose that Anbu represents an ancestral precursor at the origin of self-compartmentalization.

In this evolutionary scenario, BPH (beta-proteobacterial proteasome homolog) constitutes a comparably recent innovation that evolved by duplication of HslV. In chapter 3.2, we present the first analysis of BPH and reveal crystal structures of two representatives from of *Thiobacillus denitrificans* and *Cupriavidus metallidurans*. Remarkably, both proteins assemble in two stacked heptameric rings, unlike the hexameric rings seen in HslV. This unprecedented assembly is supported by SAXS

measurements, analytical ultracentrifugation and averaging of EM images. Its proteolytic activity is evidenced by mass spectrometry and a crystal structure with the proteasome specific inhibitor epoxomicin bound to the BPH active site. In contrast to HslV, BPH appears to act independent of AAA+ unfoldases and thus might pursue a function distinct from known bacterial self-compartmentalizing proteases, possibly in analogy to the proteasome's ATP-independent functions.

Finally (chapter 3.3), we study the first sequenced archaeal Ub tagging operon from *Caldiarchaeum subterraneum*, showing that protein ubiquitination is not restricted to eukaryotes. Via cluster analysis, we retrace the evolutionary history of this operon, revealing that its step-wise assembly was accompanied by co-evolution of increasingly Rpn11-like deubiquitinases, similar to those found in the eukaryotic proteasome. The *C. subterraneum* representative, CsRpn11, marks the apex of this process. CsRpn11 displays bimodal activity towards ubiquitin (CsUb) precursor (K_D = 14.6 µM; K_M = 24.2 µM) and isopeptide-linked ubiquitinated proteins. Our CsUb-CsRpn11 crystal structure reveals the first catalytically active deubiquitinase-Ub complex with the catalytic water oriented for cleavage. A comparison with the Ub-free structure shows a Ub-induced change between two conformations, that were previously associated with distinct binding modes in eukaryotic Rpn11 and its homolog AMSH. Consequently, the CsRpn11-CsUb interaction presents a useful model for homologous eukaryotic systems and supports an archaeal origin of protein ubiquitination.

Table of Contents

Summary 0

1 Introduction 6

1.1 Proteasome structure and function 7

1.2 Proteasome evolution and the discovery of Anbu and BPH 9

1.3 Evolution of the ubiquitin pathway 11

1.4 Ubiquitin-mediated protein degradation 14

1.5 Aims of this work 15

2 Experimental Procedures 16

2.1 General Methods 16

2.2 Methods for the characterization of Anbu 18

2.3 Methods for the characterization of BPH 24

2.4 Methods for the characterization of CsRpn11 28

3 Results and Discussion 35

3.1 The architecture of the Anbu complex reflects an evolutionary intermediate at the origin of the proteasome system 35

3.2 Structural characterization of the novel bacterial proteasome homolog BPH reveals a tetradecameric double-ring complex with unique inner cavity properties 55

3.3 Ubiquitin binding mode and cleavage mechanism of the Rpn11 protease in an ancestral archaeal ubiquitin conjugation system 70

4 Concluding remarks 90

5 Contributions 93

6 References 94

Zusammenfassung 109

Danksagung 111

1 Introduction

Living organisms are defined by an organized molecular structure. They metabolize chemicals to grow and reproduce, to adapt to environmental changes and to response to external stimuli (Koshland, 2002). As all these processes fundamentally rely on the varying interplay of thousands of proteins, a constant turnover within the cellular proteome is required and realized by synthesizing and degrading systems. One important regulator in this process is the proteasome, which degrades the great majority of proteins in eukaryotic cells via the ubiquitin (Ub) pathway (Hochstrasser, 2009; Kish-Trier and Hill, 2013). Though this complex is also found in archaea and some bacteria, it appears to recognize its substrates by different mechanisms and to adopt a less central function in these organisms (De Mot et al., 1999; Ruepp et al., 1998). Prokaryotes in general display far greater genetic diversity than eukaryotes and accordingly evolved a multitude of other, different proteolytic machineries and degradation signals (Schmidt et al., 2009). Consequently, our understanding of these pathways and mechanisms is comparably poor (Gur et al., 2011). As a way to handle the prokaryotic diversity, bioinformatic methods represent valuable tools that allow the analysis of protein relationships as well as the prediction of structure and function. By applying these methods, novel components that could complement the existing degradation pathways have been identified in prokaryotes: The proteasome homologs Anbu and BPH and the first archaeal (de-) ubiquitination system (Nunoura et al., 2011; Valas and Bourne, 2008). While being of potential physiological significance, the discovery of these proteins also sheds new light on the evolution of the ubiquitin-proteasome system (UPS) as this pathway was previously thought to be restricted to eukaryotes and the proteasome to have evolved from its only other known prokaryotic homolog, HslV (Bochtler et al., 1999; Cavalier-Smith, 2006; Gille et al., 2003). Furthermore, their analysis reveals evolutionary conserved mechanisms in simpler prokaryotic systems, that were inherited by the more complex machineries in modern eukaryotes.

1.1 Proteasome structure and function

The proteasome is a ubiquitous nano-machine for protein degradation in eukaryotes and archaea (Maupin-Furlow, 2012). It is a ~670 kDa barrel-shaped complex of four stacked rings (Figure 1; (Kish-Trier and Hill, 2013)), each composed of seven identical (archaea) or distinct (eukaryotes) subunits. The outer rings consist of catalytically inactive α subunits, whereas the inner rings are composed of proteolytic β subunits. α and β subunits are similar in sequence and structure, and are thought to have emerged by the duplication of a proto-β subunit.

The proteasome can act by itself as 20S proteasome (Pickering and Davies, 2012a), or its α-subunits may interact with various regulators that affect its choice of substrates (Forouzan et al., 2012; Fort et al., 2015). The proteasome's most prominent function, targeted protein degradation, requires interaction with hexameric unfoldases of the AAA+ (ATPase with diverse cellular functions) superfamily (Bar-Nun and Glickman, 2012), such as Rpt 1-6 within the eukaryotic 19S regulator or the archaeal VAT, which can also act as chaperones on their own (Benaroudj and Goldberg, 2000). In recent years, experiments have also emphasized the significance of the proteasome's ATP-independent functions (Ben-Nissan and Sharon, 2014), such as the degradation of oxidized proteins through interaction with PA28αβ (Pickering and Davies, 2012b) or the degradation of acetylated histones through interaction with PA200 (Qian et al., 2013).

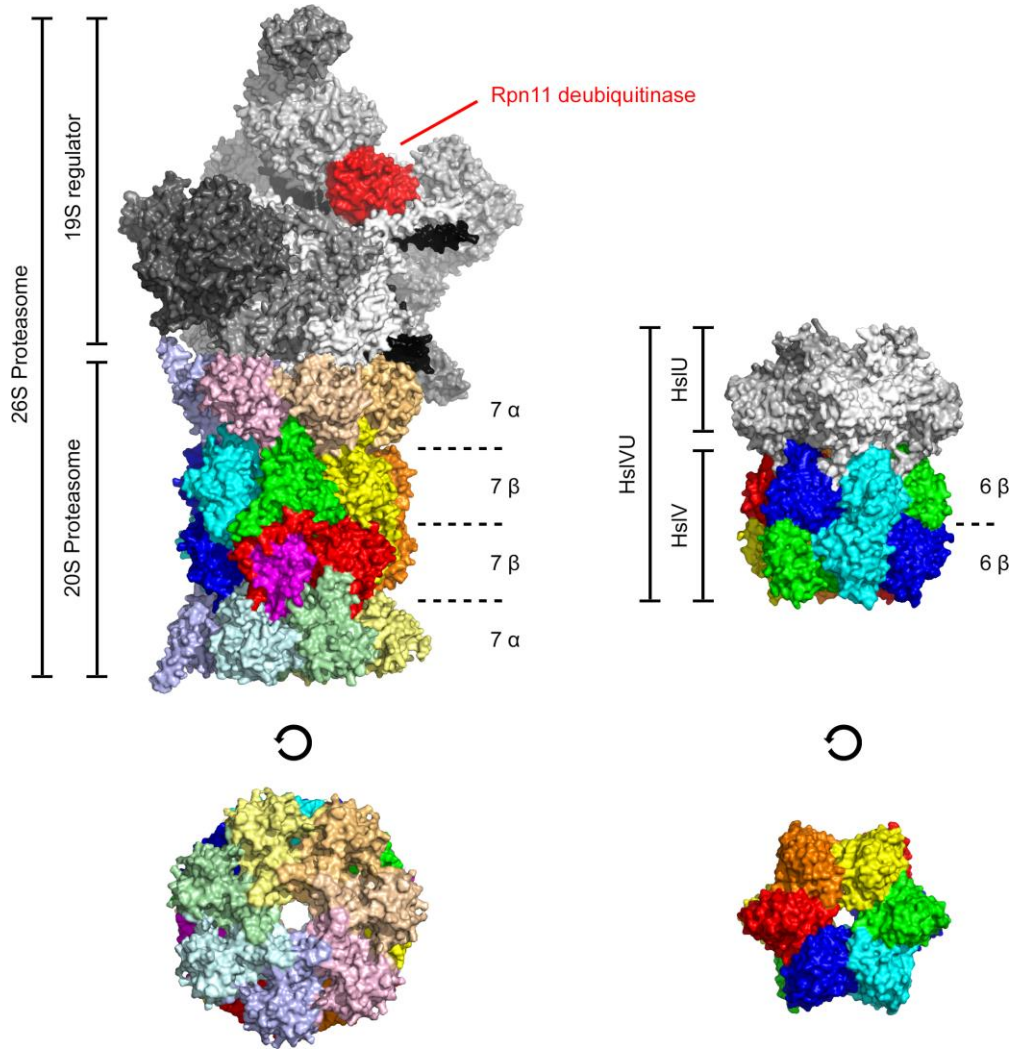


Figure 1: Structural comparison between proteasome and HslV

Shown are the architectures of the eukaryotic 26S proteasome (PDB 5WVI), an assembly composed of 20S proteasome and 19S regulator, and of the bacterial HslVU complex (PDB 1G3I), which includes the HslV protease and the HslU regulator. While the four stacked heptameric rings in the 20S proteasome are composed of seven distinct α and β subunits, HslV is a homo-oligomeric assembly β subunits, that arrange in just two hexameric rings. Furthermore, HslV appears to function only in complex with HslU (Rho et al., 2008b; Yoo et al., 1996), while the proteasome can act alone or through interaction with various regulators, such as 19S or PA200 (Fort et al., 2015). Bacterial and archaeal proteasomes are similar in architecture to those found in eukaryotes, but are typically composed of just 1-2 distinct α and β subunits and interact with a different set of regulators (Humbard and Maupin-Furlow, 2013).

1.2 Proteasome evolution and the discovery of Anbu and BPH

While the sophisticated eukaryotic proteasome is the epitome of self-compartmentalizing proteases, its evolution started from more simpler versions and, indeed, archaea have a more primordial version of the proteasome system (Campbell et al., 2015). Whereas in eukaryotes 14 different genes encode for the core α and β subunits, in archaea typically single α and β genes are responsible for the formation of homo-oligomeric rings that further assemble into the 20S proteasome. Interestingly, the proteasome is absent from bacteria, barring some branches of Actinobacteria (Lupas et al., 1994; Maupin-Furlow, 2012) and Nitrospirae (De Mot, 2007). While one theory attributes the occurrence of the proteasome in actinobacteria to horizontal gene transfer (HGT) (Gille et al., 2003), another argues that the original proteasome evolved in an ancestral actinobacterium, from where it was inherited by archaea and eukaryotes (Figure 2; (Cavalier-Smith, 2006)). Both theories, however, assume (Bochtler et al., 1999) that the proteasome as such evolved from its simpler and widely distributed bacterial homolog HslV (Heat-shock locus V). This homolog, unlike the proteasome, is a homomeric assembly of just two stacked hexameric rings (Bochtler et al., 1997) and thus lacks the antechamber constituted by the α-subunits (Figure 1). Despite this, HslV is similar to the proteasome in its ability to interact with an unfoldase of the AAA+ superfamily, HslU, which recognizes intrinsic features of misfolded proteins (Gur et al., 2011). HslU and proteasomal unfoldases, however, belong to different clades of AAA+ ATPases (Ammelburg et al., 2006) and use different interfaces for interaction with their respective proteolytic machinery (Sousa et al., 2000; Yu et al., 2010), suggesting that they were recruited independently. While several studies indicate that HslU is not always bound to HslV (Azim et al., 2005) and possesses chaperone-like activities on its own (Seong et al., 2000), HslV has not been shown to function in the cell on its own in an ATP-independent manner. Unlike the essential, constitutively expressed eukaryotic proteasome, the non-essential heat-shock induced HslV complements a set of other unrelated self-compartmentalizing proteases, such as FtsH, Lon and Clp, under stress conditions (Gur et al., 2011; Kanemori et al., 1997).

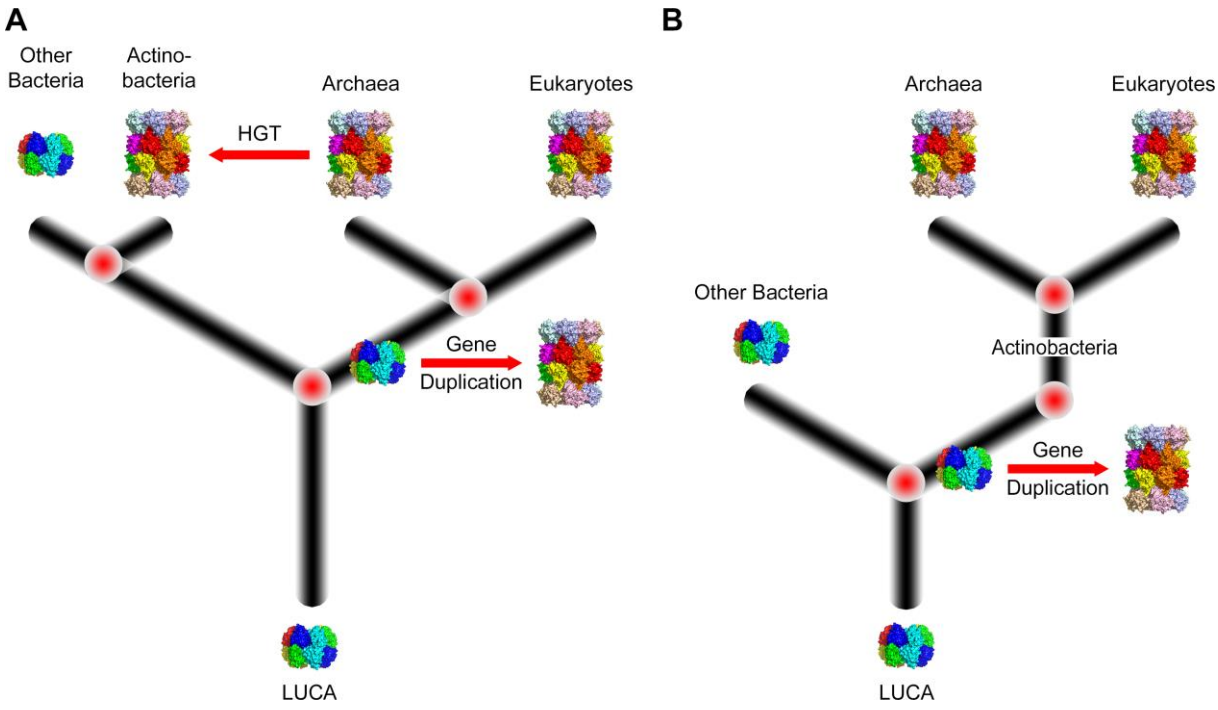


Figure 2: Proposed scenarios for the evolution of proteasome-like proteins

The β subunits of proteasome and HslV are clearly evolutionary related (Bochtler et al., 1999). HslV is found in nearly all bacterial phyla, except Actinobacteria, while the proteasome is present in all eukaryotes, archaea and most actinobacteria. This distribution was explained by two conflicting models. In both scenarios, HslV evolved first and gave rise to proteasomal α and β subunits through gene duplication. However, in one model (A), this event took place in the common ancestor of Archaea and Eukaryota and the proteasome was transferred to Actinobacteria via HGT (Gille et al., 2003), while in a second model (B), the proteasome evolved in Actinobacteria and was vertically transferred to the common ancestor of Archaea and Eukaryota (Cavalier-Smith, 2006).

A couple of years ago, bioinformatic analysis revealed the existence of two novel bacterial members of the proteasome family, Anbu (ancestral β subunit) and BPH (Valas and Bourne, 2008). It was proposed that Anbu, not HslV, gave rise to the proteasome, and that this event took place in actinobacteria. This interpretation was however questioned as Anbu, unlike other self-compartmentalizing proteases, is not associated with a AAA+ ATPase on the genomic level, but frequently co-occurs in an operon with a transglutaminase, an ATP-grasp protein with putative peptide ligase function, and a unique alpha-helical protein, Alpha-E (Iyer et al., 2009; Valas and

Bourne, 2008), hinting at a specific peptide synthesis system in which Anbu would act as a peptidase (Iyer et al., 2009).

The second hitherto unknown bacterial proteasome family member was primarily found in β-proteobacteria, hence its name BPH, short for β-proteobacterial proteasome homolog (Valas and Bourne, 2008). Based on this narrow phylogenetic distribution and maximum likelihood analysis, BPH was proposed to present a relatively young proteasome homolog, that evolved from HslV. Just like Anbu, BPH is not genomically associated with potential partner AAA+ unfoldases. Consequently, while previously only energy dependent self-compartmentalizing proteases were known in bacteria, BPH and Anbu might present a new class of ATP-independent machineries, that could adopt similar functions as the proteasome when associated with its various ATP-independent interactors.

1.3 Evolution of the ubiquitin pathway

Protein ubiquitination regulates an enormous range of physiological processes, often by acting as an adaptor in protein interactions. Substrates can be ubiquitinated via isopeptide bonds by a cascade of E1, E2 and E3 enzymes and are deubiquitinated by zinc-dependent JAMM (JAB1/MPN/Mov34) proteases, such as the proteasomal Rpn11 (Figure 1 and 3; (Hochstrasser, 2009)). Although classical ubiquitination is thought to be restricted to eukaryotes, several distinct but evolutionary related pathways are present in prokaryotes, which usually comprise just one Ub-, E1- and JAMM homolog each (Figure 3; (Burroughs et al., 2009)). These systems are often involved in cofactor synthesis, such as thiamine or molybdenum cofactor, but can also be bifunctional, as shown for the archaeon *Haloferax volcanii*. Here, the Ub-like proteins, SAMPs (small archaeal modifier protein), are also used in a similar fashion as eukaryotic Ub, to modify protein functions or to mark substrates for proteasomal degradation (Dantuluri et al., 2016; Fu et al., 2016). However, given the absence of E2 and E3 enzymes in *H. volcanii* and the evolutionary distance between SAMPs and Ub, it is unclear whether these common functions are evolutionary related or were acquired independently (Burroughs et al., 2012).

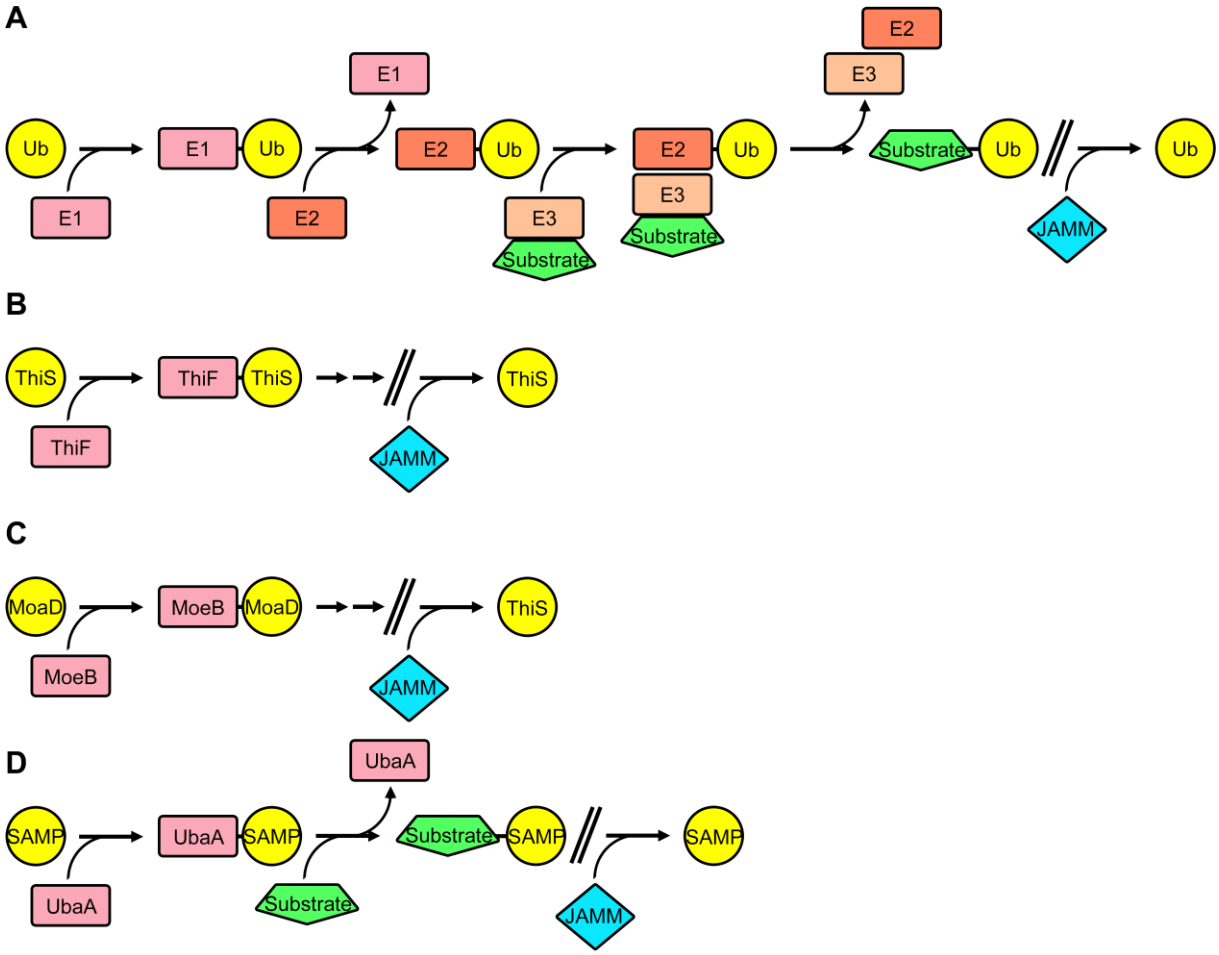


Figure 3: Schematic representation of the ubiquitination cascade and related pathways

(A) In eukaryotes, Ub is activated by ATP-dependent E1 enzymes, transferred to E2 carriers and conjugated via isopeptide bonds to target substrates, that are recognized by E3 proteins (Hochstrasser, 2009). Depending on the substrate fate, removal of the Ub-tag is catalyzed by various JAMM proteases, such as the proteasomal Rpn11 or the endosomal AMSH (Sato et al., 2008; Worden et al., 2014).

(B,C) Some prokaryotes encode for one or more pathways, in which Ub-like proteins are also activated by E1 homologs, but which typically function without E2- or E3-homologs. Examples are the first steps in the thiamine (B) and molybdopterin (C) synthesis pathways, that are followed by other, unrelated reactions (Burroughs et al., 2009). The respective genes are sometimes found in genomic context with JAMM proteases.

(D) In the archaeon *H. volcanii*, Ub-like proteins are used both in cofactor synthesis pathways and for modification of substrate proteins, that can regulate protein activity or mark substrates for proteasomal degradation (Cao et al., 2017; Fu et al., 2016).

The first complete archaeal Ub modifier system, including E2 and E3 enzymes, was discovered in 2011 in *Caldiarchaeum subterraneum* (Nunoura et al., 2011). Remarkably, all involved proteins were found to be far more similar to their eukaryotic counterparts than to any prokaryotic homologs. Thus, eukaryote-type protein ubiquitination might already have evolved in archaea. This interpretation was supported by the more recent discovery of Asgard archaea, which are thought to represent the archaeal origin of eukaryotes and whose genomes also harbor similar Ub modifier systems (Figure 4; (Zaremba-Niedzwiedzka et al., 2017)).

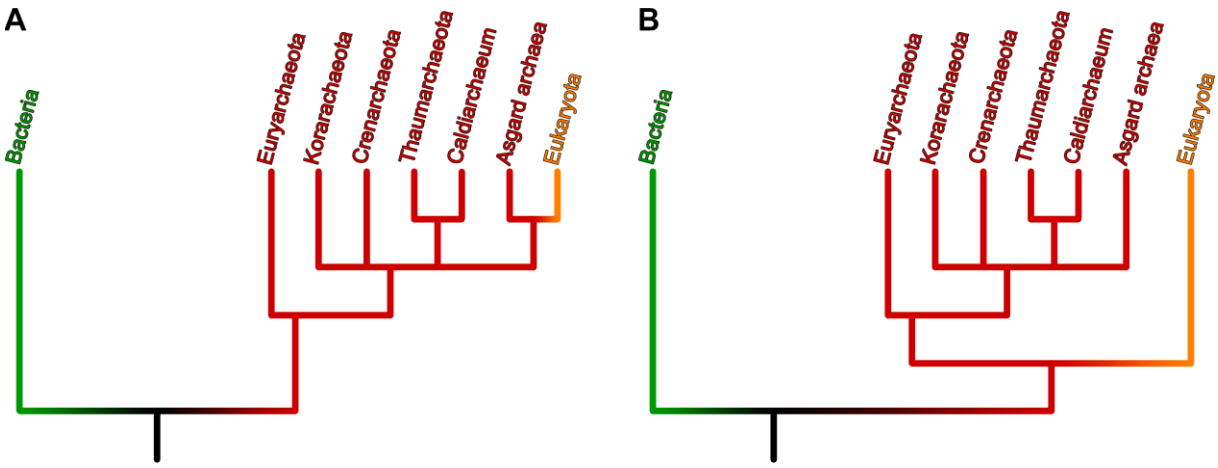


Figure 4: Two hypotheses on the origin of eukaryotes

(A) According to the "two-domain model", Eukaryota is one branch within Archaea (Williams et al., 2013). This idea is supported by the distribution of eukaryote-like genes in archaea, such as the recently identified Ub modifier systems in *C. subterraneum* and Asgard archaea. *C. subterraneum* has been proposed as member of the candidate phylum Aigarchaeota.

(B) The updated "three-domain model" as proposed by (Woese et al., 1990). Bacteria, Archaea and Eukaryota branch close to the root of the tree.

1.4 Ubiquitin-mediated protein degradation

While ubiquitination regulates a multitude of pathways and cellular processes via hundreds of interactors, its most prominent function is the targeting of substrates for proteasomal degradation (Hochstrasser, 2009). In eukaryotes, this process is realized by ubiquitin-receptors and the deubiquitinase Rpn11, which are situated within the 19S proteasomal unfoldase (Dambacher et al., 2016). As Ub resists unfolding, substrates need to be deubiquitinated prior degradation and consequently, Rpn11 activity is essential in eukaryotes (Verma et al., 2002).

Archaea lack most 19S components, with exception of the six ATPase modules Rpt1-6, that are present in form of the homomeric proteasome interactors PAN and VAT (Forouzan et al., 2012). Moreover, Rpn11-like proteases were identified within *C. subterraneum* and Asgard archaea (Nunoura et al., 2011; Zaremba-Niedzwiedzka et al., 2017). However, with the absence of other Rpn11-interacting 19S components, it is currently unclear whether this deubiquitinase is associated with the proteasome and how ubiquitinated proteins are recognized.

1.5 Aims of this work

With the discovery of the proteins and pathways described above, novel concepts in prokaryotic protein degradation and different scenarios for the origin of the ubiquitin-proteasome system can be anticipated. However, as their study is so far limited to preliminary sequence analysis, many emerging question remain unanswered: Is the concept of ATP-independent self-compartmentalizing proteases realized in bacteria through Anbu and BPH? What are the mechanistic principles and structural consequences of these functions? Is Anbu the proteasomal ancestor and was ubiquitin-mediated protein degradation already established in archaea? If so, how is ubiquitin recognized and removed with the 19S lid being absent from archaea? Through bioinformatic, structural and functional characterization, this study aims illuminate these aspects. In the following chapters, we show that the architecture of the Anbu complex reflects an evolutionary intermediate at the origin of the proteasome system. As a second unprecedented structure, we show that BPH assembles in heptameric rings with unique inner cavity properties, contrasting the hexameric rings formed by its ancestor HslV. Finally, we present a structural and mechanistic analysis of the Rpn11-ubiquitin interaction in the ancestral *C. subterraneum* ubiquitin conjugation system.

2 Experimental Procedures

2.1 General Methods

Bioinformatics

Structure based sequence alignments were generated with PROMALS3D, based on the structures of close homologs, and then manually refined. Structures were visualized using PyMOL v1.8.0.5, electrostatic potentials calculated at ± 5 kT/e with the APBS plugin and PDB2PQR) (Dolinsky et al., 2007) using default settings. DaliLite v3 (Holm et al., 2008) was used for calculating structural alignments and similarity scores.

Biophysical Methods

For electron microscopy (EM), glow-discharged carbon-coated grids were incubated with 0.1 mg/ml protein suspension, stained with 1% uranyl acetate and examined with a FEI Tecnai G2 Spirit BioTwin transmission EM at 120kV.

Thermal denaturation curves to monitor protein stability were recorded by circular dichroism spectroscopy at 220 nm using a JASCO J-810 spectropolarimeter.

Static light-scattering experiments were performed with 100 µl protein at concentrations of 0.5 mg/ml, 1 mg/ml or 2 mg/ml in 30 mM MOPS/NaOH (pH 7.2), 150 mM NaCl, using a Superdex S200 10/300 GL gel size-exclusion column (GE Healthcare) coupled to a miniDAWN Tristar Laser photometer (Wyatt) and a RI-2031 differential refractometer (JASCO). Data analysis was carried out with ASTRA V software (Wyatt).

For liquid chromatography mass spectrometry (LCMS) measurements, desalted samples were subjected to a Phenomenex Aeris Widepore 3.6 µm C4 200 Å (100 x 2.1 mm) column using an Agilent 1100 HPLC, eluted with a 30- 80% H_2O/acetonitrile gradient over 15 min at a flowrate of 0.25 ml/min in the presence of 0.05%

trifluoroacetic acid, and analyzed with a Bruker Daltonik microTOF. Eluted proteins were ionized at 4500 V and mass to charge (m/z) ratios determined in the range 800 – 3000. Data processing was performed in Compass DataAnalysis 4.2 and the m/z deconvoluted to obtain the protein mass via MaxEntropie.

Biochemical Methods

For Anbu and BPH, proteolytic activity was assayed in 20 mM Tris/HCl pH 7.5, 50 mM NaCl, supplemented with c0mplete protease inhibitor (Roche) without EDTA, which does not inhibit proteasome-like proteases. Enzyme (10 nM subunits when concentration was fixed) and fluorogenic substrates (50 µM each) were incubated at 30°C, and fluorescence changes recorded continuously for 2 h (Synergy H4 microplate reader, Biotek). Assay parameters were modified by varying pH (4.5 - 9.0), temperature (25 - 60 °C), salt (50 - 500 mM NaCl, 0 - 50 mM KCl, 0 - 5 mM $MgCl_2$/$CoCl_2$/$CaCl_2$) and enzyme concentrations (1 - 1000 nM). Assayed substrates included BODIPY-casein (Thermo), Ac-Gly-Pro-Leu-Asp-AMC, Z-Leu-Leu-Glu-AMC, Suc-Leu-Leu-Val-Tyr-AMC, Ac-Arg-Leu-Arg-AMC, Boc-Leu-Arg-Arg-AMC, Z-Gly-Gly-Leu-AMC (Enzo life sciences), Mca- Ala-Lys-Val-Tyr-Pro-Tyr-Pro-Met-Glu-Dap(Dnp) (GenScript), H-Val-AMC, H-Tyr-AMC, H-Thr-AMC, H-Pro-AMC, H-Phe-AMC, H-Asp-AMC, H-Ala-AMC, H-Ile-AMC (Bachem) and the P-check peptide library (Jena Bioscience).

2.2 Methods for the characterization of Anbu

Bioinformatics

To generate a clustermap of proteasome homologs, we searched the non-redundant protein sequence database at NCBI, comprising either bacterial, archaeal or eukaryotic proteins, employing four iterations of PSI-BLAST (Altschul and Koonin, 1998) at default settings. The following proteins were used as seeds for the first iteration of these searches: *P. aeruginosa* Anbu, *H. influenzae* HslV, *C. metallidurans* BPH, *T. acidophilum* and *M. tuberculosis* proteasome α and β. After each iteration, sequences to be included for the next iteration were manually reviewed. The sequences resulting from each of the searches were filtered down to a pairwise sequence identity of 90% using HHfilter in the MPI Bioinformatics Toolkit (Alva et al., 2016). The sequences in these reduced sets were next clustered by their all-against-all pairwise BLAST P-values in CLANS (Frickey and Lupas, 2004), to identify and remove incomplete or unrelated sequences. Sequences contained in the individual clusters of the resulting cluster maps were subsequently aligned using PROMALS3D (Pei et al., 2008), based on homologs with three-dimensional structures. The alignments were refined manually and propeptides as well as inserts of unusual lengths were removed. To further decrease redundancy, for the purpose of creating a global cluster map of proteasome homologs, we filtered these alignments down to a maximum pairwise identity of 70% using HHfilter. Next, we pooled together all sequences in these alignments and clustered them in CLANS to generate the cluster map of proteasome homologs shown in Figure 5A. Clustering was done to equilibrium in 2D at a BLAST P-value cutoff of 1e-20 and the final cluster map was made by showing all connections with a P-value better than 1e-09. The sequences in the 70% set were also used as input for HHalign (Soding, 2005) comparisons at default settings and to generate majority-rule consensus sequences with the consensus maker tool (http://www.hiv.lanl.gov). These consensus sequences were subsequently used for the manually refined PROMALS3D alignment shown in Figure 9.

The *C. hutchinsonii* Anbu-2 homology model was generated with MODELLER (Sali et al., 1995), using the Pa-Anbu structure as template.

Cloning, protein expression and purification

P. aeruginosa PAO1 and *C. hutchinsonii* genomic DNA were used to amplify and clone genes PA1733 and CHU_3460 into the pET22b expression vector using NdeI and HindIII restriction sites. The Pa-Anbu construct was the basis for "round-the-horn" (Hemsley et al., 1989) site-directed mutagenesis to obtain Pa-Anbu$^{A53C/N132C}$ for intersubunit crosslinking, a Pa-Anbu$^{\Delta 226-242}$ deletion mutant to investigate the stabilizing function of the C-terminus and a Pa-Anbu^{T1A} mutant to substantiate T1 as catalytic residue. It was also used to generate a Pa-Anbu$^{L94M/L112M/L228M}$ mutant for selenomethionine (Se-Met) labeling (Se-Met Pa-Anbu) via fusion-PCR. The Sumo-Pa-Anbu fusion was generated by cloning Pa-Anbu into a pET28b vector containing the *S. cerevisiae* sc288 Smt3p (Sumo) gene with AgeI and XhoI restriction sites. The Anbu-Con sequence was derived from an alignment of all available Anbu-1 sequences with a maximum sequence identity of 90%. The synthesized gene (Eurofins) was cloned into the pET30b vector using NdeI and HindIII restriction sites. The Cons-Anbu^{A53CN133C} mutant for intersubunit crosslinking was generated via site-directed mutagenesis. *E. coli* BL21 gold cells (Thermo) were transformed with the respective plasmids and grown at 25°C in M9 minimal medium supplemented with 50 µg/ml Se-Met, Leu, Ile, Phe, Thr, Lys and Val for Se-Met labeling, or in lysogeny broth (LB) for all other purposes. Protein expression was induced at an optical density of 0.4 at 600 nm with 0.5 mM isopropyl-β-D-thiogalactoside. Cells were harvested after 16 h, lysed by French press, and cleared from cell debris by ultracentrifugation. Soluble proteins were purified using two sequential anion-exchange columns, QHP (GE Healthcare) at pH 6.2 (20 mM MES/NaOH, 1 mM dithiothreitol (DTT), 0-500 mM NaCl) and MonoQ (GE Healthcare) at pH 8.8 (20 mM Tris/HCl, 1 mM DTT, 0 - 500 mM NaCl), followed by gel size-exclusion chromatography (Sephacryl S-300 HR, GE Healthcare) at pH 7.5 (20 mM Tris/HCl, 50 mM NaCl). Peak fractions containing the respective proteins were pooled, concentrated, supplemented with 15% glycerol, flash frozen in liquid nitrogen and stored at -80°C.

Crystallization, data collection and structure determination

For crystallization, native Pa-Anbu, the Se-Met derivative of Pa-Anbu$^{L94M/L112M/L228M}$ (Se-Met Pa-Anbu) and native Cons-Anbu were concentrated to 3.7, 11.2 and 20 mg/ml, respectively, in 50 mM NaCl, 20 mM Tris-HCl pH 7.5. Initial screening of conditions was performed in 96-well sitting-drop plates, with drops containing 300 nl of protein solution and 300 nl of reservoir solution, and a reservoir of 50 µl. For Se-Met Pa-Anbu and Cons-Anbu, crystallization conditions were further optimized in hanging drop setups, with drops containing 2.5 µl of protein solution and 2.5 µl of reservoir solution, and a reservoir of 500 µl in EasyXtal plates (Qiagen). The crystals used in the diffraction experiments grew within two days in 100 mM sodium acetate pH 4.4, 1.5 M sodium nitrate for Pa-Anbu, 100 mM sodium acetate pH 4.8, 1.2 M sodium nitrate for Se-Met Pa-Anbu, and 100 HEPES-NaOH pH 7.0, 600 mM NaF for Cons-Anbu. Prior to loop-mounting and flash-cooling in liquid nitrogen, the crystals were briefly transferred into a droplet of reservoir solution supplemented with either 25 % ethylene glycol (Pa-Anbu, Se-Met Pa-Anbu) or 3.5 M NaCl (Cons-Anbu) for cryo-protection. Data were collected at 100 K and a wavelength of either 1 Å (Pa-Anbu, Cons-Anbu) or 0.979 Å at the Selenium K-edge (Se-Met Pa-Anbu) at beamline X10SA of the Swiss Light Source (Villigen, Switzerland), using a PILATUS 6M hybrid pixel detector (Dectris Ltd.). All data were indexed, integrated and scaled using XDS (Kabsch, 2010), with the statistics given in Table 3.

For Se-Met Pa-Anbu, we employed SHELXD (Sheldrick, 2008) for heavy atom location. As we expected about three to four dozens of subunits with five selenium sites each, we decided to search for 220 heavy atoms. In the 14,344th trial, a solution with a CC All/Weak of 49.2 / 25.8 and CFOM of 75.0 was found, with a sharp drop in peak height after about 160 strong sites. After phasing and density modification with SHELXE, many secondary structure elements could be traced by ARP/WARP (Perrakis et al., 1999), and it became apparent that the asymmetric unit contains 34 subunits, arranged in a double-helix spanning the crystal. The structure was completed by cyclic chain tracing with (Buccaneer, (Cowtan, 2006)), manual modeling (Coot, (Emsley and Cowtan, 2004)), and refinement (REFMAC5, (Murshudov et al., 1999)) with strong NCS restraints. The native Pa-Anbu structure was subsequently solved on the basis of the Se-Met structure. As it did not reveal any conformational

differences, but was of lower resolution than the Se-Met structure, it was not regarded further; We therefore refer to the Se-Met Pa-Anbu structure as the Pa-Anbu structure.

The structure of Cons-Anbu was solved by molecular replacement with MOLREP (Vagin and Teplyakov, 2000), using two consecutive dimers of the Pa-Anbu structure as the search model, and the structural model completed and refined using Coot and REFMAC5. In the Cons-Anbu crystals, the double helix is built around the crystallographic 4_3 axis with 8 dimers per turn.

The structures have been deposited in the Protein Data Bank (PDB). Refinement statistics are given in Table 3, together with the PDB accession codes.

Small-Angle X-ray Scattering (SAXS)

SAXS experiments were performed at beamline B21, Diamond Light Source (Didcot, UK), with a X-ray wavelength of 1 Å and a PILATUS 2M detector at a distance of 3.9 m. A sample of 50 µl Pa-Anbu at a concentration of 19 mg/ml was delivered at 20 °C via an in-line Agilent HPLC with a Shodex Kw-403 column and a running buffer consisting of 20 mM TRIS pH 8.0, 150 mM NaCl, 2 mM TCEP and 1% sucrose. The continuously eluting sample was exposed for 300 s in 10 s acquisition blocks, and the data pre-processed using in-house software. Frames recorded immediately before the elution of the sample were used for buffer subtraction. Buffer subtraction and further analysis were performed with ScÅtter version 2.2b. According to the Kratky plot and a Porod exponent of 3.8, Pa-Anbu forms a well-folded particle with low intrinsic flexibility.

For fitting the SAXS profile, models of different geometries were prepared using a dimer from the Pa-Anbu crystal structure and the program pdbset of the CCP4 suite (Winn et al., 2011). Closed 6- or 7-rings of different radii were constructed by placing the dimer at different distances in 1 Å intervals to the origin and applying sixfold symmetry or 51.43° rotations. Helix sections with six or seven dimers were directly cut out of the Pa-Anbu crystal structure.

For the refinement of helical models, the translational components and the three Euler angles of the coordinate transformation from one dimer to the next in the crystal

structure were determined. Then, a series of helical models of six dimers was constructed by applying the same transformation five times using pdbset, by varying the translational components and individual Euler angles between different models in 1 Å and 1° steps. The fits of these models to the SAXS profile were assessed using the program FoXS (Schneidman-Duhovny et al., 2016), with the best model resulting in χ^2=2.83. For further refinement, this best-fitting model was subjected to rigid body modeling in sasref (Petoukhov and Svergun, 2005), with two loose contact restraints between each pair of dimers, i.e. ten restraints in total. Multiple such trials did not improve the fit further and did not change the helical nature of the complex.

Biophysical Methods

Melting curve and light scattering measurements were performed as described in the General Methods.

For electron microscopy (EM), disulfide crosslinks were generated by incubation of 140 µg/ml protein (5 µM subunits) with 1.5 mM Cu(II)-Phenanthroline for 5 min at room temperature, followed by addition of 100 mM EDTA. The proteins were then visualized as described in the General Methods.

For liquid chromatography mass spectrometry (LCMS) measurements, 1 mg/ml (38 µM subunits) Pa-Anbu, Pa-Anbu^{T1A} or Sumo-Anbu were incubated with either a 20x molar excess of epoxomicin or a 2x molar excess of MG132 for 48 h at 4°C (20 mM HEPES/NaOH pH 7.5, 150 mM NaCl). The masses of the Pa-Anbu protomers were then determined as described in the General Methods.

To validate the Pa-Anbu N-terminus as site of epoxomicin modification, the epoxomicin treated Pa-Anbu sample was also used for AspN digestion and subsequent LC-MS/MS analysis. Desalted samples were subjected to an EasyLC nano-HPLC (Proxeon Biosystems) coupled to an LTQ Orbitrap XL (Thermo Scientific) at a flow rate of 200 nl/min, and eluted with a segmented gradient of 8 - 26 - 40 - 64 - 72% H_2O/acetonitrile in the presence of 0.1% formic acid. Full scans were acquired at a resolution of 60,000. The target values were set to 5000 charges for the LTQ

(MS/MS) and 10^6 charges for the Orbitrap (MS), respectively. The MS data were processed with MaxQuant software suite v.1.5.2.8 (Cox and Mann, 2008). MS/MS spectra were searched with the Andromeda module (Cox et al., 2011) against a database consisting of 4,313 protein entries from *E. coli*, 285 commonly observed contaminants and Pa-Anbu. Protein N-terminal acetylation and methylation, methionine oxidation, and epoxomicin modification were set as variable modifications. Initial precursor mass tolerance was set to 4.5 parts per million (ppm), and at the fragment ion level 0.5 dalton (Da) was set for CID fragmentation. False discovery rates were estimated by the target/decoy approach (Elias and Gygi, 2007) and set to 1%. MS/MS spectra were obtained for both oxidized and non-oxidized versions of the modified N-terminal Pa-Anbu peptide (TYCVAM[ox]HLA). In these spectra, the epoxomicin moiety was represented as the peptide IITL with N-terminal acetylation and methylation, and the morpholino adduct between epoxomicin and Pa-Anbu Thr-1 as C_3H_6O.

Biochemical Methods

Native Western blots were performed using the Novex 4-16% Bis-Tris protein gels (Thermo) according to manufacturer's instructions. Pa-Anbu was detected using Pa-Anbu specific antibody raised in rabbit (Davids Biotechnology) and visualized with HRP-coupled goat anti-rabbit IgG (Sigma) and TMB (Thermo).

Anbu's proteolytic activity was investigated as described in the General Methods.

2.3 Methods for the characterization of BPH

Bioinformatics

The clustermap of proteasome homologs shown in Figure 15A was created as described in the methods for the characterization of Anbu.

Cloning, protein expression and purification

C. metallidurans CH34 (DSM 2839) and *T. denitrificans* AB7 (DSM 12475) genomic DNA were used to amplify and clone genes Rmet_1198 (Cm-BPH) and B059_01700 (corresponding to *T. denitrificans* ATCC 25259 gene Tbd_1847) (Td-BPH) into pET22b and pET30b expression vectors using NdeI and HindIII restriction sites. The constructs for the pET22b vector were extended via PCR with C-terminal TEV cleavage motifs followed by His_6-tags. The constructs for the pET30b vector were extended with C-terminal Strep-, Myc- and HA-tags or used without protein tag. Acidic pore loop mutants of Cm-BPH (E97A, D98G, E99A, D100G, D101A) and Td-BPH (E98A, D99G, K100A, E101G, D102A) were generated by site-directed mutagenesis (Hemsley et al., 1989). Wild-type and Trap Cm-BPH mutant (T1A, T2A, C3A) genes were also cloned into pBAD/Myc-His vectors, using XhoI and HindIII restriction sites. *E. coli* BL21 gold or KY2266 (Kanemori et al., 1997) cells (Thermo) were transformed with the respective plasmids and grown at 25°C in M9 minimal medium supplemented with 50 µg/ml Se-Met, Leu, Ile, Phe, Thr, Lys and Val for Se-Met labeling, or in lysogeny broth (LB) for all other purposes. Protein expression was induced at an optical density of 0.4 at 600 nm with 0.5 mM isopropyl-β-D-thiogalactoside (BL21 gold) or 1 % arabinose (pBAD-vector constructs in KY2266). Cells were harvested after 16 h, lysed by French press, and cleared from cell debris by ultracentrifugation. Soluble proteins were purified via two successive anion exchange steps with HiTrap Q HP (20 mM MES-NaOH pH 6.2, 1 mM DTT, 50 - 500 mM NaCl) and MonoQ HR 16/10 (20 mM Tris-HCl pH 8.5, 50 - 500 mM NaCl) columns or via HisTrap HP columns (20 mM Tris-HCl pH 8.0, 250 mM NaCl, 20 - 250 mM imidazole). His_6-tags were removed by incubation with 0.1 molar equivalents of His_6-TEV protease for 16 hrs at 4°C, followed

by another HisTrap HP purification step. Finally, all proteins were applied to a gel size-exclusion chromatography (Sephacryl S-300 HR, GE Healthcare) using buffer A (20 mM HEPES-NaOH, 100 mM NaCl). Purified proteins were supplemented with 15% glycerol, flash frozen in liquid nitrogen and stored at -80°C.

Crystallization, data collection and structure determination

For crystallization, Cm-BPH, Td-BPH or its Se-Met derivative were concentrated and set-up for crystallization as listed in Table 1. Initial screening of conditions was performed in 96-well sitting-drop plates, with drops containing equal volumes of protein and reservoir solution (250 - 400 nl), and a reservoir of 50 µl. For crystallization of epoxomicin-conjugated Td-BPH, hanging drops containing 2.5 µl protein solution with 1 mg/ml epoxomicin and 2.5 µl of reservoir solution were set-up over a reservoir of 500 µl in EasyXtal plates (Qiagen). The crystals used in the diffraction experiments usually grew within two days. Prior to loop-mounting and flash-cooling in liquid nitrogen, the crystals were briefly transferred into a droplet of reservoir solution supplemented with the indicated cryo-protectant (Table 1).

Table 1: Crystallization conditions for CsRpn11 and CsUb proteins.

Protein	Protein concentration [mg/ml]	Protein buffer	Reservoir solution	Cryo-protectant
Cm-BPH	8	10 mM Hepes-NaOH pH 7.5, 100 mM NaCl, 1 mM DTT	1M Na+-malonate pH 3.5	30% ethylene glycol
Td-BPH (Se-Met)	17.5	20 mM HEPES-NaOH pH 7.5, 50 mM NaCl	1.26 M NaH2PO4, 0.14 M KH2PO4	30% glycerol
Td-BPH (Epoxomin)	10	20 mM HEPES-NaOH pH 7.5, 50 mM NaCl, 1 mg/ml Epoxomicin (insoluble)	1.26 M NaH2PO4, 0.14 M KH2PO4	30% glycerol

Data were collected at beamline X10SA of the Swiss Light Source (Villigen, Switzerland) and the structures solved by Marcus Hartmann. The final models have not yet been deposited in the PDB.

Small-Angle X-ray Scattering (SAXS)

SAXS experiments were performed at beamline B21, Diamond Light Source (Didcot, UK), with a X-ray wavelength of 1 Å and a PILATUS 2M detector at a distance of 3.9 m. Samples of 50 µl Cm-BPH and Td-BPH at concentrations of 8.25 mg/ml and 15 mg/ml were delivered at 20 °C via an in-line Agilent HPLC with a Shodex Kw-403 column and a running buffer consisting of 20 mM HEPES-NaOH pH 8.0, 150 mM NaCl, 2 mM TCEP and 1% sucrose. The continuously eluting samples were exposed for 300 s in 10 s acquisition blocks, and the data pre-processed using in-house software. Frames recorded immediately before the elution of the sample were used for buffer subtraction. Buffer subtraction and further analysis were performed with ScÅtter version 2.2b. The crystallographic models of dodecameric Cm-BPH and tetradecameric Td-BPH were fit to both resulting SAXS profiles using the program FoXS (Schneidman-Duhovny et al., 2016).

Biophysical Methods

Melting curve and light scattering measurements were performed as described in the General Methods.

For averaging of BPH complexes in electron microscopy (EM) images (see General Methods), particles were selected manually, and image processing was carried out using EMAN2 (Tang et al., 2007).

For liquid chromatography mass spectrometry (LC-MS) measurements, 1 mg/ml Cm-BPH or Td-BPH were incubated with either a 20x molar excess of epoxomicin or a 2x molar excess of MG132 for 48 h at 4°C in buffer A. The masses of the BPH protomers were then determined as described in the General Methods.

Sedimentation velocity experiments were carried out in a BeckmanCoulter Optima XL-I Analytical Ultracentrifuge using both interference and absorbance detection. Experiments were performed with 0.5 mg/ml protein in buffer A at 20°C, an angular velocity of 40 krpm was applied. Samples were filled in titanium cuvettes with an optical pathlength of 20 mm. Solvent density and viscosity were calculated according to the buffer composition with Sednterp v.2. The partial specific volumes of Cm-BPH (0.729 mL/g) and Td-BPH (0.732 mL/g) were incrementally calculated from the amino acid sequence. Evaluation was carried out as global fitting to approximate solutions of the Lamm equation with Sedfit v.14.6. The frictional properties of the molecules in terms of the frictional ratio f/f0 were treated as floating parameters.

Biochemical Methods

To identify BPH interactors, C-terminally HA-, Myc- or Strep-tagged Cm-BPH proteins were produced in *E. coli* and bound to Streptavidin-, Anti-HA or Anti-Myc coated magnetic beads in buffer B (25mM HEPES-NaOH 7.2, 130mM NaCl, 10mM KCl, 0.1% NP40, 1x c0mplete protease inhibitor (Roche)). Beads were then incubated with *C. metallidurans* mid-log phase extracts at 4°C for 12 hrs (buffer B without NP40, but with DNAse and 1 mM $MgCl_2$). After six washes with buffer B, BPH and bound proteins were eluted with either 2 mg/ml HA- or Myc-peptides or 2.5 mM desthiobiotin and analyzed by SDS-PAGE followed by mass spectrometry.

The *in vivo* effect of BPH expression on H_2O_2 susceptibility was studied in the *E. coli* mutant strain KY2266 defective in cytosolic proteases (Δ*lon,* Δ*hslVU,* Δ*clpPX*). Log-phase cells expressing either wild type or trap mutant Cm-BPH were diluted to OD 0.1 and the medium (LB + 1% arabinose) supplemented with the indicated H_2O_2 concentrations. Growth curves were recorded for 4 hrs at 30°C and compared to those of KY2266 without plasmid and wild-type MC4100 cells.

The concentration of carbonyl groups in oxidized proteins was quantified photometrically with 2,4-dinitrophenylhydrazine (DNPH) (Levine et al., 1990). Oxidative stress was brought about by a prolonged 16 hrs stationary phase incubation at 30°C in the presence or absence of 10 µM MG132. The resulting lysate was

separated from cell debris via ultracentrifugation and incubated with 12.4 µM DNPH in the presence of 2 M HCl for 15 min at 25°C. Modified proteins were precipitated with 25% trichloroacetic acid, separated from residual DNPH by three successive washes and resuspended in 6 M guanidine, 100 mM Tris-HCl at pH 8.8. The relative concentration of modified carbonyl groups was quantified by measuring absorbance of samples at both 370 nm (DNPH) and 280 nm (Protein). The numbers shown in Figure 17 represent mean values determined in four independent experiments.

The proteolytic activity of BPH was investigated as described in the General Methods.

2.4 Methods for the characterization of CsRpn11

Bioinformatics

To gather sequences of Rpn11 homologs, we searched the non-redundant protein sequence database at NCBI, comprising either bacterial, archaeal or eukaryotic proteins, employing at least four iterations of PSI-BLAST (Altschul and Koonin, 1998) at default settings. JAMM core domains of the following proteins were used as seeds for the first iteration of these searches: *Homo sapiens* Rpn8 and Rpn11, *H. volcanii* JAMM1 and JAMM2 and *Frankia Alni* JAMM. The sequences in these sets were then clustered by their all-against-all pairwise BLAST P-values in CLANS (Frickey and Lupas, 2004), to identify individual JAMM protein types and their relative relationships. Based on this preliminary map, the ten most distinct sequences of each cluster were extracted and used in further rounds of iterative PSI-BLAST searches to identify all members of the respective protein type. After each iteration, sequences to be included for the next iteration were manually reviewed. Next, we pooled together all sequences, aligned them based on homologs with three-dimensional structures using PROMALS3D (Pei et al., 2008) and truncated all elements N- and C-terminal of the JAMM core domain. To further decrease redundancy, we filtered the sequences down to a maximum pairwise identity of 70% and clustered them in CLANS to generate the cluster map shown in Figure 22A. Clustering was done to equilibrium in 2D at a BLAST

P-value cutoff of 1e-15 and the final cluster map was made by showing all connections with a P-value better than 1e-10.

The Ub cluster map (Figure 22B) was generated using a similar approach, except that the initial PSI-BLAST search was seeded with *Homo sapiens* Ub, SUMO, UFM1 and Urm1 and *H. volcanii* as well as *E. coli* ThiS and MoaD. Clustering was done at a BLAST P-value cutoff of 9e-6 and the final cluster map was made by showing all connections with a P-value better than 1e-4.

Cloning, protein expression and purification

C. subterraneum genes Csub_C0080, Csub_C0800 (proteasomal AAA candidates), Csub_C702 and Csub_C1603 (SAMP candidates), Csub_C1473 (CsRpn11) and Csub_C1474 (CsUb) were synthesized (Eurofins) using codons optimized for expression in *E. coli*. These genes were then cloned into a modified pET30b expression vector using NdeI and HindIII restriction sites, so that the resulting proteins were produced with N-terminal His_6-tags and TEV restriction sites.

The CsRpn11 vector was the basis for "round-the-horn" (Hemsley et al., 1989) site-directed mutagenesis to obtain a series of CsRpn11 mutants with deletions in the sequence C-terminal of the JAMM domain, $CsRpn^{\Delta 175-202}$, $CsRpn^{\Delta 167-202}$, $CsRpn^{\Delta 149-202}$ and $CsRpn^{\Delta 145-202}$. In a similar fashion, the CsUb vector was used to generate vectors for expression of CsUb-mat, CsUb-mat C33A and CsUb-pre G77A, G78A, G77A/G78A, G77V, G78V, G78E, C79G, C79E, C79W, G80W and E81W mutant genes. Via fusion PCR, CsUb was ligated to either *Mus musculus* DHFR, mature *Bos taurus* Lactalbumin, GyrA intein, superfolder GFP (Pedelacq et al., 2006) or ssrA-linked superfolder GFP (Gerega et al., 2005) to produce His_6-CsUb-Lactalbumin, His_6-CsUb-DHFR, $CsUb^{C33A}$-GyrA, CsUb-GFP, GFP-CsUb, CsUb-GFP-ssrA and GFP-CsUb-ssrA fusions. An alternative variant of the CsUb-DHFR fusion contained a three amino acid linker, Cys-Gly-Glu, between both genes. In a similar manner, the Csub_C0080 and Csub_C0800 vectors were used to obtain truncated versions of both proteins with deletions in the AAA domains and hybrid proteins, where one or both

AAA domains were replaced with the respective domain of *T. acidophilum* VAT (Figure 30).

E. coli BL21 gold cells (Thermo) were transformed with the respective plasmids and grown at 25°C in in lysogeny broth (LB). Protein expression was induced with 0.5 mM isopropyl-β-D-thiogalactoside at an optical density of 0.4 at 600 nm. Cells were harvested after 16 h, lysed by French press, and subjected to ultracentrifugation. To purify insoluble proteins (all CsRpn11 variants, Lactalbumin fusions, VAT hybrids and VAT N-terminal domains), the ultracentrifugation pellet washed with buffer A (50 mM HEPES-NaOH pH 7.5, 250 mM NaCl, 10 mM DTT, 1 M Urea, 1% Triton X-100), resuspended in buffer B (50 mM Tris-HCl pH 8.0, 250 mM NaCl, 10 mM Imidazol, 8 M Urea) and purified using HisTrap HP columns (GE Healthcare). CsRpn11 variants, VAT hybrids and VAT N-terminal domains were then refolded via dialysis against buffer C (20 mM HEPES-NaOH pH 7.5, 500 mM NaCl) at room temperature and Lactalbumin fusions by dialysis against buffer D (20 mM HEPES-NaOH pH 7.5, 150 mM NaCl, 2.5 mM glutathione (GSH), 0.5 mM glutathione disulfide (GSSG), 1 mM $CaCl_2$). Soluble proteins were purified via two successive anion exchange steps with self-packed HiTrap Q HP and MonoQ HR 16/10 columns (GFP fusion proteins) or via HisTrap HP columns (all other proteins; GE Healthcare). GFP-fusions, CsUb-DHFR (S75), $CsUb^{C33A}$-GyrA and the N-terminal domains of Csub_C0080 and Csub_C0800 were subsequently separated via Superdex 75 10/300 GL columns and VAT hybrid proteins via Superdex 300 10/300 GL columns (GE Healthcare) using buffer C. Wild-type and mutant CsUb proteins were not purified via gel filtration and instead dialyzed against buffer C and separated from heat-sensitive proteins by a 10 min incubation at 70°C. *T. acidophilum* VAT and 20S proteasomes were cloned and purified as described in (Forouzan et al., 2012).

Crystallization, data collection and structure determination

CsUb and CsRpn11 proteins were set up for crystallization in 96-well sitting-drop plates, with drops containing 250 nl of reservoir solution and 250 nl of protein solution over a reservoir of 50 µl (Table 2). Prior to loop-mounting and flash-cooling in liquid

nitrogen, the crystals were briefly transferred into a droplet of reservoir solution supplemented with cryoprotectant. Data were collected at beamline X10SA of the Swiss Light Source (Villigen, Switzerland) and the structures solved by Marcus Hartmann. The final models have not yet been deposited in the PDB.

Table 2: Crystallization conditions for CsRpn11 and CsUb proteins.

Protein	Protein concentration [mg/ml]	Protein buffer	Reservoir solution	Cryo-protectant
Holo-CsRpn11	12	20 mM Tris-HCl pH 8.0, 50 mM NaCl, 1 M Urea	100 mM HEPES-NaOH pH 7.0, 15% PEG 4000	20% glycerin
Apo-CsRpn11$^{\Delta 149\text{-}202}$	8.5	20 mM HEPES-NaOH pH 7.5, 250 mM NaCl, 1 M Urea, 2 mM EDTA	100 mM sodium acetate pH 4.6, 1.5 M ammonium sulfate, 25% PEG 4000	30% glycerin
Apo-CsRpn11$^{\Delta 149\text{-}202}$-CsUb	6.07 CsRpn11 + 3.33 CsUb	20 mM HEPES-NaOH pH 7.5, 150 mM NaCl	100 mM sodium acetate pH 4.6, 1.5 M ammonium sulfate	15% PEG200
Holo-CsRpn11$^{\Delta 149\text{-}202}$-CsUb	Apo-CsRpn11$^{\Delta 149\text{-}202}$-CsUb crystals were soaked for 24 h in reservoir solution supplemented with 5 mM $ZnCl_2$			30% ethylene glycol

Biophysical methods

Melting curves, light scattering signals and electron micrographs were recorded as described in the General Methods.

For liquid chromatography mass spectrometry (LCMS) measurements (Figure 26A), 10 µM CsUb-pre were incubated with 1 µM CsRpn11 in buffer C for 20 min at 45°C and compared to an untreated CsUb-pre sample. The masses were then determined as described in the General Methods.

For MST-measurements, purified CsRpn11$^{\Delta 149\text{-}202}$ was labeled with fluorophore using the NT-647-NHS labeling kit (Nanotemper). Next, a serial 1:1 dilution of CsUb raging from 47 nM to 770 µM was prepared and mixed with 50 nM labeled CsRpn11$^{\Delta 149\text{-}202}$. MST measurements were performed at a Monolith NT.115 (Nanotemper), using various MST power and laser intensity settings to test the general validity of the obtained data. The results shown in Figure 26D represent mean values obtained in three independent experiments and were measured at a temperature of 45°C, using MST power 80% and laser intensity 40%. The shown binding curve was fitted to the data, using the NT Analysis 1.5.41 software (Nanotemper).

Biochemical methods

CsRpn11 activity towards CsUb-pre, Csub_C0702 (ThiS) and Csub_C1603 (MoaD) and eukaryotic diubiquitin (Enzo Life Sciences) was assayed by incubating 0.2 mg/ml Ub-like substrate with 0.2 mg/ml CsRpn11$^{\Delta 149\text{-}202}$ for 10 min at 37°C in buffer C (Figure 27).

To investigate its metal binding characteristics (Figure 28C), 50 µM CsRpn11$^{\Delta 149\text{-}202}$ were incubated with 2 mM EDTA, dialyzed two times against buffer C and a third time against buffer C supplemented with either 1 µM $CoCl_2$ or $ZnCl_2$. The proteins were then concentrated via centrifugal filters (Amicon Ultra-0.5 mL 3 kDa MWCO), incubated for 10 min with 6 M guanidine and supplemented with 100 µM 4-(2-pyridylazo)resorcinol (PAR). Absorption spectra were recorded with a Synergy H4 microplate reader (Biotek) and subtracted from spectra of the respective buffers.

To analyze the effect of the active site metal ion on catalysis (Figure 28A), CsRpn11$^{\Delta 149\text{-}202}$ was incubated with an excess of EDTA (2 mM) and dialyzed against buffer C. The resulting catalytically inactive Apo-CsRpn11$^{\Delta 149\text{-}202}$ sample contained less than 1 µM EDTA to avoid trace amounts of metal within the buffer from binding. 1.75 µM Apo-CsRpn11$^{\Delta 149\text{-}202}$ was then supplemented with 40 µM of either $CuCl_2$, $NiCl_2$, $CoCl_2$, $FeCl_2$ or $MnCl_2$, incubated with 14 µM CsUb-pre at 45°C for 5 min. The reaction was stopped with 1 mM EDTA and cleavage efficiency of the samples compared via SDS-PAGE.

For kinetic measurements, the indicated CsUb-pre concentrations were incubated for 21 min at 45°C in buffer C with 400 nM Zn^{2+}-bound CsRpn11, 400 nM Zn2+-bound CsRpn11$^{\Delta 149-202}$ or 76 nM Co^{2+}-bound CsRpn11$^{\Delta 149-202}$. Cleavage was stopped with 1 mM EDTA and the samples loaded on 12% Nu-PAGE Bis-Tris gels (Thermo). The gels were stained with Coomassie G-250 and the resulting band intensity quantified using ImageJ. The results shown in Figures 26C and 28B represent mean values of three independent experiments. The binding curve was fitted to the data using the SigmaPlot 12.3 enzyme kinetics tool with the Michaelis-Menten model.

The influence of amino acid substitutions in the CsUb C-terminus (Figure 29A) was assayed by incubating 14 µM of the respective CsUb-pre mutants with 1.75 µM zinc-bound CsRpn11$^{\Delta 149-202}$ in buffer C at 45°C for 8 min. The reactions were stopped with 1 mM EDTA and cleavage efficiency of the samples compared via SDS-PAGE.

The effect of the substrate folding state was assayed with CsUb-Lactalbumin and CsUb-DHFR fusions (Figures 29B and C). 6.15 µM CsUb-DHFR fusions were digested for 5 min with 14 µM CsRpn11$^{\Delta 149-202}$ at 45°C in presence or absence of 444 µM methotrexate and NADPH. The reaction was stopped with 1 mM EDTA and soluble and insoluble fractions separated via centrifugation at 20000 g for 3 min. CsUb-Apo-Lactalbumin was obtained by reducing the refolded Holo-form (see protein purification) with 10 mM TCEP and 10 mM EGTA, followed by dialysis against buffer E (20 mM HEPES, 150 mM NaCl, 5 mM TCEP, 0.5 % Chelex 100 resin (Biorad)). 12 µM of both CsUb-Lactalbumin forms or 14 µM of CsUb-pre were then incubated with either proteinase K or CsRpn11$^{\Delta 149-202}$ under oxidizing (buffer D) or reducing (buffer E) conditions. Proteinase K digests were performed with 0.1 µg/ml enzyme for 30 min at 10°C and stopped with 4 mM PMSF. CsRpn11$^{\Delta 149-202}$ were digests performed at 45°C for 17 min with 14 µM enzyme in case of CsUb-Lactalbumin and for 8 min with 1.75 µM enzyme in case of CsUb.

To obtain isopeptide-linked TEV-protease (Figure 29D), CsUb^{C33A}-GyrA was treated with cysteamine and allowed to react with purified TEV protease as described in (Chatterjee et al., 2010). 0.25 mg/ml of the resulting reaction products were then incubated with 3.33 µM CsRpn11 in buffer C at 37°C and the reaction stopped at the indicated times with 1 mM EDTA.

ATPase activity was measured by photometric quantification of the complex formation between inorganic phosphate and malachite green. To this end, 0.5 µM ATPase were incubated for 5 min at 37°C in buffer F (20 mM HEPES-NaOH pH 7.5, 300 mM NaCl, 120 mM $MgCl_2$) in presence of of 1.9 mM ATP. The reaction was stopped with an excess of EDTA, complex formation initiated with malachite green solution (0.034% malachite green, 1.05% ammonium molybdate, 0.1% Triton X100, 1 M HCl). After 10 min, citric acid was added to final concentration of 4% and absorbance of the samples was determined at 640 nM (Figure 30B).

To investigate the influence of CsUb on VAT mediated unfolding or proteasome mediated degradation, we measured the fluorescence change of CsUb-GFP fusions (Figure 30, D-J). To this end, 350 nM of the indicated substrates were incubated with 250 nM *T. acidophilum* proteasome (20S) alone or 250 nM *T. acidophilum* 20S-VAT in buffer F at 45°C. Fluorescence was excited at 488 nM and measured at 510 nM using a Synergy H4 microplate reader (Biotek). For analysis, values in presence of 25 mM ATP were compared to negative controls without ATP.

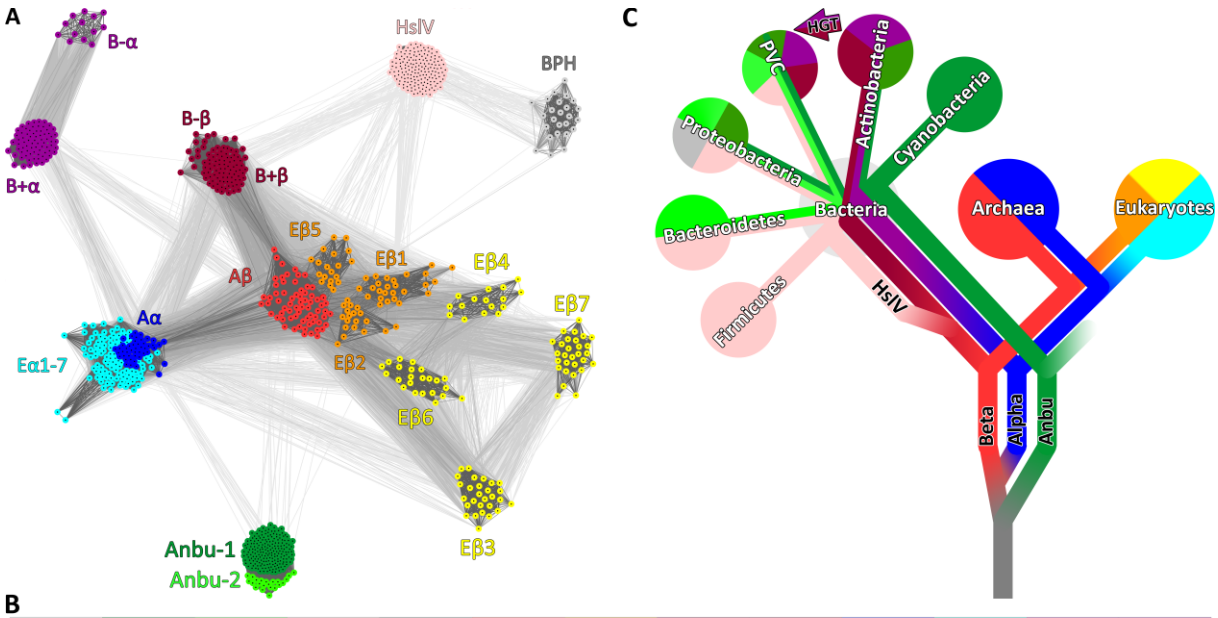
3 Results and Discussion

3.1 The architecture of the Anbu complex reflects an evolutionary intermediate at the origin of the proteasome system

In this chapter, we study the molecular characteristics of Anbu and, based on bioinformatic analysis, its place in proteasomal evolution. We determined crystal structures of two Anbu proteins and could decipher the Anbu structure in solution via small-angle X-ray scattering (SAXS). We find that Anbu forms a dodecameric open-ring assembly that locally resembles the architecture of the self-compartmentalizing proteasome, but is not closed for steric reasons. Based on these findings, we draft a scenario in which the Anbu complex constitutes an evolutionary intermediate at the origin of the proteasome system.

Anbu Dates Back to the Last Universal Common Ancestor

To investigate the evolutionary history of Anbu in the context of proteasome evolution, we searched for its homologs in the nonredundant protein sequence database and clustered them by pairwise sequence similarity. In the resulting map, the archaeal α- and β-subunits form two distinct, but tightly connected, central clusters, which exhibit the highest sequence similarity of all α-β pairs and from which all the other groups radiate (Figure 5A). Also, α-subunits of all kingdoms are closer in sequence space to the archaeal β than to any other β sequences and likewise all β-subunits are more similar to archaeal α than to any other α sequences (Figure 5B), suggesting that the archaeal proteasome is closest to the original form of the proteasome.



B

E-value	Anbu-1	Anbu-2	HslV	BPH	Aβ(E)	Eβ5	B+ β	B- β	Aα(E)	Eα5	B+ α	B- α
Anbu-1	X	6E-58	1E-01	3E-02	8E-09	1E-05	6E-04	6E-03	4E-09	3E-05	4E-03	1E-01
Anbu-2	3E-58	X	2E-03	7E-01	4E-06	5E-07	1E-03	2E-03	1E-05	1E-03	1E-01	3E-01
HslV	2E-01	1E-03	X	2E-22	1E-14	1E-08	1E-13	7E-09	3E-05	3E-04	2E-01	9E-01
BPH	5E-02	4E-02	7E-23	X	2E-10	8E-09	4E-11	9E-06	3E-05	1E-03	3E-02	5E-01
Aβ(E)	7E-09	3E-06	1E-13	1E-10	X	2E-53	1E-41	3E-27	3E-41	2E-33	1E-10	3E-06
Eβ5	3E-07	2E-08	9E-09	7E-10	7E-54	X	2E-29	5E-22	5E-24	7E-22	1E-03	1E-02
B+ β	4E-03	1E-03	3E-13	1E-10	5E-41	2E-28	X	1E-57	5E-20	2E-15	4E-12	4E-05
B- β	4E-02	9E-04	2E-08	6E-06	2E-27	1E-21	7E-57	X	7E-13	9E-11	4E-03	1E-04
Aα(E)	2E-09	3E-06	7E-04	6E-05	6E-41	2E-23	5E-19	6E-12	X	3E-86	1E-14	5E-09
Eα5	1E-06	5E-04	4E-03	5E-03	4E-33	1E-21	1E-14	2E-10	2E-85	X	7E-12	1E-07
B+ α	1E-02	1E-01	4E-01	4E-02	1E-11	7E-04	2E-06	1E-02	3E-16	5E-13	X	2E-30
B- α	1E-01	7E-02	6E-01	7E-01	4E-07	3E-02	5E-06	3E-04	2E-10	1E-08	5E-30	X

Figure 5: Cluster map and evolutionary scenario for the origin of proteasome-like proteins

(A) A cluster map of 969 proteasome-like sequences, with a maximum pairwise identity of 70%, was prepared using CLANS based on their all-against-all pairwise similarities as measured by BLAST P-values. Sequences are represented by dots and the line coloring reflects BLAST P-values; the darker a line, the lower the P-value. Proteasome subunits are abbreviated E for eukaryotic, A for archaeal and B+/B- for gram-positive/negative bacteria.

(B) Pairwise similarity of structure-guided sequence alignments. Alignments of selected clusters in (A) were compared and their pairwise similarity expressed as HHalign P-values (Soding, 2005). HHAlign uses HMM-HMM (hidden Markov model) comparisons to provide a statistical measure for the relationship between alignments of the indicated proteasome homologs. Thus, this analysis complements the cluster map, where individual sequences are compared.

(C) A schematic simplified tree of life (see also Figure 4) showing the phylogenetic distribution of proteasome-like proteins and our proposed scenario for their evolution in the same colors as in (A). Anbu may have shared a direct ancestry with the proto-β-subunit in the last universal common ancestor (LUCA). A subsequent duplication of the proto-β-subunit resulted in the

emergence of the proteasome, which was then linearly inherited by actinobacteria, archaea and eukaryotes. HslV evolved by gene duplication of proteasome β at the root of bacteria and replaced the proteasome in many bacterial phyla. Gram-negative bacteria acquired the proteasome by horizontal gene transfer (HGT) from actinobacteria, as previously proposed (De Mot, 2007). Anbu was lost in the common ancestor of eukaryotes and archaea, but was inherited by most bacteria, where it diverged into two subtypes with presumably different substrate specificity (Anbu-1, Anbu-2). Notably, the other proteins of the bacterial Anbu operon are still found sparsely distributed among archaea. The PVC-superphylum comprises Planctomycetes, Verrucomicrobia and Chlamydiae.

The closest neighbors of these core clusters are the eukaryotic α- and β-subunits, which are also connected to each other tightly. While the seven α types group together with the archaeal α-subunits in a large cluster, the seven β types are more divergent and form distinct subclusters, of which the catalytic ones (β1, β2, and β5) lie closer to the archaeal β-subunits, whereas the non-catalytic ones are further removed.

By comparison to their archaeal and eukaryotic counterparts, the bacterial α- and β-subunits only show residual sequence similarity to each other and are organized into two separate subclusters each (Figure 5A and B). While one comprises subunits of Gram-positive actinobacteria and exhibits statistically significant sequence connections to the archaeal core clusters, the other radiates from it and contains subunits of the Gram-negative armatimonadetes, nitrospirae and verrucomicrobia. This relationship is consistent with the suggestion (De Mot, 2007) that the Gram-negative proteasome may have been acquired by HGT from actinobacteria.

BPH is found at the periphery of the map and forms strong connections to HslV. Since HslV is more similar to any other analyzed homolog (Figure 5A and B), BPH might have emerged by duplication of HslV and consequently appears to play no role in the evolution of the proteasome (see chapter 3.2). The other bacterial proteasome homologs, Anbu and HslV, form two unconnected clusters. HslV sequences group into a single cluster, comprising sequences from almost all bacterial phyla, barring cyano- and actinobacteria. In contrast, Anbu is organized into two closely related subclusters: One comprising sequences from phylogenetically diverse bacteria, mainly from α-/β-/γ-proteo-, cyano- and actinobacteria (in the following referred to as Anbu-1), and the other (Anbu-2) comprising sequences with a more narrow phylogenetic spectrum

(planctomycetes, bacteroidetes, verrucomicrobia and δ-proteobacteria). Curiously, both Anbu subtypes sometimes co-occur in the same phyla or even in the same organism, e.g. in the α-proteobacterium *Azospirillum brasilense,* indicating that their sequence divergence possibly reflects functional diversification. Furthermore, the distribution of Anbu and HslV in deep-branching bacteria suggests that they were already present in an early bacterium.

As Anbu and HslV are more similar to proteasome subunits of all kingdoms than to each other (Figure 5A and B), their evolution is likely not directly linked to each other, but to the proteasome. Notably, HslV exhibits equally significant similarity to the actinobacterial β and to the archaeal β, but only residual similarity to α-subunits, suggesting that it arose from a β subunit after the diversification of α and β subunits, very early in bacterial evolution. In contrast, Anbu exhibits only weak similarity to the bacterial α and β subunits, but equally strong similarity to both subunits of the archetypal archaeal proteasome (Figure 5A and B). Anbu is thus likely to have either shared a direct ancestry with the proto-β-subunit or may even have given rise to it.

These relationships suggest that Anbu and the proteasome were already present in the Last Universal Common Ancestor (LUCA) and that the proteasome was linearly inherited by bacteria, archaea and eukaryotes (Figure 5C). This scenario is, in contrast to previous proposals (Figure 2), not dependent on any HGT events. Notably, this linear inheritance is also supported by the presence of homologs of archaeal proteasome interactors, such as the AAA ATPase PAN or proteasome assembly chaperones (PAC), in actinobacteria (De Mot, 2007; Grana et al., 2009). Since these factors do not occur in a single operon with proteasome α and β in archaea, their co-acquisition through HGT events seems highly unlikely. Furthermore, PAC variants with similarities to both actinobacterial and archaeal sequences are also found sporadically in actinobacteria-related phyla (Hug et al., 2016), such as chloroflexi, which do not contain a proteasome (data not shown). This strongly suggests that both the proteasome and its associated factors were already established in the LUCA. Since Anbu represents the precursor for the proteolytic core of the proteasome system in this scenario, we decided to tackle it experimentally.

Anbu forms a dodecameric assembly via the lateral association of six dimers

For experimental characterization, we recombinantly expressed the Anbu-1 protein from the pathogenic bacterium *Pseudomonas aeruginosa* (Pa-Anbu) in *Escherichia coli*. The purified protein was well folded with a thermal melting point (T_m) of 55 °C. It migrated as a single species in size exclusion chromatography with an average molecular mass of ~316 kDa as determined by static light scattering (Figure 6A), and as a single oligomeric species in native PAGE (Figure 6B and C). These results indicated a dodecameric assembly and intuitively suggested an architecture of two stacked hexameric rings similar to HslV.

Crystallization trials yielded a monoclinic crystal form with large cell dimensions that led us to expect about three to four dozens of subunits per asymmetric unit. In the absence of clearly assignable non-crystallographic symmetry and without detectable pseudo-translation, molecular replacement trials with different search models based on various HslV and proteasome β-ring structures failed. We therefore prepared a seleno-methionine derivative, which allowed us to solve the structure based on 170 selenium sites belonging to 34 subunits (Table 3).

Unexpectedly, these 34 subunits in the asymmetric unit constitute the repeating unit of a continuous helix spanning the crystal (Figure 7A). The individual subunits are essentially all in the same conformation, assuming the same fold as the proteasomal β-subunit (Figure 8C). Moreover, a comparison of the active sites shows the catalytic threonine and most other catalytically important residues in essentially the same conformation between Anbu, proteasomal β, and HslV, which suggests the same proteolytic mechanism for all three proteins (Figure 8D). Within the helix, the subunits are arranged in 17 dimers of opposing protomers, distributed over two unequal helical turns. Locally, the subunits are assembled analogous to the β-subunits in the double-β-ring. However, a prominent difference to the latter lies in the extensive contacts formed between the two protomers within the Anbu dimers. Via an extended C-terminal α-helical segment, Anbu forms a dimeric anti-parallel coiled coil-interaction across this subunit interface (Figures 8 and 9), yielding a three-fold increased interface

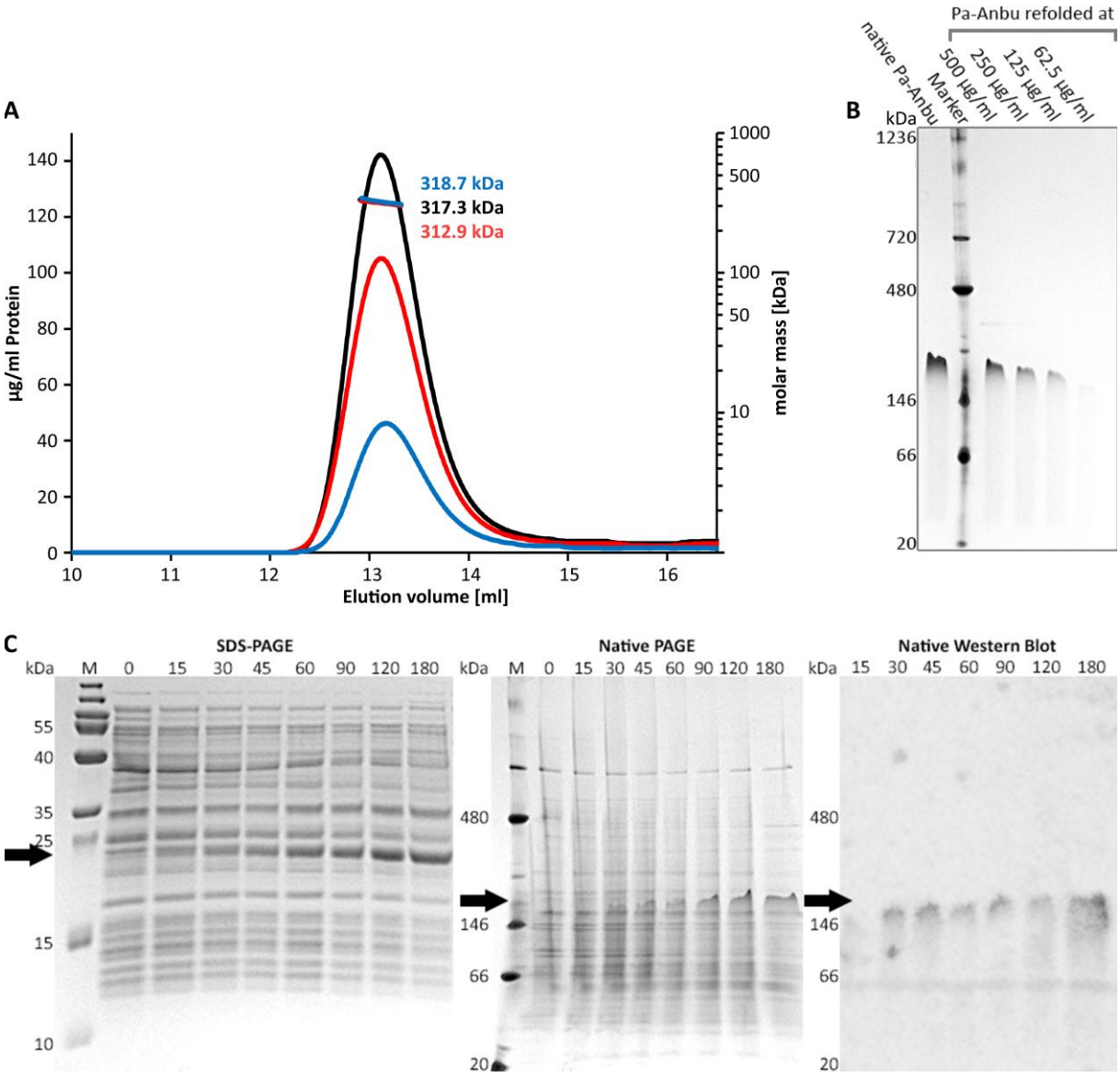


Figure 6: Pa-Anbu is a dodecamer in solution

(A) Light scattering profiles of Pa-Anbu at three different protein concentrations. 100 µl Pa-Anbu at 2 mg/ml (black profile), 1 mg/ml (red) or 0.5 mg/ml (blue) was subjected to a S200 size exclusion column and the mass of the eluted particles in the peak area was analyzed via static light scattering. The protein concentration in the eluted fractions, plotted on the left axis, was determined by integration of the UV signal of all samples.

(B) Native PAGE of refolded Pa-Anbu. Pa-Anbu was denatured in 8 M urea and subsequently refolded at the specified concentrations by dialysis against buffer (20 mM HEPES-NaOH pH 7.5, 150 mM NaCl). The gel shows 10 µl of the refolded soluble fractions, migrating at the same size as native Pa-Anbu (left lane).

(C) Oligomeric state of Pa-Anbu at different expression levels. Pa-Anbu was recombinantly produced via the T7 expression system in *E. coli*. Cells were grown at 37°C in LB medium and induced at an OD of 0.4 with 1 mM IPTG. Protein expression was stopped at the indicated time points (in minutes), and corresponding protein extracts were analyzed by Coomassie Blue-stained SDS-PAGE (left panel), native PAGE (middle panel), and Western blot of a native PAGE using anti-Pa-Anbu antibody (right panel).

Table 3: Data collection and refinement statistics

	Native Pa-Anbu	**Se-Met Pa-Anbu$^{L94M/L112M/L228M}$**	**Native Cons-Anbu**
Wavelength (Å)	1.0	0.979	1.0
Space group	$P2_1$	$P2_1$	$P4_32_12$
Cell dimensions	a=150.1 Å, b=230.0 Å, c=172.8 Å, β=108.0 °	a=150.1 Å, b=230.1 Å, c=171.7 Å, β=108.4 °	a=b=170.7 Å, c=92.10 Å
Monomers / ASU	34	34	4
Resolution range data collection (Å)	39.7 - 3.15 (3.34 – 3.15)	39.6 - 2.90 (3.07 – 2.90)	38.7 - 2.50 (2.65 – 2.50)
Completeness (%)	99.2 (97.3)	99.6 (98.2)	99.8 (98.9)
Redundancy	3.47 (3.46)	14.2 (14.1)	10.9 (10.7)
I/σ(I)	9.79 (2.03)	14.3 (2.31)	19.0 (2.50)
R_{merge} (%)	12.3 (74.8)	17.2 (119)	11.7 (99.1)
CC(1/2)	99.3 (70.1)	99.8 (82.2)	99.9 (78.9)
Resolution range refinement (Å)		39.6 - 2.90 (2.97 – 2.90)	38.3 - 2.05 (2.17 - 2.05)
R_{cryst} (%)		19.8 (35.6)	18.1 (34.0)
R_{free} (%)		21.8 (39.9)	22.2 (40.7)
RMSD Bond angles / lengths		1.45 / 0.012	1.49 / 0.012
Ramachandran statistics (%)		91.1 / 8.9 / 0	93.1 / 6.9 / 0
PDB accession code		5LOX	5LOY

area for this dimer as compared to the respective subunits in the proteasome. Consequently, the intra-dimer-interface is virtually invariant and does not exceed 0.4 Å RMSD in a superposition of all 17 dimers. This is contrasted by a high degree of variability in the lateral interface between the dimers, enabling the formation of the

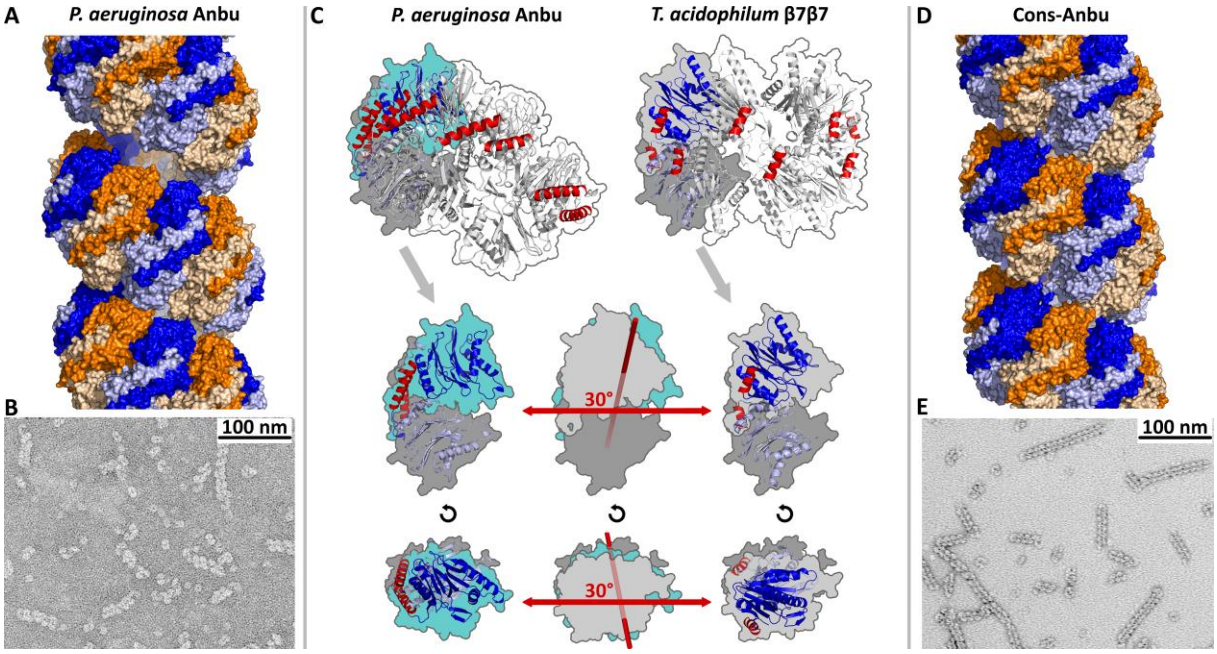


Figure 7: Crystal structures and EM micrographs of Pa-Anbu and Cons-Anbu

(A, D) About 2.5 turns of the continuous helices are shown for each crystal structure. The dimeric subunits of the helices are alternatingly colored orange and blue, with the lower protomers in lighter shades and the upper protomers in darker shades. Adjacent dimers were crosslinked in the preparation of the EM samples (B, E).

(C) Comparison of the geometry of Anbu dimers with the corresponding dimers in proteasomal β-rings. The dimers are shown in the context of the Anbu helix or β-ring (top row), and isolated in a view orthogonal to the helix- or ring main axis (side view, middle row) and in a view along the main axis (top view, bottom row). The lower protomers (dark grey) of the dimers are aligned in their orientation, highlighting the different orientation of the upper protomer (cyan for Anbu, light grey for proteasomal β), totaling in an angular difference of 30°. The respective axis of this rotation is indicated in the middle panels in a superposition of the outlines, which indicate the components of the rotational difference orthogonal and parallel to the main axis. The C-terminal helices of Anbu and proteasome β are highlighted in red.

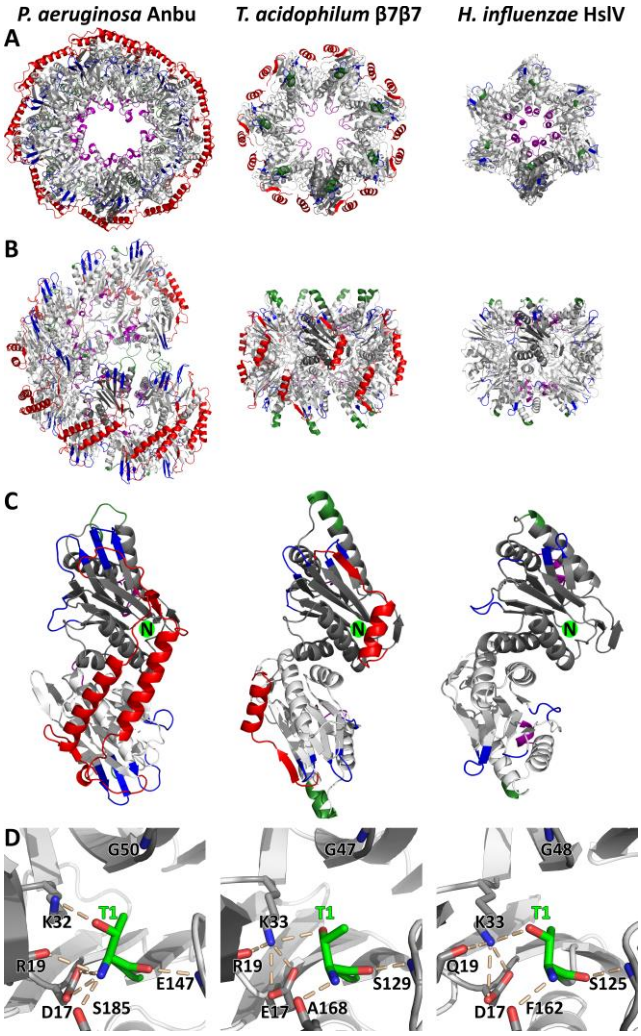


Figure 8: Structural comparison of proteasome homologs

Shown are top (A) and side views (B), as well as dimers (C) of a helical section of the Pa-Anbu crystal structure, proteasome β7β7 (*T. acidophilum*, PDB 1PMA (Lowe et al., 1995)), and HslV (*H. influenzae*, PDB 1G3K (Sousa et al., 2000)). Divergent sequence elements are highlighted according to the alignment in Figure 9: The extended C-terminal region in Pa-Anbu and proteasome β (red); The gate-keeping (Park et al., 2013) pore loop, which is elongated in Pa-Anbu (purple); The main interface between proteasome β- and α- subunits, which is present in unstructured form in Pa-Anbu, but not in HslV (green); Extended loop regions in Pa-Anbu that could possibly serve as a binding motif (blue). N-termini are indicated by green spheres in the upper subunits in (C).

(D) Comparison of the active sites, highlighting the catalytically important residues (Groll et al., 1999). The reaction requires the T1 hydroxyl group to be deprotonated by the T1 backbone amino group. The resulting K32/K33-stabilized T1 oxyanion can subsequently attack the substrate peptide bond, resulting in a G47/G48/G50-stabilized oxyanion-intermediate.

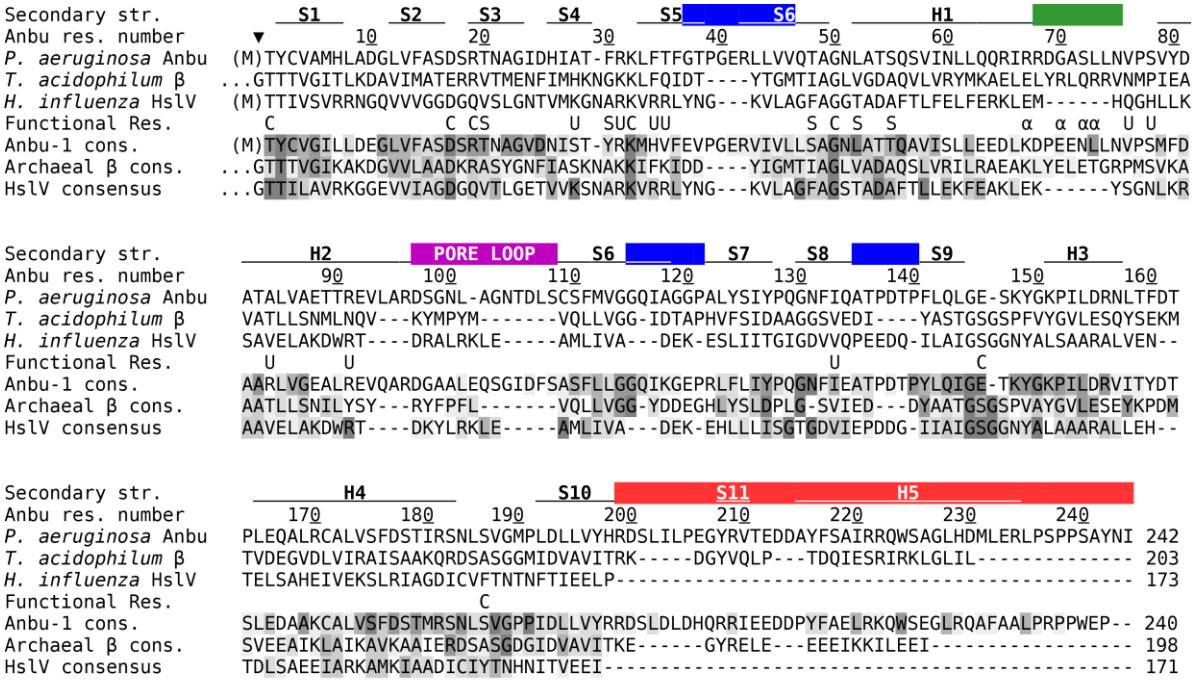


Figure 9: Structure based sequence alignment

The color scheme is as in Figure 8. The degree of conservation is indicated by shades of grey in the consensus sequences of the respective proteins. Important functional residues are marked by C (catalytic center), S (substrate specificity pocket S1) U (HslU interface in HslV) and α (α-interface in proteasome β). The arrow indicates the cleavage site of the pro-peptide. The *H. influenzae* protein is an unusual HslV representative as it lacks this pro-peptide.

irregular helix with its two unequal turns that differ in the number of subunits and in the rise per turn: A superposition of all possible pairs of neighboring subunits yields RMSD values of up to 1.6 Å. To elucidate the importance of the stability and integrity of the dimer interface, we prepared a Pa-Anbu deletion mutant with the C-terminal helix truncated to the length of the respective helix in β-subunits (Pa-Anbu$^{\Delta 227\text{-}242}$), which would be too short to form an interaction across the dimer-interface. Indeed, the mutant protein was unable to assemble into complexes (not shown), substantiating an assembly pathway based on the lateral association of stable dimers. This proposed assembly pathway is fundamentally different to that of the archaeal and eukaryotic proteasome, in which assembly of the β-rings depends on preassembled α-rings as scaffolds (Kunjappu and Hochstrasser, 2014). A pathway employing stable dimers of opposing subunits could thus reflect an ancient trait to form stable proto-β-double rings before the emergence of the α-subunits.

The shape of the Anbu subunits determines helical geometry

To find a structural rationale for the unexpected tendency to form helical assemblies, we analyzed the geometry of the basic Anbu dimers in superimpositions to β-subunits. Compared to the relative orientation of two opposing subunits in the proteasomal β-rings, the two protomers of the Anbu dimers are inclined at a different angle (Figure 7C). The angular difference totals to ~30° and can be decomposed into two components. The first and larger component is equivalent to a closing hinge motion about an axis perpendicular to the main axis of the ring or helix, yielding a more compact dimer. The second component is equivalent to a shearing motion between the two protomers about an axis parallel to the main axis. While this second component is more subtle, it appears to be sufficient to prevent the assembly into closed rings and drive the crystallization of helices.

As continuous helices do obviously not reflect the native dodecameric assembly, we employed electron microscopy (EM) for its visualization, but could not detect well-defined particles in negative-stain images, neither rings nor helices. We then reasoned that the native Anbu particles might potentially be stabilized - and thus visualized - by inter-dimer crosslinks. To this end, we introduced cysteine residues in juxtaposed positions of adjacent dimers (Pa-Anbu$^{A53C/N132C}$) and promoted disulfide bond formation with Cu-phenanthroline. The EM images of crosslinked mutant proteins, however, showed elongated helices reminiscent of those seen in the crystal structure (Figure 7B), suggesting that closed rings indeed do not exist in solution.

The crystal structure of a designed Anbu protein

In order to test the general validity of our findings, we sought to determine the structure of a second Anbu protein. We attempted to express Anbu from several other organisms but did not succeed in obtaining soluble and folded protein. Therefore, we resorted to a different strategy and designed an Anbu-1 consensus protein (Cons-Anbu), an artificial protein with a sequence corresponding to the most frequently used

amino acids in the individual positions of the sequence alignment of all members of the Anbu-1 cluster (Figure 5A). The concept of consensus protein design has emerged as a promising tool in the engineering of stable proteins by making use of evolutionary information embedded in protein sequences (Porebski and Buckle, 2016). Cons-Anbu was indeed highly soluble and even exceeded the thermostability of Pa-Anbu with a T_m of 80 °C. Crystallization trials yielded a tetragonal crystal form with two dimers in the asymmetric unit (Table 3). Around the crystallographic four-fold screw axis, these two dimers assemble into a helix with eight dimers per turn (Figure 7D). The dimers superimpose closely on those of Pa-Anbu with an RMSD of 1.0 Å, and the only noteworthy difference between the two structures lies in the geometry of the helices, which is highly regular in the consensus structure. Also for Cons-Anbu, we did not observe well defined particles in negative-stain EM, but could visualize elongated helices upon cysteine crosslinking (Cons-Anbu^{A53CN133C}) as for Pa-Anbu (Figure 7E).

Anbu forms an open ring of defined geometry in solution

As the structure of the native dodecamer still remained elusive, we decided to study Pa-Anbu in solution via Small Angle X-ray Scattering (SAXS) and obtained a SAXS profile of a well-folded oligomeric particle with low intrinsic flexibility as judged from the Kratky plot (Figure 10B). We initially tried to fit the profile with modeled closed rings of 6 and 7 Anbu dimers with different diameters, but could only obtain reasonable fits when using helical segments of 6 dimers from the Pa-Anbu crystal structure (Figure 10A), clearly indicating that the solution structure is not a closed ring. However, a helical fragment with open ends poses the question of how the assembly process could be terminated after the association of 6 dimers. We were hoping to find the answer by refining the solution structure in a simple rigid-body approach. Based on six rigid Anbu dimers, we systematically constructed helical segments of different geometries, starting from the geometry of a continuous helix as in the crystal structure. Within each model, the same relative coordinate transformation was applied between the six dimers, yielding symmetric helical segments. We thus constructed a large set of such models, by varying the three Eulerian angles of the rotational component in 1° steps, based on different translations, sampling a conformational continuum of open

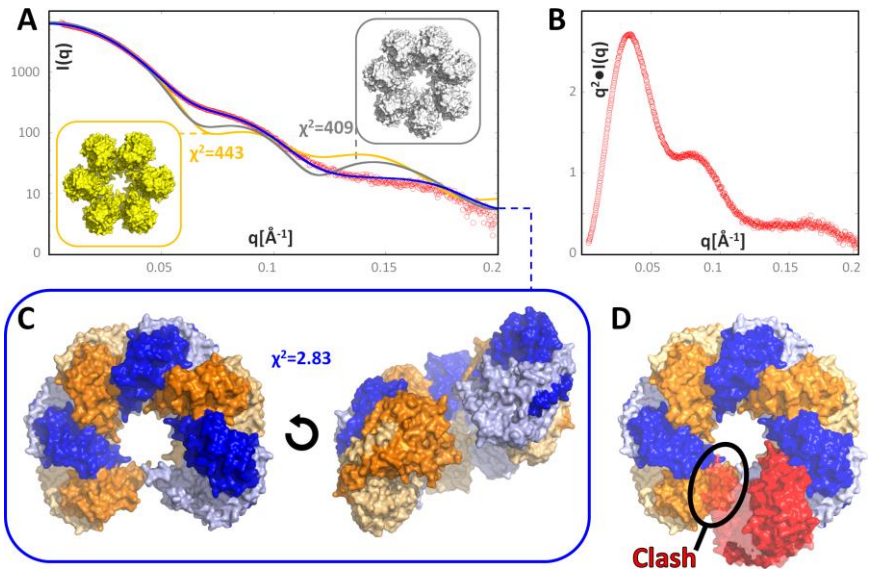


Figure 10: The solution structure of the Anbu complex determined by SAXS

(A) Experimental SAXS data (red) are plotted together with the theoretical profiles of the refined helical complex of six dimers (blue) and the best-fitting models of closed rings with six (yellow) and seven (grey) dimers with the resulting χ^2 values.

(B) The Kratky plot with characteristic shoulders of multi-domain particles, converging to the q-axis at high q, indicating low flexibility.

(C) Top and side view of the refined solution structure with the subunits colored in shades of blue and orange as in Figure 7.

(D) Association of a seventh dimer (red) would result in a steric clash with the first dimer.

ring structures. When these models were fit to the SAXS data, the best agreement was found for models within a small range of parameters, for which the best model with a χ^2=2.83 is depicted in Figure 10. This refined solution structure is an open ring resembling a helical turn with a rise per subunit that is too low to be continued into a second turn – the gap between the first and sixth dimer is too tight for the association of another subunit. As this model – owing to the simple modeling procedure of the refinement approach – is perfectly symmetric, we attempted to refine it further by using established SAXS quaternary structure modeling software without symmetry restraints. However, while these attempts reproducibly resulted in similar overall geometries, their fits to the SAXS profile yielded worse χ^2 values. These results

suggest that the interfaces between the six Anbu dimers are indeed mostly identical in solution, substantiating the fully symmetric open-ring model of the initial approach as the relevant solution structure. Interestingly, functional split-ring structures and spiral architectures have recently been also described for AAA+ ATPases (Glynn and Chien, 2016; Yokom et al., 2016).

The active site differs between Anbu-1 and Anbu-2 but shares common traits with the proteasomal β-subunit

The comparison of Anbu with the proteasomal β-subunit reveals a highly similar active site geometry with essentially all conserved catalytic residues in place (Figure 8D). While Anbu does not possess a propeptide, in contrast to most β-subunits and HslV (Figure 9), its catalytic threonine is exposed at the N-terminus in crystal structures and mass spectra (see below), indicating removal of the start-Met by methionine aminopeptidase *in vivo* (Xiao et al., 2010). In the proteasome, the hydroxyl group of this threonine is deprotonated by the N-terminal amino group and thereby enabled to perform a nucleophilic attack at the peptide bond (Marques et al., 2009). To test this functionality in Pa-Anbu, we assayed Thr-1 binding to the proteasome-specific inhibitors epoxomicin and MG132 via mass spectrometry. To date, no cellular targets aside from the active eukaryotic proteasome subunits are known for epoxomicin, since its reaction mechanism requires both activated threonine hydroxyl and amino groups and appropriate positioning by the proteasome substrate binding channel (Huber et al., 2015). We found that both MG132 and epoxomicin were efficiently bound to wild-type Pa-Anbu (Figures 11A and B), but not to Pa-Anbu with mutated Thr-1 (Pa-Anbu^{T1A}; Figures 11C and D) or modified N-terminus (Sumo-Pa-Anbu; Figures 11E and F). In agreement, MS/MS-spectra of AspN digested Pa-Anbu unambiguously confirmed Thr-1 as site of epoxomicin modification, indicating that this residue indeed acts as N-terminal nucleophile in Pa-Anbu (Figure 12).

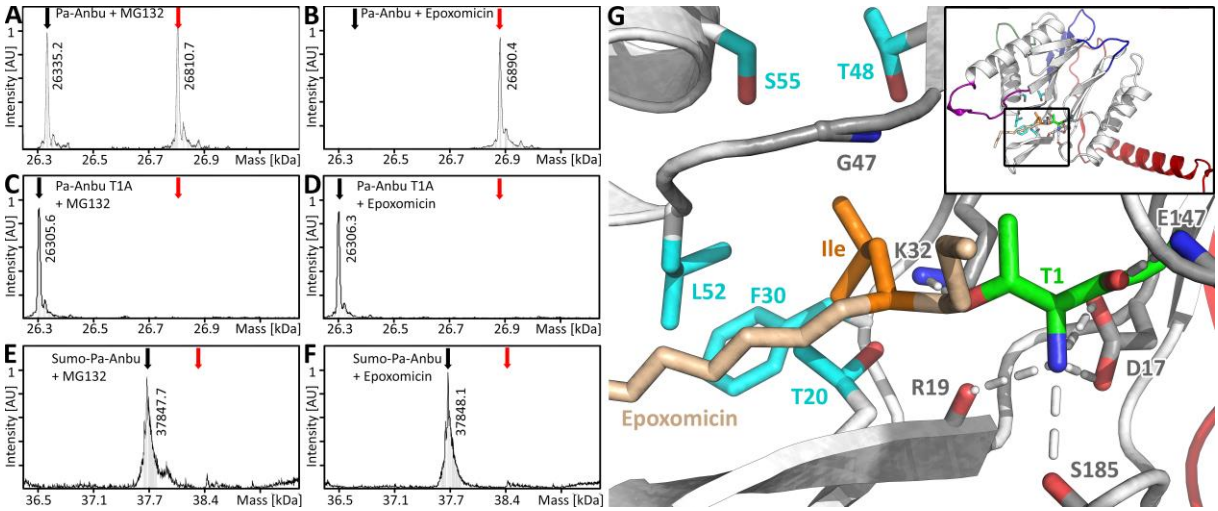


Figure 11: Mass spectrometric analysis shows the binding of epoxomicin and MG132 to Anbu

Pa-Anbu (A, B), Pa-Anbu^{T1A} (C, D) and Sumo-Pa-Anbu (E, F) were incubated with the proteasome specific inhibitors MG132 or epoxomicin. The resulting protomer masses were determined via mass spectrometry and are shown in comparison with the theoretical masses of the conjugated (red arrow) and unconjugated (black arrow) protomers.

(G) Model of the Pa-Anbu catalytic center in complex with epoxomicin. The protein backbone is shown as cartoon (white) with the catalytic T1 (green) as well as other residues thought to be crucial for catalysis (grey) and the residues forming the S1 pocket (cyan) represented as sticks. Hydrogen bonds to T1 are indicated by dotted white lines. The structure of the inhibitor epoxomicin, whose binding to Anbu resulted in poorly diffracting crystals, is taken from the structure of the conjugated *S. cerevisiae* β5-subunit (PDB 1G65 (Groll et al., 2000)), which was superimposed on the Pa-Anbu catalytic center with a RMSD of 0.67 Å. For simplicity, only the epoxomicin backbone (pale orange) and the terminal leucin-sidechain at P1 (orange) are shown.

Interestingly, when we tested an Anbu-2 version from *C. hutchinsonii* for this activity, we found the protein unmodified by both inhibitors (data not shown). A similar preference for these inhibitors was observed for various proteasome β-subunits and ascribed to their different substrate specificities. These are primarily governed by the composition of the so-called S1-pocket, which binds the side-chain N-terminal to the scissile bond (P1), positioning the latter for cleavage (Huber et al., 2012). The three catalytic eukaryotic β-subunits can be distinguished by using different residues in the corresponding positions, most importantly residue 45, and their consistent activity

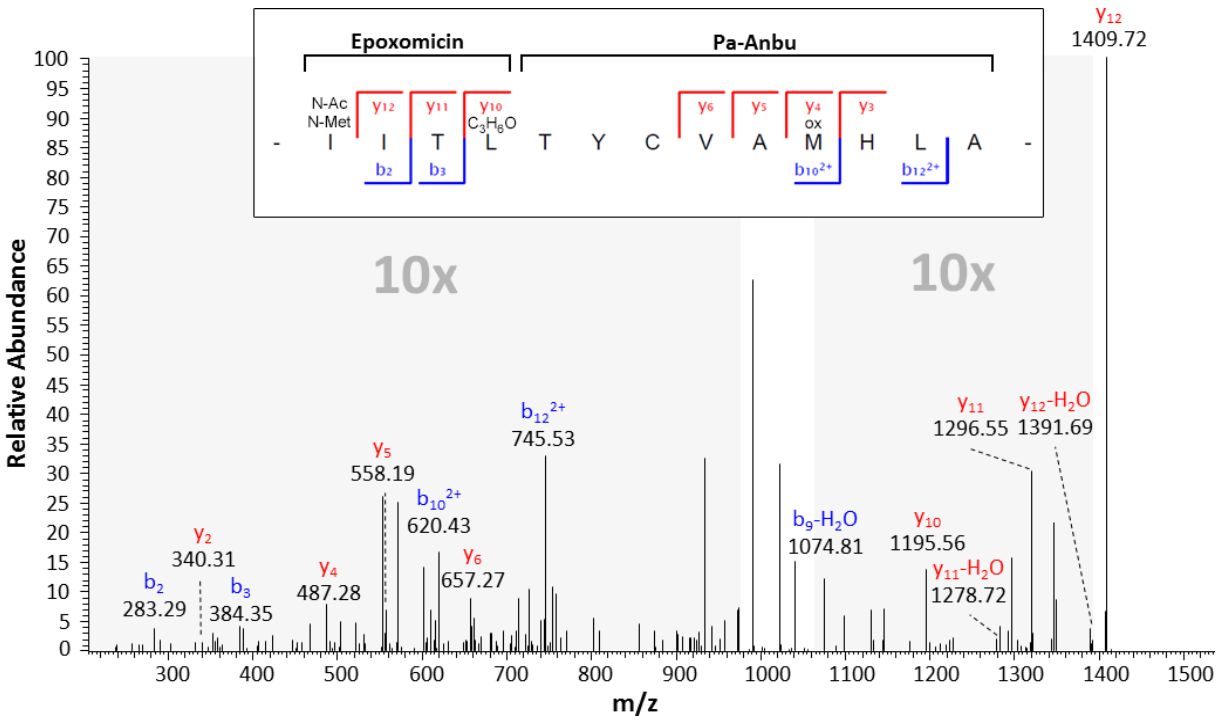


Figure 12: MS/MS spectrum of the epoxomicin-modified N-terminal Pa-Anbu fragment after AspN digestion

Epoxomicin-treated Pa-Anbu (Figure 11B) was digested with AspN and subjected to MS analysis. A mass shift corresponding to the molecular weight of epoxomicin (554.36 Da) was detected only for the N-terminal fragment, but not for other fragments in the digest. The sequence (see inset) of the modified N-terminal fragment was traced via the shown MS/MS-spectrum. For this purpose, the epoxomicin moiety was interpreted as the peptide IITL with N-terminal acetylation and methylation. The morpholino adduct between epoxomicin and Pa-Anbu Thr-1 is represented by its chemical formula, C_3H_6O (Wei et al., 2012). Within the spectrum, b and y series ions are indicated. For better visualization, grey areas in the spectrum were magnified tenfold.

against acidic (*S. cerevisiae* β1 S1-pocket: T20-T31-R45-A49-T52), spacious basic (*S. c.* β2 S1: S20-C31-G45-A49-T52-E53) or hydrophobic (*S. c.* β5 S1: A20-V31-M45-A49-C52) peptide substrates (Marques et al., 2009). Anbu was found to be inactive against a standard set of these substrates under numerous conditions tested, which could possibly be ascribed to the requirement for activators or post-translational modifications which were absent in our experiments. However, the Pa-Anbu S1 pockets superimpose exceptionally well with those in the proteasome, revealing a relatively narrow, amphiphilic S1-pocket, in which F30 and L52 form a hydrophobic patch, whereas the moderately hydrophilic T20, T48 and S55 are oriented for

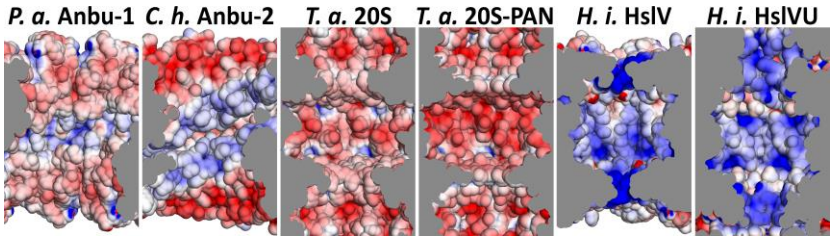


Figure 13: Comparison of inner cavities of proteasome homologs

Shown are *H. influenza* HslV in the presence (PDB G3I) or absence (PDB G3K) of the AAA interactor HslU, the *T. acidophilum* proteasome in presence (PDB 3IPM) or absence (PDB 1PMA) of the AAA interactor PAN (PDB 1PMA), Pa-Anbu and a *C. hutchinsonii* Anbu-2 model. ±5 kT/e Poisson-Boltzmann electrostatic potentials are plotted on the surface of the cut-open rings using APBS (Baker et al., 2001), with negative potentials in red and positive potentials in blue. The *C. hutchinsonii* Anbu-2 cavity was prepared by homology modelling with Pa-Anbu as a template.

hydrogen bonding to various positions in P1. Remarkably, while these S1-pocket residues are highly conserved in Anbu-1 (Figure 9), all Anbu-2 sequences are seen to have an entirely different S1-pocket (*C. hutchinsonii* Anbu S1: I20-A29-R46-R50-T53). Its highly basic characteristic potentially explains why this Anbu-type is incapable of binding MG132 and epoxomicin, which contain hydrophobic residues at P1. Since only S1-pockets with hydrophobic properties are found in archaeal β, bacterial β and HslV sequences, Anbu-2 might represent the first prokaryotic proteasome homolog tuned for acidic cleavage specificity. This is mirrored by the properties of the inner cavity surface, which is more positively charged in Anbu-2 (Figure 13).

Although there are no known interactors, there are two unique structural features on the surface of the Anbu complex that could play a role in the docking of potential cofactors. The most obvious of the two features is the unused oligomerization-interface on either end of the open ring, which is almost fully solvent-exposed. The second feature concerns the outer rims of the ring, which protrude more prominently towards the outside than in the β-rings (blue in Figures 8 and 9). As both these potential binding surfaces form a right angle with respect to each other at either end of the ring, it is tempting to speculate that they might jointly mount a cofactor each on top and on the bottom of the complex.

The biological function of Anbu

What can be discerned from our structural studies about the physiological role of Anbu? The presence of many identical cleavage sites in a ring-shaped structure, as seen here, argues for processivity, a key feature of self-compartmentalizing proteases (Akopian et al., 1997). Substantiating a protease function is the binding of specific proteasome inhibitors which mimic the first steps of the cleavage reaction (Figure 11), and the evolution of different substrate binding pockets in Anbu-1 and Anbu-2. Thus, our characterization agrees with a function of Anbu in protein degradation, making the previously proposed involvement in a peptide-synthesis system (Iyer et al., 2009) less likely. As different proteasome homologs frequently co-occur within a given organism, rather than being mutually exclusive as originally thought (De Mot et al., 1999), their different substrate specificities could provide the cell with a broader functionality, like it is accomplished by the presence of alternative proteasome subunits and interactors in eukaryotes (Fort et al., 2015).

While a partner AAA ATPase for Anbu appears to be absent, the proteasome provides a precedent for the exertion of ATP-independent functions, both with non-ATPase interactors and as uncapped 20S proteasome, which has been implicated in the degradation of oxidized and otherwise damaged proteins (Pickering and Davies, 2012a). While being expressed constitutively in moderate amounts, neither Anbu nor the other operon proteins are essential for bacterial growth under non-stress conditions (Molina-Henares et al., 2010). Their expression does not change under various environmental conditions (Dotsch et al., 2015), including nitrosative stress (Firoved et al., 2004), a situation in which the mycobacterial proteasome is required for resistance to nitric oxide (Darwin et al., 2003). Anbu is upregulated, however, in nitrogen-limited *Pseudomonas putida* cultures (Hervas et al., 2008), suggesting a contribution to a higher protein turnover under these conditions. The activity and active conformation of Anbu could be further regulated by covalent modifications, as is known for the proteasome (Hirano et al., 2016). However, attempts to obtain a different Anbu form by purifying the endogenous protein from *P. aeruginosa* or by co-expression with the other operon proteins were not met with success (not shown). Thus, a key task for the future will be to find physiological substrates and potential interactors.

Anbu as a precursor of self-compartmentalizing proteases

Our bioinformatic analysis places Anbu at the very early stages of proteasomal evolution in the LUCA, but it cannot clearly assign the relative ancestry of Anbu and the proto-β-subunit. However, our structural characterization suggests an evolutionary scenario in which the Anbu complex represents a precursor of the self-compartmentalizing proteases (Figure 14). Generally, in evolving a homo-oligomeric assembly, two distinct, complementary interaction surfaces have to be established on the protomer. In a first trial, it is highly unlikely that the relative orientation of these interfaces yields either a perfectly straight polymer or a closed ring without helical rise. Most likely, the orientation of the interfaces would yield a curved polymer with non-zero helical rise. However, a helical rise larger than the height of the subunits would potentially yield infinite polymers, which would not be desirable in many cases. If the helical rise per turn is sufficiently small though, the assembly cannot proceed into a second turn, but instead forms a defined oligomer, as seen in Anbu. Such, a primordial Anbu- or Anbu-like protein had already established the ability to oligomerize into proteasome-like proteolytic chambers in the LUCA. However, its helical geometry was subsequently superseded by the formation of closed ring assemblies that allowed the construction of modular systems as the proteasome, facilitating the recruitment of further rings of regulatory proteins. A key necessity in this process was presumably the trimming of the C-terminal helices to disrupt the rigid coiled-coil interface of the dimeric subunits. While Anbu was then lost in archaea, a gene duplication of proteasome β led to the emergence of HslV in the common ancestor of bacteria, which subsequently replaced the proteasome in non-actinobacterial strains (Figure 5C). The Anbu proteins we characterized here are presumably direct descendants of the ancient evolutionary intermediate that survived in a functional niche alongside the modular bacterial self-compartmentalizing proteases.

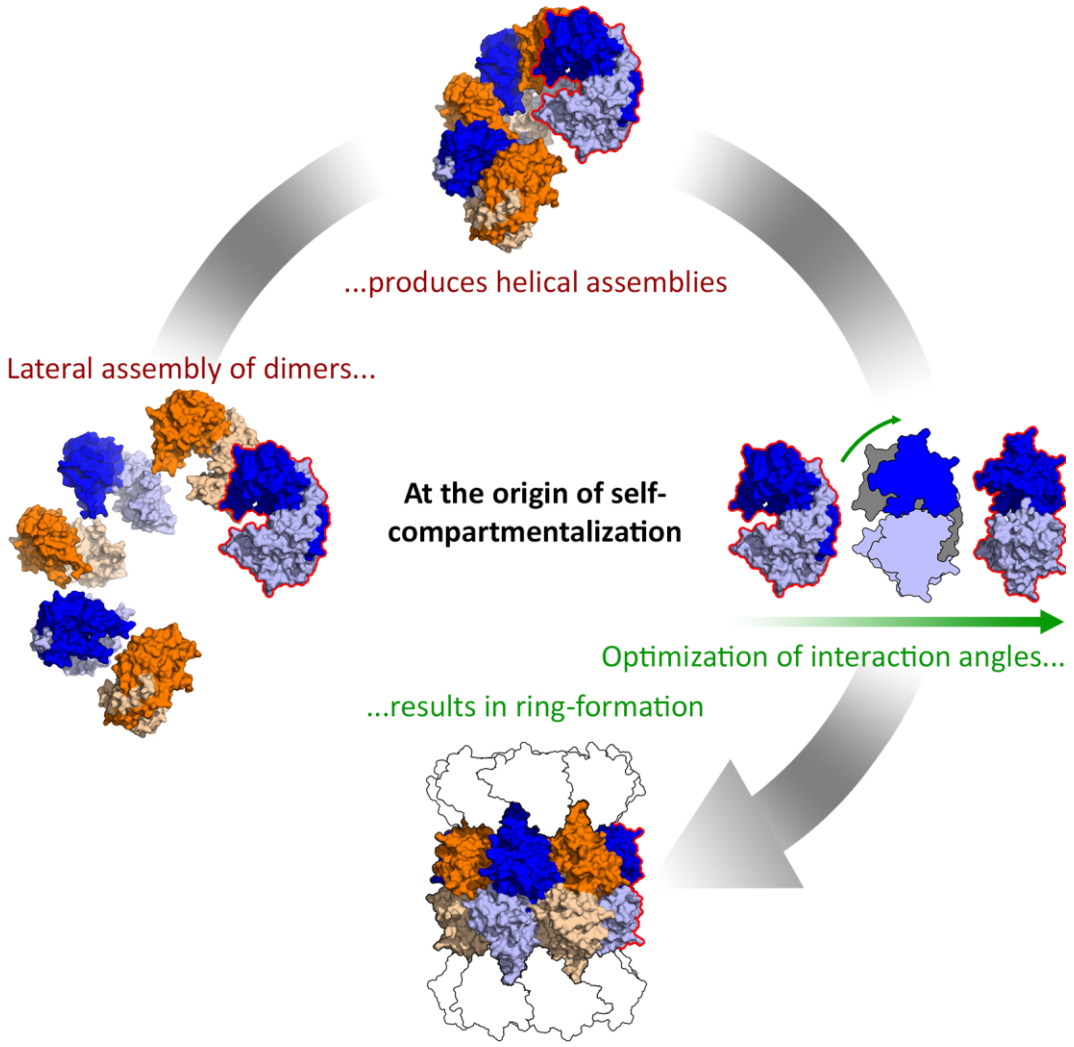


Figure 14: Summary of the evolutionary concept presented in this chapter

Our bioinformatic analysis suggested that Anbu and the proteasome share a common root in the LUCA, but could not reveal their relative ancestry. Based on its helical open-ring structure, we propose that Anbu presents an evolutionary intermediate that resulted from a lateral assembly of dimers and that gave rise to the ring-shaped proteasome through an optimization of interaction angles.

The results presented here have been published in the Structure Journal as (Fuchs et al., 2017). Articles and Figures of this publication have been used - with permission - in this chapter. In the meantime, other Anbu structures, that support our findings but were not accompanied by publications so far, have been deposited in the PDB. These are: 5NYF, 5NYG, 5NYJ, 5NYP, 5NYQ, 5NYR, 5NYW.

3.2 Structural characterization of the novel bacterial proteasome homolog BPH reveals a tetradecameric double-ring complex with unique inner cavity properties

While our study of Anbu revealed an ancestral complex that was placed at the origin of the proteasome system, BPH might present a relatively young proteasome homolog, that is more similar to HslV. The biology of BPH is terra incognita, virtually nothing is known about its structure and function. In this chapter, we present the first experimental characterization of two bacterial BPH proteins. Their crystal structures reveal homo-oligomeric double-barrel complexes with internalized active sites, akin to those of proteasome, HslV and Anbu. Based on a combined analysis including SAXS and electron microscopy averaging, we show that BPH forms tetradecamers in solution, unlike the dodecamers seen in HslV. While the highly acidic inner surface of BPH is in striking contrast to the cavity characteristics of the proteasome and HslV, a classical proteasomal reaction mechanism can be inferred from the covalent binding of the proteasome-specific inhibitor epoxomicin. A ligand-bound structure implies a role for the elongated BPH inner pore loop in regulating substrate access to the active site. The apparent lack of a partner unfoldase and other unique adaptations, such as Ser replacing Thr as catalytic residue in certain BPH family members, suggest a proteolytic function for BPH distinct from known bacterial self-compartmentalizing proteases.

BPH is an evolutionary descendant of HslV

To investigate the evolutionary history of BPH, we reviewed our cluster analysis of proteasome homologs (Figure 15A; see chapter 3.1 for a detailed analysis). The map shows BPH sequences at the periphery, with strong connections to HslV. However,

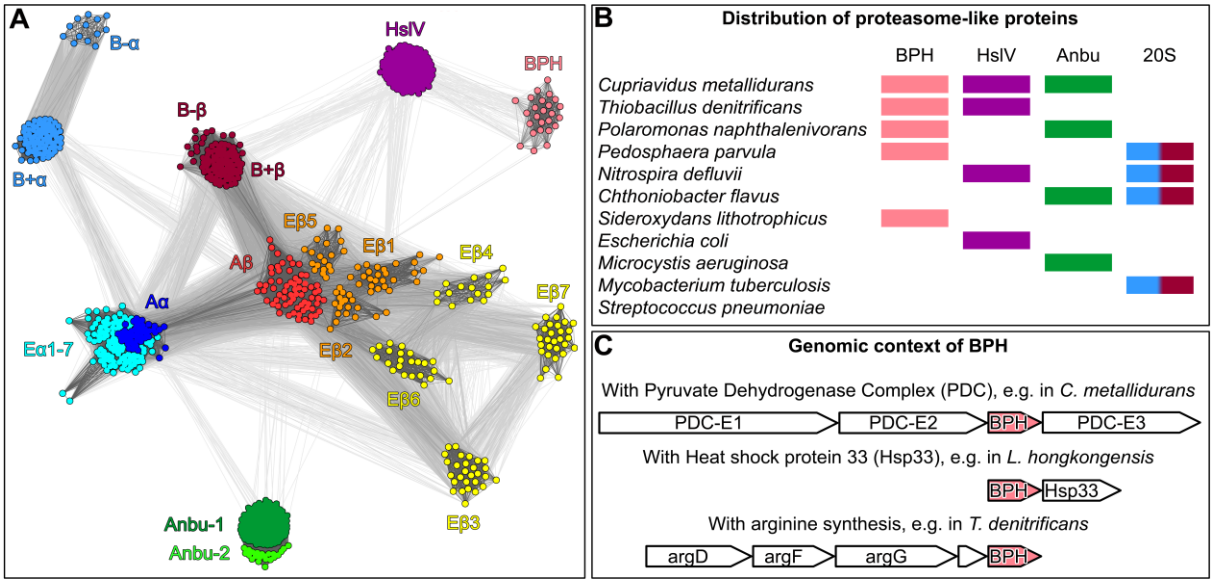


Figure 15: Phylogeny and occurrence of BPH genes

(A) Cluster map of 969 proteasome-like sequences, with a maximum pairwise identity of 70%, was prepared using CLANS (Frickey and Lupas, 2004) based on their all-against-all pairwise similarities as measured by BLAST p values. Sequences are represented by dots and the line coloring reflects BLAST p values (the darker a line is, the lower the p value). Proteasome subunits are abbreviated E for eukaryotic, A for archaeal, and B+/B– for Gram-positive/-negative bacteria.

(B) The co-occurrence of BPH, HslV, Anbu and the proteasome (20S) in representative organisms; bars are in the same color scheme as in (A).

(C) Genomic context of BPH in representative organisms.

HslV is more similar to any other analyzed homolog, indicating that BPH presents a more recent addition to the proteasome family, that emerged by gene duplication of HslV. Strictly speaking, however, the acronym BPH (β-proteobacteria proteasome homologue) does not accurately reflect the distribution of this protein in bacterial species as, in variance with the initial description (Valas and Bourne, 2008), BPH is also found outside the β-proteobacteria, e.g. in members of verrucomicrobia, α- and γ-proteobacteria, cyanobacteria or acidobacteria. The phylogenetically widespread, though sparse occurrence of BPH outside the β-proteobacteria could be attributed to horizontal gene transfer. Arguing against this is the finding that all BPH sequences outside the β-proteobacteria surprisingly feature Ser-1 as active site residue, in

contrast to the prototypical Thr-1 present in β-proteobacterial BPH and the proteasome family as a whole, suggesting that already early on different BPH variants existed which then diverged in their evolutionary paths. It seems plausible that BPH originated from HslV before the emergence of β-proteobacteria, which still seem to provide a suitable ecological niche for this protease, whereas in many other bacterial species the gene disappeared over time. In extant organisms, BPH would be a vestige of the past that survived due to a beneficial function or activity. As there is a fair amount of redundancy amongst bacterial cytosolic proteases, such an activity does not have to be unique, but could be overlapping with that of its evolutionary relatives. In fact, BPH occurs in bacterial species in various combinations together with Anbu, HslV or the 20S proteasome, or is even the sole proteasome family member present, e.g. in *Sideroxydans lithotrophicus* (Figure 15B). Finally, its organismal distribution does not correlate with that of its ancestor HslV, in line with the notion that also HslV was lost from a number of genomes at later times.

BPH genes are constitutively expressed genes occurring in stress-regulated operons

Before embarking on the biochemical characterization of BPH proteins, we tried to extract information about their physiological role from the genomic context of their genes. In prokaryotes, genes that functionally and/or physically interact are often grouped together in their respective genomes (Osbourn and Field, 2009). When BPH genes were first analyzed (Valas and Bourne, 2008), three types of operon environments were found to prevail (Figure 15C). In some organisms, e.g. *Thiobacillus denitrificans*, BPH is grouped in an operon with argininosuccinate synthase (argG) and adjacent to argD and argF, proteins involved in arginine biosynthesis. In other organisms, e.g. *Cupriavidus metallidurans*, BPH is found in an operon with the three genes encoding the pyruvate dehydrogenase complex (PDC). Finally, in other organisms, e.g. *Laribacter hongkongensis*, BPH is in the same operon as heat shock protein 33 (Hsp33). None of these operon contexts hint at the existence of regulatory partner proteins for BPH, nor of companion AAA ATPases that could act as unfoldases. Moreover, when we searched BPH-containing genomes for suitable

ATPase candidates, none were found. This raises the possibility that BPH acts autonomously in an ATP-independent fashion. To address the issue of interaction partners, we performed pull-down experiments with bacterial cell extracts, using purified BPH as bait. As we had found that BPH from the β-proteobacterium *C. metallidurans* was amenable to biochemical and structural studies (see below), we used this organism as model system for our experiments. Cm-BPH was recombinantly expressed in *E. coli* and formed a stable high-molecular weight complex, also when C-terminally tagged. Pull-downs from *C. metallidurans* extracts were done on HA-, Strep- or Myc-magnetic beads and candidate bands obtained on SDS-gels were analyzed by mass spectrometry. No obvious interaction partners could be identified though (data not shown). Notably, neither PDC subunits (from the same operon) nor HslU (as potential partner ATPase) were found to be associated with BPH.

Even though at first glance the three different operon contexts do not seem to have much in common, it turns out that they are all linked to the cellular stress response. Hsp33 is a chaperone expressed under heat shock conditions and activated under oxidative stress (Winter et al., 2005). Similarly, arginine biosynthesis is increased during heat shock and oxidative stress in several species (Helmann et al., 2001; McLean et al., 2010; Zhang et al., 2006). Arginine is the substrate for NO-synthase, whose product nitric oxide interferes with the formation of reduced thiols and the recycling of ferrous iron, thus diminishing the generation of damaging hydroxyl radicals (Gusarov and Nudler, 2005). Finally, PDC subunits are prone to carbonylation, an irreversible oxidation of amino acid side-chains caused by hydroxyl radicals (Fredriksson et al., 2005), which, among many other proteins, inactivates PDC and turns it into a target for regulated degradation. While available Microarray data indicate that BPH is also expressed constitutively (Barrett et al., 2013), the correlations described above suggest a stress-related protease function for BPH which we set out to test. We first assayed the degradation of carbonylated proteins by BPH. Carbonylation can be detected by derivatization of the carbonyl groups with dinitrophenylhydrazine (DNPH) (Levine et al., 1990). However, the carbonylation patterns of *C. metallidurans* extracts that were grown under oxidative stress conditions did not change upon incubation with BPH (Figure 16A). For a second test system, we

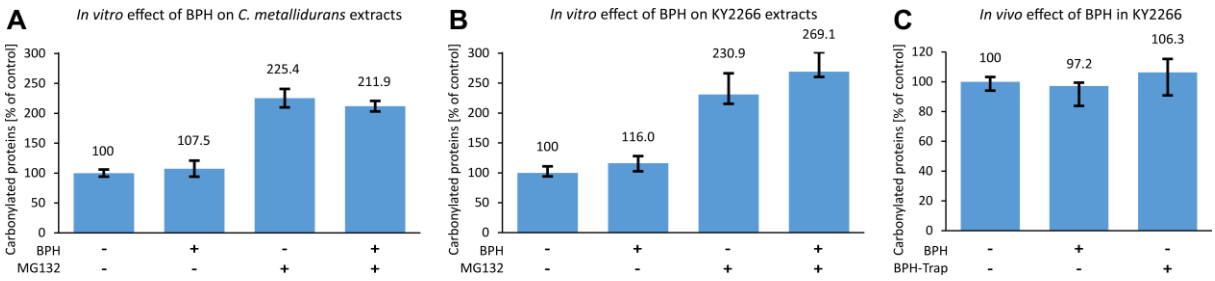


Figure 16: Effect of BPH on the levels of carbonylated proteins under oxidative stress

(A,B) *C. metallidurans* (A) or *E. coli* KY2266 (Δ*lon,* Δ*hslVU,* Δ*clpPX* (Kanemori et al., 1997)) (B) extracts were produced from stationary-phase cultures grown with or without the proteasome inhibitor MG132 and treated with Cm-BPH. The levels of carbonylated proteins were quantified photometrically with 2,4-dinitrophenylhydrazine.

(C) *E. coli* KY2266 cells were transformed with an arabinose-inducible Cm-BPH gene or a Cm-BPH trap gene with mutated catalytic residues. Oxidative stress was induced by a prolonged 16 hrs stationary phase incubation, and the levels of carbonylated proteins were subsequently quantified as in (A) and (B).

resorted to the *E. coli* strain KY2266 in which the major cytosolic proteases are deleted (Δlon, ΔhslVU, ΔclpPX), providing us with minimal protease background and redundancy (Kanemori et al., 1997). Again, addition of BPH to KY2266 extracts did not change the protein carbonylation pattern (Figure 16B). For corresponding in vivo experiments, we transformed KY2266 with the Cm-BPH wild-type gene or a trap mutant with mutated catalytic residues, under control of an inducible arabinose promoter. After induction of oxidative stress through prolonged incubation at the stationary phase, cells were lysed and carbonylated proteins were quantified. Like with the in vitro extracts, BPH did not affect carbonylation in vivo (Figure 16C). Interestingly though, while KY2266, with or without Cm-BPH, grew slower than its parent strain MC4100 in LB and also at low H_2O_2 concentrations (1 mM), at higher H_2O_2 concentrations (2.5-5 mM) the protease-deficient mutant strain thrived better than the wild-type (Figure 17). The enhanced stress resistance of KY2266 could be explained by the increased levels of heat shock proteins and stabilization of stress transcription factor σ^{32} in the mutant strain, caused the absence of the cytosolic proteases (Kanemori et al., 1997). In summary, we conclude that BPH cannot substitute for the missing self-compartmentalizing proteases in our assays.

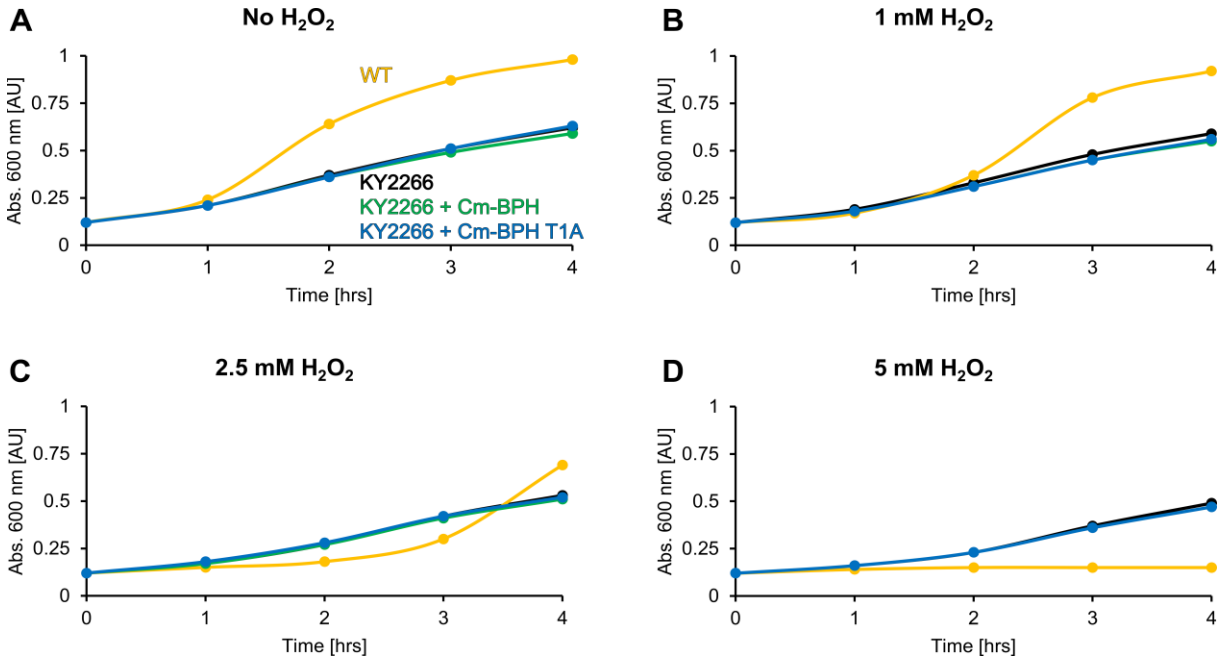


Figure 17: Effect of BPH-expression on *E. coli* growth rates in the presence of hydrogen peroxide

Growth curves for the wild type *E. coli* strain MC4100 or its mutant strain KY2266 (Δ*lon,* Δ*hslVU,* Δ*clpPX*) were recorded photometrically in the absence (A) or presence of 1 mM (B), 2.5 mM (C) or 5 mM (D) H_2O_2 and compared to KY2266 cells expressing Cm-BPH or a Cm-BPH trap mutant with mutated catalytic residues (see Figure 16C).

BPH assembles into tetradecameric double-ring complexes with a unique acidic inner surface

In the absence of obvious interaction partners, the question arises as to how BPH recognizes its substrates. Moreover, how is access granted to the active sites unless they are not shielded? We tried to find answers to these questions in the molecular architecture of BPH and did a structural characterization of this novel protein complex. We opted for the analysis of BPH from *C. metallidurans* and *T. denitrificans*, two representatives of the β-proteobacteria, yet with different lifestyles. *C. metallidurans* is an aerobic chemolithoautotroph which degrades xenobiotics and is adapted to survive heavy metal stress (Janssen et al., 2010), whereas *T. denitrificans* is a facultatively anaerobic chemolithoautotroph oxidizing inorganic sulfur compounds (Beller et al.,

2006). Both BPH sequences are shown in the alignment in Figure 18A. All active site residues (yellow) are conserved in BPH proteins, with the noticeable presence of Ser as N-terminal nucleophile in BPH-containing organisms outside the β-proteobacteria. The BPH sequences do not feature propeptides, suggesting that, contrary to the situation in the proteasome, the subunits should be in their active state right after assembly. As one would predict, residues mediating the interaction of HslV with HslU (purple) are not conserved in its descendant BPH, which became independent of the accessory ATPase.

When recombinantly expressed in *E. coli*, both Cm-BPH and Td-BPH formed stable complexes (Tm 63°C and 61°C respectively). We crystallized both proteins and obtained crystals diffracting at 2.1 (Cm-BPH) and 2.0 Å (Td-BPH) resolution. Unexpectedly, the structure of Cm-BPH (Figure 18B and C) shows a double-ring cylinder with sixfold symmetry, whereas the Td-BPH double-ring has sevenfold symmetry. One way to explain this difference is to invoke structural diversity within the BPH family; after all, both sixfold (HslV) and sevenfold (proteasome) symmetries persist in related proteins of the same fold. Alternatively, one of the crystal structures does not reflect the solution state of the protein. To resolve this issue, we subjected Cm-BPH and Td-BPH to biophysical characterization including SAXS, analytical ultracentrifugation and electron microscopy (Figure 19A and B). In a similar approach as described in chapter 3.1, we plotted the theoretical SAXS profiles of the dodecameric (grey curve in Figure 19A and B) and tetradecameric (blue curve) BPH complexes against the experimental scattering data. Interestingly, the profile for the tetradecameric model closely delineated the Cm-BPH dataset, while the profile for the dodecameric model differs in a prominent minimum. A similar result is observed for the Td-BPH dataset, even though it is not mirrored by lower $\chi 2$ values in this case. These results suggest that Cm-BPH might assemble as a tetradecamer in solution, just like Td-BPH.

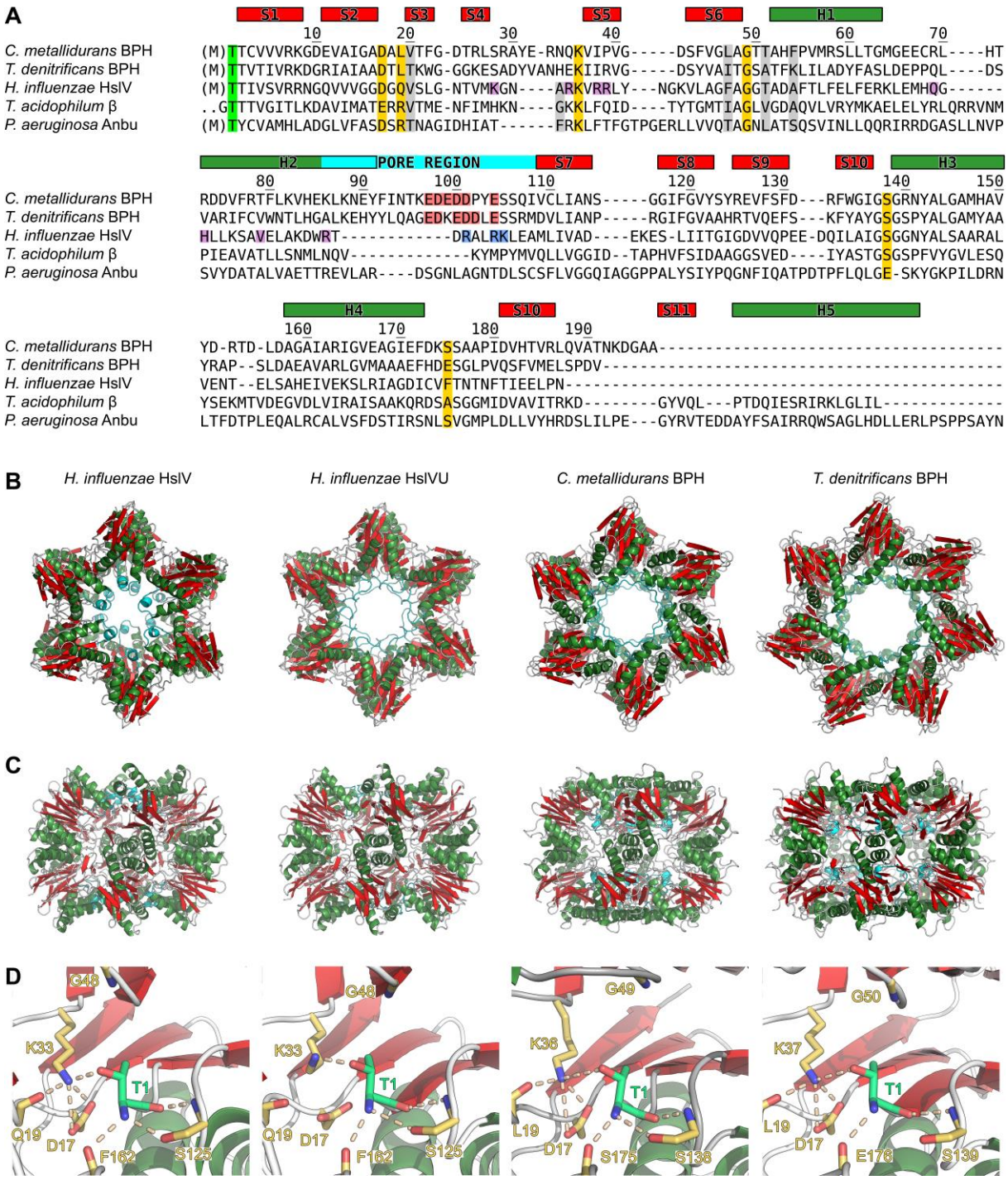


Figure 18: Sequence and structure comparison of BPH and HslV complexes

(A) Structure-based sequence alignment of proteasome-like proteins. Active site residues (yellow; Thr1 in green), residues forming the substrate specificity pocket S1 (grey; (Huber et al., 2012)), and HslU interface residues (purple) are highlighted. Residues accounting for the acidic BPH inner cavity characteristics are shown in red and the corresponding HslV Arg-86/Arg-89/Lys-90 basic cluster in blue. Colored structural elements in B and C are indicated

by red (sheets), green (helices) or cyan (pore region) bars. The residue number corresponds to to *C. metallidurans* BPH.

(B, C) Top- (B) and side-view (C) representations of *C. metallidurans* and *T. denitrificans* BPH crystal structures, in comparison with *Haemophilus influenzae* HslV in the presence (PDB 1G3I) or absence (PDB 1G3K) of HslU.

(D) Active sites comparison, highlighting the catalytically important residues (Groll et al., 1999).

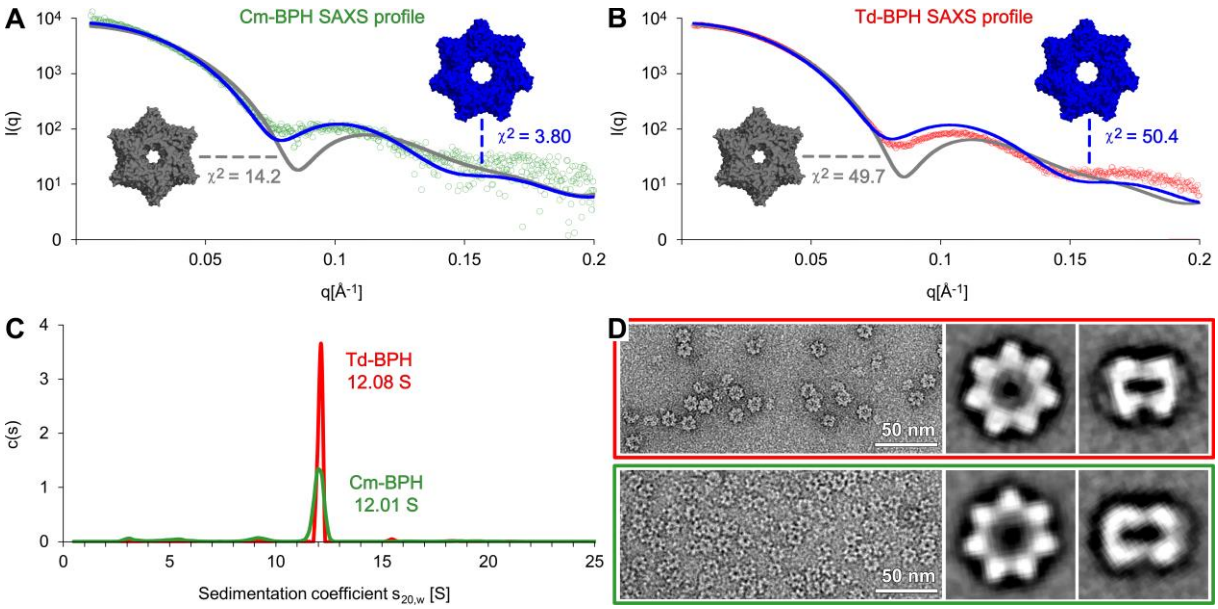


Figure 19: BPH forms a tetradecameric double-ring in solution

(A,B) Experimental SAXS data are plotted together with the theoretical profiles of dodecameric (grey) or tetradecameric ring-structures (blue).

(C) Sedimentation coefficient distribution of Cm-BPH and Td-BPH complexes, as determined by analytical ultracentrifugation.

(D) Electron micrographs of Td-BPH (red box) and Cm-BPH (green box). Particles in these preparations were classified and averaged, resulting in top- (middle) and side-view (right) representations of both proteins.

In analytical ultracentrifugation runs, almost identical $s_{20,w}$ sedimentation coefficients of 12.01 and 12.08 were obtained for Cm-BPH and Td-BPH respectively (Figure 19C). Finally, when we averaged negative-stain electron micrographs of Cm-BPH and Td-BPH particles, exclusively seven-membered rings were obtained for both proteins (Figure 19D). In summary, both BPH proteins unequivocally form double-rings with sevenfold symmetry in solution. The sixfold symmetry X-ray structure of Cm-BPH can thus be seen as the result of subunit rearrangements under crystallization conditions. Given that BPH evolved from a protein with sixfold symmetry, its subunits apparently have not unlearned to assemble into closed rings of ancestral symmetry, even though in their natural cellular environment they can be assumed to form distinct tetradecameric oligomers.

If one looks at the inner dimensions of various proteasome-like barrels, it is evident that both Cm-BPH and Td-BPH display somewhat larger pores inside the cylinder in comparison to HslV. In the presence of HslU, however, the HslV pore is known to widen considerably (Figure 18B and C). Taking a closer look at the inner cavity, one finds that BPH proteins are distinguished from other proteasomal proteins by their rather long pore loops. Whereas in HslV the loop is merely a short linker between H2 and S7, the extended BPH pore loop is long enough to make contacts with its adjacent subunit (Figure 18, A-C). The result is an almost mesh-like lining of the pore surface. Moreover, as BPH pore loops have a very acidic sequence lineup, exemplified by five successive Asp/Glu residues in Cm-BPH, this results in a unique inner surface environment (Figure 20), which is in striking contrast to that of its ancestor HslV, where the basic residues Arg-86, Arg-89 and Lys-90 protrude towards the inner cavity at the pore entrance. Interestingly, point mutations in these residues cause a severe decrease of HslV's proteolytic activity (Ramachandran et al., 2002) and HslU binding induces structural changes in this region, resulting in a switch of cavity characteristics (Figure 20).

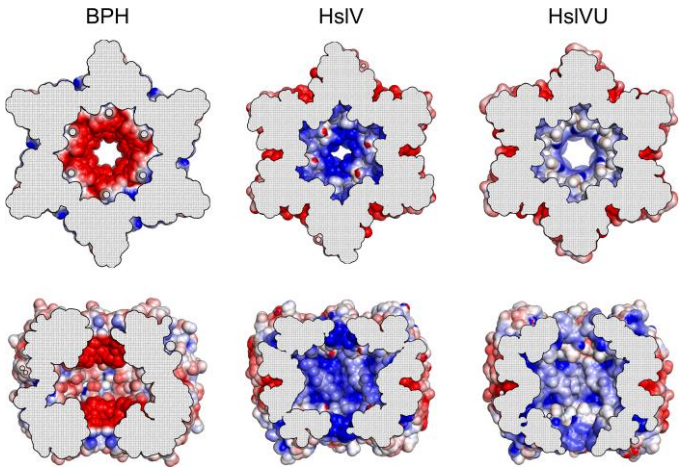


Figure 20: Comparison of BPH and HslV inner cavity characteristics

Top- and side-views of cut-open rings from Cm-BPH, *H. influenzae* HslV (PDB 1G3K), and *H. influenzae* HslV in the presence of HslU (HslVU; PDB 1G3I) are shown with Poisson-Boltzmann electrostatic potentials (±5 kT/e) plotted on the surface (Unni et al., 2011). Negative potentials are visualized in blue and positive potentials in red.

Active site characteristics of BPH – Serine as an alternative catalytic residue

The prototypical BPH active sites display all the features that are characteristic for proteasome-like proteins. Prominent residues in Cm-BPH are Thr-1, Asp-17 and Lys-37 (Figure 18D), with Thr1 acting as catalytic nucleophile. Lys-37 corresponds to Lys-33 in the proteasome, its function is to facilitate deprotonation of Thr-1 (see chapter 3.1; (Huber et al., 2016)). Asp-17 assists in this reaction by orienting Lys-37 and making it prone to protonation. Equally present are e.g. conserved Gly-49, which helps to provide the oxyanion hole that stabilizes the transition state, and Ser-138, which forms hydrogen bonds with Th-r1. The S1 pocket, determining substrate specificity (Huber et al., 2012), is of hydrophobic character, priming it for the accommodation of uncharged amino acids. An interesting and highly unusual feature is the presence of Ser as N-terminal residue in all BPH proteins outside the β-proteobacteria. Aside from that peculiarity, the other active site residues in these cases conform to the typical proteasome family pattern (Figure 18D). To our

knowledge, apart from the aforementioned BPH proteins, Ser occurs in only two proteasome family members as catalytic residue, one Anbu representative (Amb0901 from *Magnetospirillum magneticum AMB-1*), and one HslV protein (CodW from *Bacillus subtilis* (Rho et al., 2008a)), which is active only in the presence of its partner ATPase HslU (CodX) (Rho et al., 2008a). In enzymatic activity assays using a wide range of standard substrate peptides and the intrinsically disordered protein casein, we found Cm-BPH and Td-BPH to be inert (data not shown), even though mass spectrometry (Figure 21A) and crystal structures (Figure 18B and C) show that Thr-1 is exposed and not modified. The apparent lack of protease activity could be explained by the need for an as yet unknown cofactor, required to allosterically activate BPH or to deliver substrate proteins in a conformation that we do not mimic *in vitro*. In addition, BPH might be less promiscuous than the housekeeping proteases with a very limited substrate spectrum. To ascertain the functional nature of the BPH active site, we covalently modified it with the proteasome-specific inhibitor epoxomicin. This compound features an epoxyketone electrophilic trap, the warhead, which renders it highly selective towards the catalytic Thr-1 of proteasome-like proteins (see chapter 3.1; (Schrader et al., 2016)). The ligand binding affinity resides in a peptide-like moiety, resembling Ile-Ile-Thr-Leu side-chains. When we incubated Cm-BPH and Td-BPH with epoxomicin or the inhibitor MG132, we readily obtained modified protein, as seen in mass spectra (Figure 21A), and a crystal structure with bound ligand (Figure 21B). Comparable co-structures with epoxomicin are available only for the proteasome, but not for HslV or Anbu (Groll et al., 2000). The crystal structure of Td-BPH with bound epoxomicin is consistent with productive binding of the inhibitor. The peptide-like moiety lines up in the substrate binding pocket and contacts the same type of relevant residues as in the proteasome (Huber et al., 2012). In particular, the modification of Td-BPH Thr-1 by the epoxomicin warhead confirms this residue as the active nucleophile in the catalytic reaction cycle. Thus, all prerequisites for functionality are met by the layout of the BPH active site. Interestingly though, in the ligand-free form, the Cm-BPH acidic pore loop appears to restrict access to the active site (Figures 21C and D). While this region is unresolved in the ligand-free Td-BPH structure, a more defined pore loop electron density is observed upon epoxomicin conjugation, revealing a distinct conformation that renders the active site now more accessible (Figure 21B) and at the same time

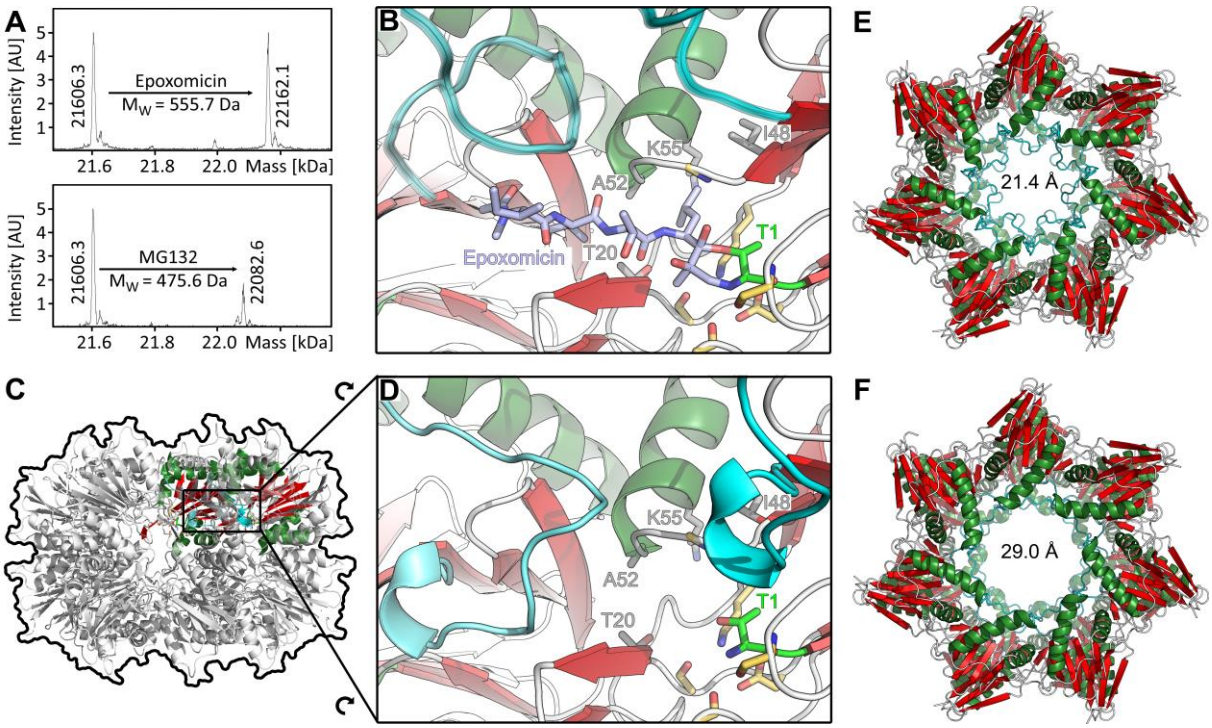


Figure 21: Binding of the proteasome-specific inhibitor epoxomicin to the Td-BPH active site is accompanied by structural transitions in the pore loop

(A) Epoxomicin and MG132 modify the Cm-BPH active site. Cm-BPH was treated with the proteasome-specific inhibitors MG132 or epoxomicin and analyzed by mass spectrometry. The shifts in protomer masses correspond exactly to the molecular masses of the respective inhibitors.

(B) Crystal structure of epoxomicin-conjugated Td-BPH with catalytic nucleophile Thr1 (green), residues thought to be crucial for catalysis (yellow), and residues forming the substrate specificity pocket S1 (grey) represented as sticks. The active site Thr-1 is inhibited through formation of a morpholino adduct, as described for the proteasome (Groll et al., 2000). The Leu side-chain N-terminal of the reactive epoxyketone moiety is coordinated by the Td-BPH S1-pocket residues. The pore loop region (cyan) is resolved in a conformation that permits access to the active site.

(C,D) Side-view representation and close-up view of the ligand-free Td-BPH active site. As the pore loop is not resolved in the ligand-free Td-BPH crystal structure, an overlay with the same region of Cm-BPH (cyan) is shown to indicate possible structural changes in the loop. In the depicted conformation, substrate access to the active site would be hampered.

(E,F) Top-view representations of epoxomicin-modified Td-BPH (E) and ligand-free Td-BPH (F), the latter one with a modeled Cm-BPH pore loop. The inner cavity diameter is indicated.

narrows the inner cavity (Figure 21E and F). As it seemed possible from these results that our inactive BPH proteins are frozen in a closed conformation in solution, we sought to obtain versions with different pore properties. However, when we replaced the acidic loops in Cm-BPH and Td-BPH with small uncharged amino acid residues, the BPH mutant proteins still assembled into stable double-ring complexes but showed no activity (data not shown). Nonetheless, given that the acidic pore loops can adopt different states that vary in the accessibility of the active site, a gating function for this structure motif remains an attractive hypothesis to explore in the future.

Conclusions

Our structural analysis of BPH proteins has revealed the formation of homo-oligomeric barrels with shielded active sites at the inside of the cavity. Despite its descendance from dodecameric HslV, BPH proteins have acquired a sevenfold symmetry like the proteasome, indicating inherent flexibility of the proteasome fold to undergo symmetry rearrangements and to form different types of assemblies. Early in descendance from HslV, the initial co-occurrence in the same cell might have made it necessary to structurally distinguish BPH and to keep HslU interaction reserved for HslV. This would have BPH allowed to move towards ATP-independent functions, for which HslV seems to have little or no propensity (Yoo et al., 1996; Yoo et al., 1997). The evolution of the distinct acidic pore loops with a potential gating function can be seen as part of these mechanistic and structural changes. We may miss ATP-independent functions in the universally used standard activity assays, that do not comprehensively cover all possible binding and cleavage specificities. Importantly, also the 20S proteasome typically exists in a latent form that first must be tweaked with the detergent SDS to show significant activity in these assays (Shibatani and Ward, 1995; Yamada et al., 1995). A rather idiosyncratic BPH activity is also suggested by the presence of Ser1 as catalytic residue in several BPH proteins. It was shown that Ser1 is less efficient compared to Thr1 in proteasomes because of its unfavorable orientation towards incoming substrates (Huber et al., 2016). As this requires repositioning of the N-terminal residue, substrate turnover is expected to be delayed. For ClpP, it was recently shown that the protease exhibits remarkable specificity for certain types of

fluorescent peptides *in vitro*, which is not mirrored in the proteolysis of endogenous substrates, where ClpP seems to be more promiscuous (Gersch et al., 2016). Effective local concentrations of active sites and substrate in the protease barrel reaching the high mM range have been proposed to cause this relaxed cleavage specificity (Gersch et al., 2016). This could also apply to BPH and Anbu, for which no *in vitro* activity has been obtained either (see chapter 3.1). Consequently, future studies will have to focus on the identification of the physiological substrate spectrum of BPH.

3.3 Ubiquitin binding mode and cleavage mechanism of the Rpn11 protease in an ancestral archaeal ubiquitin conjugation system

While previous chapters focused on novel proteolytic complexes, the following chapter comprises the first study of an archaeal (de-) ubiquitination system. Ubiquitin (Ub) functions as a posttranslational modification by regulating interactions with other proteins, including the proteasome. In eukaryotes, Ub is linked to its substrates via a cascade of E1, E2 and E3 enzymes and removed by action of zinc-dependent JAMM (JAB1/MPN/Mov34) proteases, such as the proteasomal Rpn11 (Figure 3; (Hochstrasser, 2009)). This system was long thought to be restricted to eukaryotes and to have evolved from simpler prokaryotic pathways, such as archaeal SAMPylation (see chapter 1.3). This view was challenged with the discovery of *Caldiarchaeum subterraneum* and Asgard archaea, whose genomes encodes for eukaryote-type Ub modifier systems, suggesting that protein ubiquitination might already have evolved in archaea.

C. subterraneum Ub (CsUb), E1, E2 and E3 are organized in a single operon, adjacent to a second E3 protein and a JAMM protease (CsRpn11) with high sequence similarity to Rpn11. In eukaryotes, this deubiquitinase is positioned in the proteasome lid, where it forms intimate contacts with an inactive homolog, Rpn8 (Dambacher et al., 2016). This organization is retained in paralogous eukaryotic structures, known as PCI-complexes (Proteasome - COP9 signalosome - Initiation factor 3), but has to date not been found in archaea (Cavadini et al., 2016; des Georges et al., 2015). While the structures of lid-associated and isolated Rpn11-Rpn8 have been determined, analysis of the exact Ub-binding mode is hampered by the lack of an Rpn11-Ub co-structure (Dambacher et al., 2016; Pathare et al., 2014; Worden et al., 2014). Valuable insights came from the respective co-structure of a monomeric Rpn11 homolog, AMSH-Ub, though it was anticipated that the interacting regions in both deubiquitinases use

different strategies to position Ub for cleavage (Sato et al., 2008; Shrestha et al., 2014; Worden et al., 2014).

In this chapter, we analyze the evolutionary relationships of Rpn11 and Ub proteins, revealing that CsRpn11-CsUb is indeed ancestral to the eukaryotic system. We show, that CsRpn11 processes isopeptide-linked substrates and analyze the cleavage site by mutational studies. Finally, our CsRpn11-CsUb co-structure provides novel insights in the evolutionary conserved Ub binding mode and cleavage mechanism.

Rpn11-like proteins co-evolved with the ubiquitin modifier system in archaea

To investigate the evolutionary relationship of JAMM proteases, we searched for homologs in the non-redundant protein sequence database and clustered them by pairwise sequence similarity (Figure 22A). In the resulting map, the active PCI-complex deubiquitinases (CSN5, Rpn11; eIF3H is inactive) form a single cluster, that is located central between the other eukaryotic clusters and forms the best connections to prokaryotic sequences. Further removed are the inactive PCI-complex subunits CSN6, eIF3F and Rpn8, that also cluster in close vicinity to one another, highlighting the paralogous relationship between Rpn11-Rpn8, eIF3H-eIF3F and CSN5-CSN6.

In contrast to the well-characterized eukaryotic proteins, many prokaryotic JAMMs are of unknown function. Recently, two groups have been described (Cao et al., 2017): JAMM-2 lacks independent activity, whereas isolated JAMM-1 can act as desampylase. While the clusters formed by these sequences are located distant from eukaryotic clusters, our cluster analysis reveals a third group (JAMM-3) with higher similarity to eukaryotic Rpn11 and CSN5 but with poor connections to other eukaryotic JAMMs. In contrast to these, the eukaryote-like JAMMs from *C. subterraneum* and Asgard are located central between Rpn11, Brcc36 and AMSH and are typically more similar to each of these clusters than Rpn11, Brcc36 and AMSH are to each other. With up to four such Rpn11-like sequences in Asgard genomes, a radiation resulting

in the evolution of Rpn11, AMSH and BRCC36 might have already started at the time of archaea.

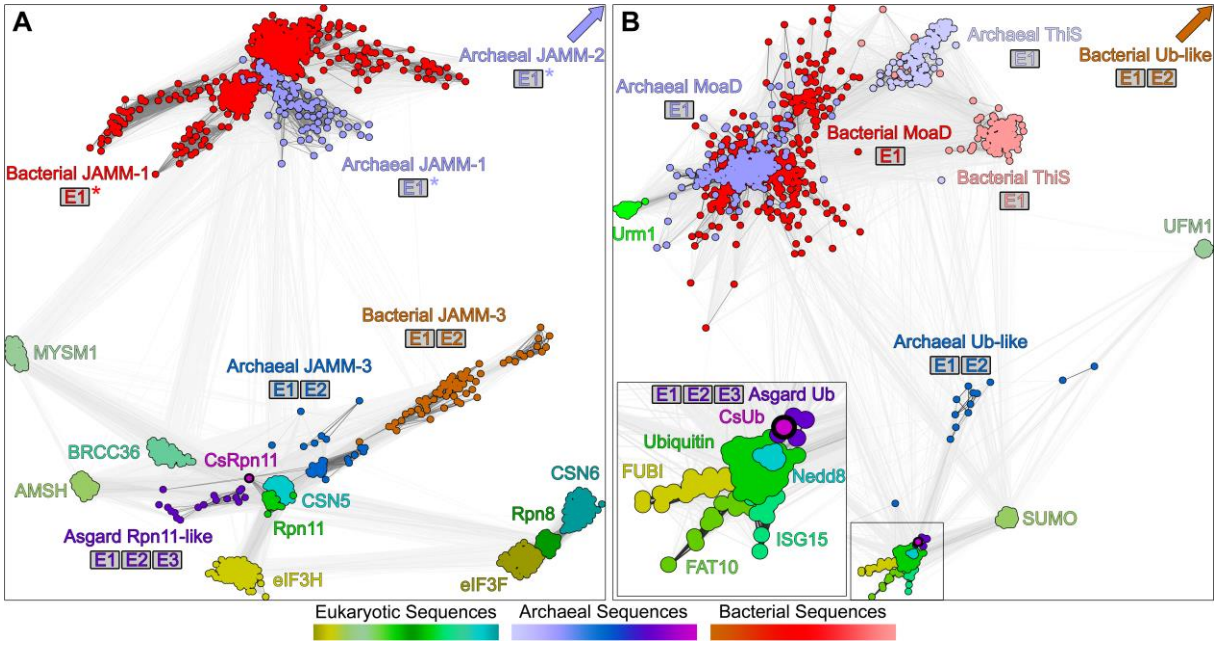


Figure 22: Cluster analysis of JAMM and Ub-like proteins

(A) A cluster map of 1543 JAMM core domain sequences, with a maximum pairwise identity of 70%, was prepared using CLANS (Frickey and Lupas, 2004) based on their all-against-all pairwise similarities as measured by BLAST p-values. Sequences are represented by dots and the line coloring reflects BLAST p-values; the darker a line, the lower the p-value. For simplicity, the map does not include eukaryotic sequences of the prp8, KIAA0157 and FLJ14981 type as well as bacterial lysM-associated JAMM sequences. For prokaryotic JAMMs, the genomic or functional (*) association with E1/E2/E3 proteins is shown. MYSM1 is a histone H2A deubiquitinase, BRCC36 and AMSH are K63-specific deubiquitinases and CSN5 removes Ub-like Nedd8 (Eletr and Wilkinson, 2014).

(B) A cluster map of 1473 Ub-like sequences, prepared in a similar fashion as the JAMM cluster map. UFM, Urm1, Sumo, Nedd8, ISG15, FAT10 and FUBI are eukaryotic Ub-like modifiers that can be conjugated to target proteins in a similar fashion as Ub (Burroughs et al., 2012). MoaD and ThiS function in prokaryotic cofactor synthesis, but can also modify target proteins in some archaea (Dantuluri et al., 2016).

Remarkably, while the prokaryotic JAMMs that cluster distant from Rpn11 are typically functionally associated with just a Ub- and E1-homolog, sequences in the clusters that lie closer to Rpn11 are found in an operon with a Ub-, E1, and E2-homologs. Finally, a RING-type E3-homolog completes the classical ubiquitination machinery in the novel Asgard phylum and *C. subterraneum*, which harbor the most Rpn11-like sequences. Thus, it a appears that the step-wise completion of the Ub modification operon was accompanied by increasingly Rpn11-like sequences, with CsRpn11 at the apex of this process.

Likewise, a cluster analysis of Ub-like proteins (Figure 22B) reveals eukaryote-like sequences in E2- and E3- containing archaea, while the other prokaryotes encode for more distant homologs. Urm1, previously proposed as potential Ub ancestor, is found distant from classical Ub and forms the best connections to prokaryotic MoaD and ThiS (Hochstrasser, 2009). The most Ub-like prokaryotic sequences are found in Asgard archaea and Caldiarchaeum, substantiating the interpretation of their operon-structures as closest precursors to the eukaryotic Ub modifier system.

The CsRpn11-CsUb co-structure reveals novel insights in an evolutionary conserved ubiquitin binding and cleavage mechanism

We therefore were interested in how the CsRpn11-Ub pair would function and whether it could participate in the proteasomal degradation machinery. Since *C. subterraneum* is not culturable, we synthesized CsUb (Csub_C1474) and CsRpn11 (Csub_C1473) genes and produced the proteins in *E. coli*. While CsUb could be purified from the soluble fraction of the cell extract, CsRpn11 aggregated at protein concentrations above ~2 mg/ml and was therefore refolded from inclusion bodies. We consequently optimized the buffer conditions for maximum protein stability (pH 7.5, 500 mM NaCl) and created a series of mutants with deletions in the sequence C-terminal of the JAMM domain, CsRpn$^{\Delta 175\text{-}202}$ CsRpn$^{\Delta 167\text{-}202}$ CsRpn$^{\Delta 149\text{-}202}$ and CsRpn$^{\Delta 145\text{-}202}$, of which CsRpn$^{\Delta 149\text{-}202}$ proved most stable (data not shown). Using these constructs, we were able to obtain crystal structures for CsUb, zinc-bound (Holo-) CsRpn11, zinc-free

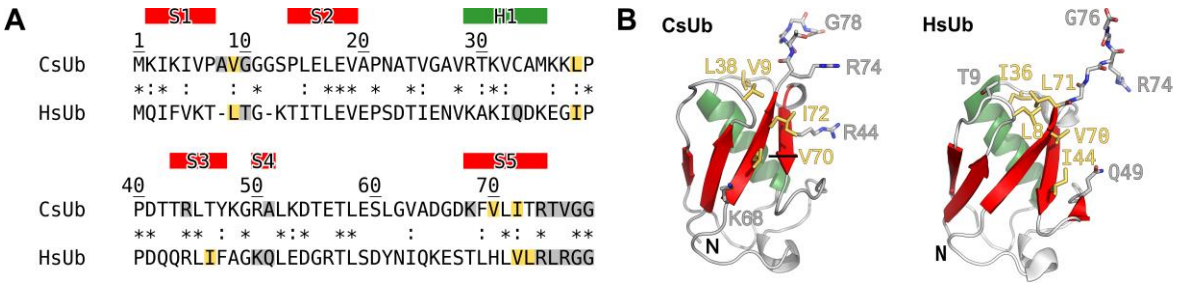


Figure 23: Sequence alignment (A) and structure comparison (B) of CsUb and *homo sapiens* Ub (HsUb; PDB 1D3Z)

The five stranded sheet (red) and the central helix (green) are highlighted. Residues that form important polar (grey) or hydrophobic (yellow) interactions with CsRpn11 or AMSH are shown as sticks (see Figure 24). In the alignment, identical (*) and similar (:) residues are indicated.

(Apo-) CsRpn11$^{\Delta 149\text{-}202}$, and the co-structures Apo-CsRpn11$^{\Delta 149\text{-}202}$-CsUb and Holo-CsRpn11$^{\Delta 149\text{-}202}$-CsUb (Figures 23 and 24). CsUb adopts the classical β-grasp fold, with a structural equivalent for each of the 76 residues of eukaryotic Ub plus a two residue insertion in the first loop region (Figure 23). In agreement with the sequence similarities in the clustermap (Figure 22B), the CsUb structure aligns best with eukaryotic Ub (DALI Z-Score 12.2 for human Ub), but is distinct from so far characterized prokaryotic Ub homologs, such as *H. volcanii* MoaD/SAMP1 (Z = 6.9) or ThiS/SAMP2 (Z = 5.4). In a structure-based sequence alignment (Figure 23A), CsUb shares 31% sequence identity with eukaryotic Ub, but except for K29, all of the eight lysines are located differently. In eukaryotes, these residues are used for formation of poly-Ub chains, which regulate divergent cellular processes (Suryadinata et al., 2014). Though the CsUb lysines are found in different positions, a similar formation of poly-CsUb chains remains conceivable, since other eukaryotic homologs, such as Sumo or Nedd8, also use differently located lysines for these assemblies (Jones et al., 2008; Matic et al., 2008).

CsRpn11 forms a crystallographic dimer, but the dimerization interfaces differ in each determined structure and diverge from eukaryotic Rpn11-Rpn8. Furthermore, light scattering measurements identify only a single mass close to the molecular weight of the monomer (Figure 25). Thus, similar to the monomeric homologs AfJAMM and *S.*

pombe AMSH, which also crystallized as dimers, CsRpn11 appears to exist only as monomer in solution (Ambroggio et al., 2004; Shrestha et al., 2014).

In structure alignments, CsRpn11 displays the highest similarity to eukaryotic Rpn11 (DALI Z-Score 21.1), but is less similar to so far determined structures of other prokaryotic JAMMs, such as *P. furiosus* JAMM-1 (Z = 15.1; (Cao et al., 2017)). Common to all these proteins is a circular core of seven β-strands with the topology S1-S3-S2-S4-S5-S6-S7-S1, that is flanked by three helices (Figure 24). This region is highly conserved between CsRpn11 and eukaryotic Rpn11 with a shared sequence identity of 33%, but eukaryotic Rpn11 diverges in an insertion between S6 and S7 known as ins-2 ((Worden et al., 2014); Figure 24A and 3B) and in its longer helices at the C-terminus (not resolved in the crystal structure (Figure 24B)). Both regions are highly variable, do not interact with the Ub substrate and function primarily in anchoring Rpn11 within the proteasome lid (Dambacher et al., 2016). Thus, the adaptations in CsRpn11 might be explained with the absence of other lid subunits from archaeal genomes. Nevertheless, the C-terminal helices found in CsRpn11 and Asgard Rpn11 are absent from other prokaryotic JAMM sequences and several observations indicate, that they could present a distinct interaction sites or possibly even a precursor structure to the eukaryotic C-termini: (1) Just like in eukaryotes, the helices protrude from the JAMM core structure and promote aggregation of the isolated protein (Worden et al., 2014). (2) The absence of electron density in the CsRpn11 crystal structure for residues 174-202 indicates a high degree of flexibility, similar to what can be expected for the C-termini of free eukaryotic Rpn11. (3) Some Asgard Rpn11 variants contain both ins-2 and the longer C-terminal helices, although other proteasome lid subunits are absent from their genomes. However, due to the extremely low conservation (see above) of these sequences, it remains unclear whether these features are homologous.

The CsRpn11-CsUb interaction is stabilized by a network of hydrogen bonds and a hydrophobic patch (Figures 23B, 24E and 24F). The overall Ub interface topology is very similar in AMSH and CsRpn11, although the individual contacts are often formed by different amino acid pairs (Figures 23 and 24A). Further contacts with Ub are formed by a region known as ins-1, that has been associated with different Ub binding

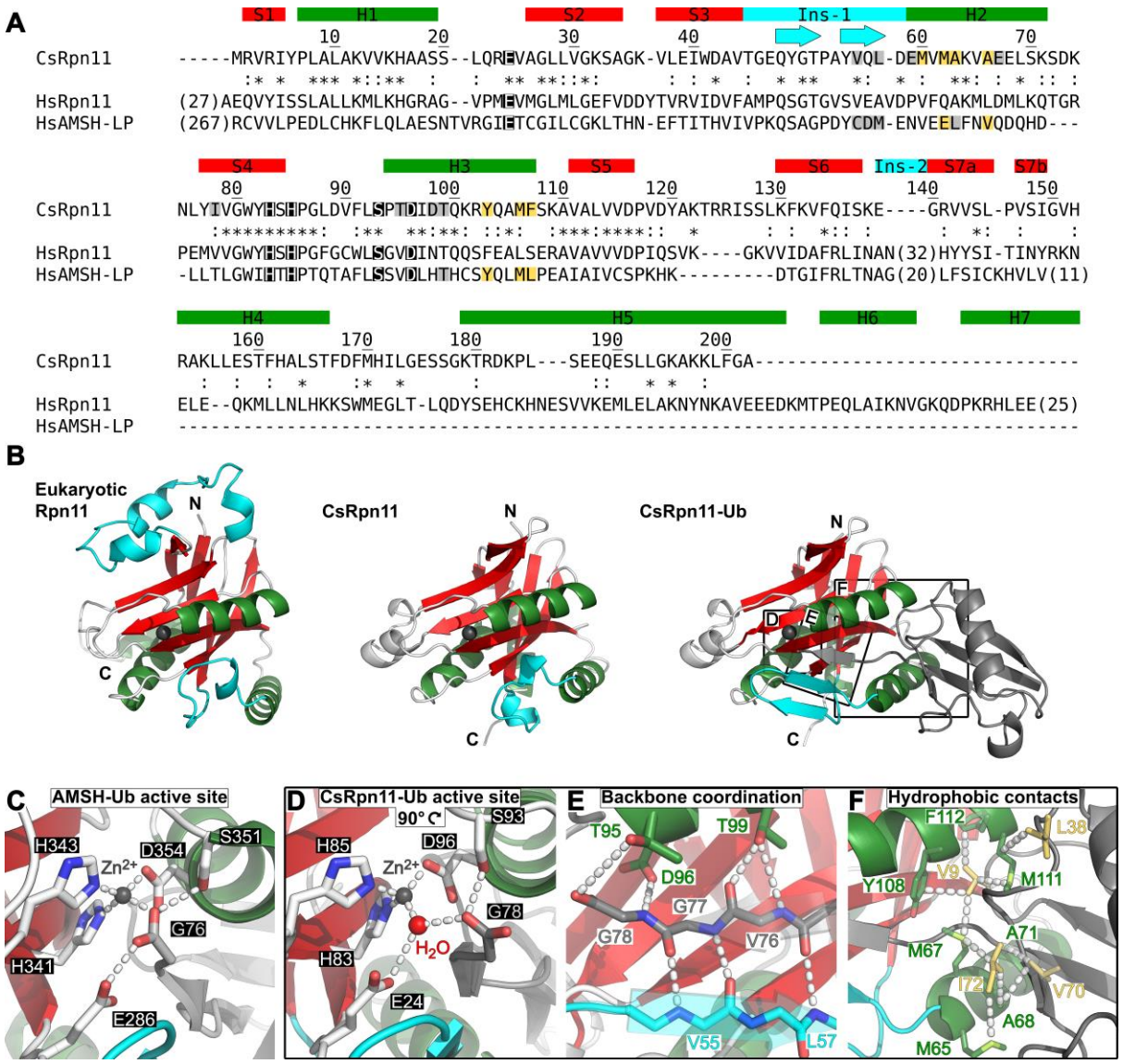


Figure 24: Structure comparison and sequence alignment of *homo sapiens* Rpn11 (HsRpn11), AMSH-LP (HsAMSH-LP) and CsRpn11

(A) Structure-based sequence alignment with the same color scheme as in Figure 23, except that catalytic residues are wrapped in black boxes and ins-1 / ins-2 regions (see main text) are shown in cyan. The Ub-induced β-sheets are indicated by arrows. AMSH ins-2 and C-terminus differ from the same elements in Rpn11 and the respective sequences are therefore not shown in the alignment.

(B) The crystal structures of C-terminally truncated eukaryotic Rpn11 from S. cerevisiae (PDB 4O8X), CsRpn11 and CsRpn11-CsUb with the active site zinc ion shown in black. The ins-2 region, which is not resolved in the high-resolution *S. cerevisiae* Rpn11 structure, is shown as seen in the low-resolution structure of proteasome associated human Rpn11 (PDB 5T0C). Residues 174-202 are not resolved in the CsRpn11 structure and the corresponding region is

also absent from the truncated *S. cerevisiae* Rpn11 structure. CsUb is colored grey and the location of structural elements that are shown in detail in the panel below is indicated by boxes.

(C, D) Close-up views of the AMSH-Ub (PDB 4MSQ) and the CsRpn11-Ub active site in a 90° rotated orientation compared to the structures in (B). Catalytic residues are shown as sticks and polar interactions indicated by dotted lines.

(E, F) Critical interactions between CsRpn11 and CsUb in the same orientation as in (B). Shown are the backbone coordination of the CsUb C-terminus, that results in β-sheet formation and hydrophobic contacts between CsUb and CsRpn11. For simplicity, only side-chains that are involved in the indicated contacts are shown.

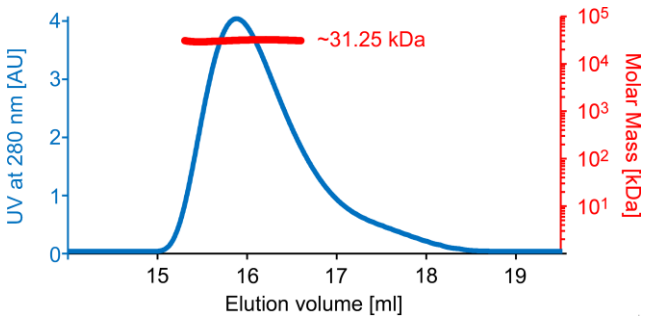


Figure 25. Light scattering profile of CsRpn11.

CsRpn11 was subjected to a size-exclusion column and the scattering data (red) determined for the UV (blue) peak area. The determined mass, 31.25 kDa, is close to the expected size of a single CsRpn11 protein chain (24.5 kDa) and suggests that CsRpn11 exists as monomer in solution.

strategies in eukaryotic JAMMs. In the structures of both AMSH and AMSH-Ub, ins-1 forms a short two-stranded β-sheet, while it adopts an extended loop formation in Ub-free eukaryotic Rpn11 and CSN5 that appears to block access to the active site (Cavadini et al., 2016; Pathare et al., 2014; Sato et al., 2008). Thus, it was anticipated that Rpn11 and CSN5 are autoinhibited by ins-1 and that a remodelling of ins-1 would be required for substrate binding, though it remained unclear to which conformation and by which mechanism. In cryo-EM reconstructions of substrate-engaged proteasome lids a bridging density at the ins-1 region was observed, that was

controversially discussed: While it was interpreted as AMSH-like ins-1 conformation in one study (Matyskiela et al., 2013), it was thought to represent a loop structure reminiscent to that of the isolated Rpn11 ins-1 in another (Worden et al., 2014). Our structures shed new light on the function of ins-1 by showing that a transition between both conformations is possible: In the absence of CsUb, CsRpn11 ins-1 forms an unstructured loop that blocks the active site, just like in eukaryotic Rpn11 and CSN5, but this loop becomes ordered via an induced-fit binding mechanism to form a three stranded β-sheet with the Ub C-terminus, as seen in the AMSH-Ub structure (Figure 24B). We therefore propose, that remodelling of the similar Rpn11 and CSN5 regions functions via the same induced-fit mechanism and consequently, that AMSH, Rpn11 and CSN5 essentially use the same conserved Ub binding mode, that was already established in their archaeal ancestor. In support of this idea, any mutations in Rpn11 ins-1 lead to severe catalytic defects (Worden 2014) and a mutation in the equivalent of CsUb V55 (hsRpn11 V80A), which forms the β-sheet contacts in vicinity of the Ub cleavage motif (Figure 24E), completely abolishes deubiquitinase activity (Worden et al., 2014). Consequently, the Rpn11 ins-1 loop seems to have no autoinhibitory functions but instead appears to participate in Ub binding via a substrate-induced structural transition.

Our CsUb-CsRpn11 structure also shows first catalytically active deubiquitinase-Ub complex with the catalytic water oriented for cleavage (Figure 24D). JAMM proteases use a zinc ion and a glutamate residue to polarize this water molecule, enabling it to perform a nucleophilic attack at the amide bond C-terminal of the Ub Gly-Gly motif (Sato et al., 2008). However, the catalytic water is absent from the AMSH-Ub co-structure and glutamate and zinc form strong contacts to the Ub C-terminal carboxylate instead (Figure 24C; (Shrestha et al., 2014)). This led to the conclusion, that the processed Ub product could remain tightly bound to AMSH and that unconjugated Ub might therefore efficiently inhibit cytosolic AMSH. Possibly, the regeneration of the active site after the cleavage reaction by binding of a new catalytic water molecule facilitates Ub dissociation from the complex by disrupting the contacts between the Ub C-terminal carboxylate, glutamate and zinc. This formation is represented by our CsRpn11-CsUb co-structure, which shows the Ub C-terminus positioned for catalysis, displaced by the catalytic water, which is bound by both Glu-

24 and zinc. Thus, both co-structures might depict two consecutive steps in the reaction cycle.

CsRpn11 exhibits bimodal activity towards ubiquitin precursor and isopeptide-linked ubiquitin conjugates

Seeing how the CsRpn11-CsUb structure depicts ancestral traits of eukaryotic deubiquitinase-Ub interactions, we chose to investigate the catalytic capabilities of this system. The CsUb gene encodes for a Ub precursor with 9 additional residues succeeding the C-terminal Gly-Gly motif, that serves as ubiquitination site and Rpn11 cleavage motif in eukaryotes (Suryadinata et al., 2014). Though most of its physiological substrates are linked to this motif via isopeptide bonds, eukaryotic Rpn11 also processes linear amide bonds at this site (Pathare et al., 2014). Consequently, we studied whether CsRpn11 shares this ability by assaying CsUb precursor (CsUb-pre) cleavage via mass spectrometry. The determined masses show, that CsUb-pre is indeed cleaved precisely after the Gly-Gly motif (Figure 26A), while the other *C. subterraneum* SAMP precursors, Csub_C0702 (ThiS) and Csub_C1603 (MoaD), and eukaryotic diubiquitin variants are not processed (Figure 27). In this respect, CsRpn11 diverges from the promiscuous *P. furiosus* JAMM1, that cleaves all three *H. volcanii* SAMPs and even human Ub (Cao et al., 2017).

To support our interpretation of the CsRpn11 C-terminus as potential interaction site, independent from the catalytic region, we studied its role in CsUb binding and cleavage. To this end, we determined the respective kinetic parameters with full-length CsRpn11 and CsRpn$^{\Delta 149\text{-}202}$ via gel band quantification and microscale thermophoresis (MST; Figures 26B-D). Truncated CsRpn11 processed CsUb-pre with a K_M of 24.2 µM and a K_{cat} of 3.5/min and bound the reaction product with a K_D of 14.6 µM. With a K_M of 31.6 µM and a K_{cat} of 4.0/min, the full-length version displayed similar reaction parameters, indicating, that the C-terminus is indeed not crucially involved in catalysis. Though reaction conditions differ within the assembled proteasome (Dambacher et al., 2016), free eukaryotic Rpn11-Rpn8 processes the linear substrate Ub-TAMRA at similar rates (K_M = 20 µM; K_{cat} = 0.95/min (Worden et

al., 2014)) while AMSH cleaves its preferred substrate, K63-linked diubiquitin, much faster (K_{Cat} = 90/min; K_M = 18.8 µM (Shrestha et al., 2014)).

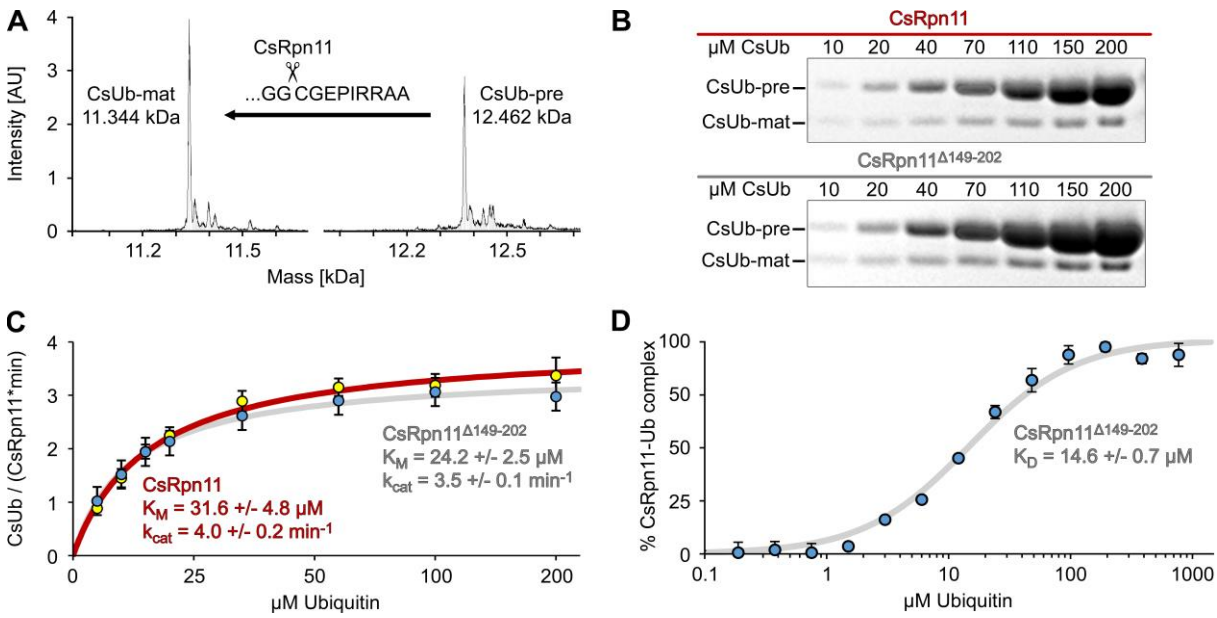


Figure 26: Analysis of CsUb-precursor (CsUb-pre) cleavage by CsRpn11

(A) Mass spectrometrical profile of CsUb-pre before and after incubation with CsRpn11. The mass of the reaction product, mature CsUb (CsUb-mat), corresponds to CsUb-pre without the nine amino acid fragment succeeding the Gly-Gly motif.

(B-C) Sample gels used for KM determination of CsUb-pre processing by full-length and truncated CsRpn11 and the resulting Michaelis-Menten plot.

(D) MST binding curve for KD determination of the CsRpnΔ149-202-CsUb-mat interaction. The respective KD for full-length CsRpn11 could not be determined, due to protein precipitation. All values in C and D represent mean values of three independent experiments, the error bar reflects the standard deviation.

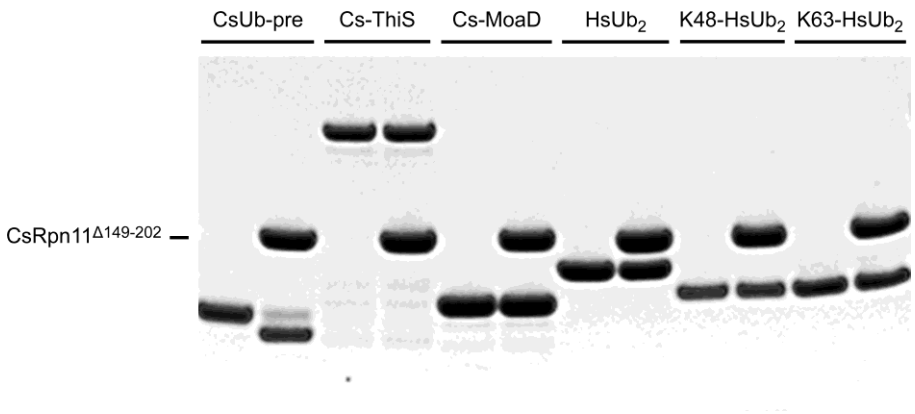


Figure 27: CsRpn11 cleaves CsUb-pre but is inactive towards *C. subterraneum* ThiS (Cs-ThiS), MoaD (Cs-MoaD) and linear, K48-, or K63-linked human diubiquitin ($HsUb_2$)

The gel shows two lanes for each of the indicated substrates, one before (left lane) and one after CsRpn11Δ149-202 treatment (right lane).

Next, we investigated the influence of the active site metal ion. As expected, $CsRpn^{\Delta 149\text{-}202}$ was inactive towards CsUb-pre, when its active site zinc ion was stripped with EDTA (Figure 28A). However, complete activity could be restored upon addition of Zn^{2+} and Fe^{2+}, while Ni^{2+} and Mn^{2+} were less effective and Cu^{2+} could not be used in catalysis.Thus, similar to other metalloproteases (Fukasawa et al., 2011), $CsRpn^{\Delta 149\text{-}202}$ is rather promiscuous in its choice of metal ions. Surprisingly, the reaction speed was dramatically increased in presence of Co^{2+} (K_{cat} = 17.8/min), though the CsUb binding efficiency appeared largely unaffected (K_M = 17.8 µM; Figure 28B). Since a comparison of the cobalt- with the zinc-bound CsRpn-CsUb structure revealed no differences in the active site geometry, we could find no structural rationale for the cobalt effect (data not shown). However, in experiments with the metal-chelating chromophore PAR (McCall and Fierke, 2000), Zn^{2+} appeared to bind to CsRpn11 with much higher affinity than Co^{2+}, indicating that Zn^{2+} is more likely to present the physiological cofactor (Figure 28C).

In our co-structure (Figure 24B), CsRpn11 specifically interacts with CsUb residues N-terminal of G78 but could be expected to be oblivious to the region C-terminal of the cleavage site. To test these observations, we produced a series of mutations in the CsUb-pre C-terminus. Surprisingly, a G78A mutant was still processed by CsRpn11,

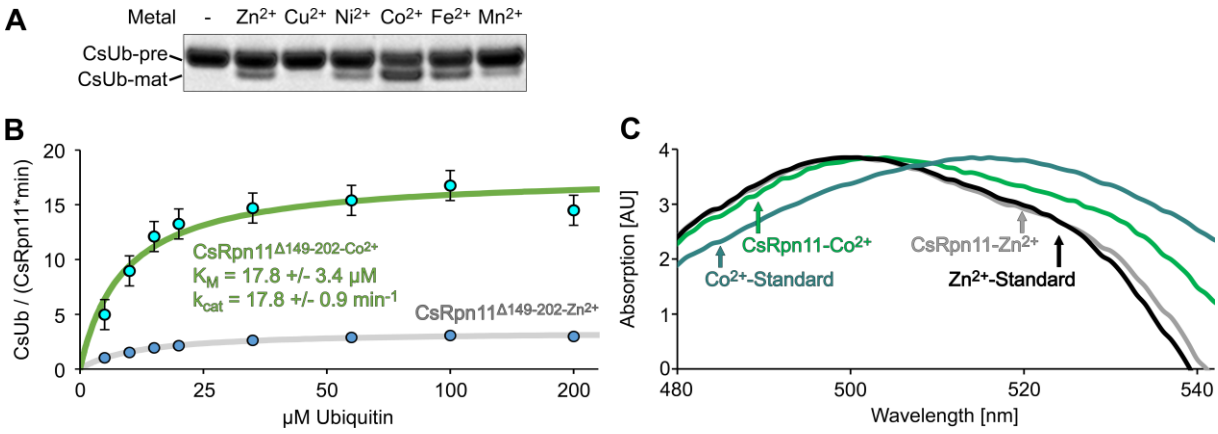


Figure 28: Influence of the CsRpn11 active site metal ion on CsUb-pre cleavage rates and colorimetric analysis of CsRpn11 metal binding characteristics

(A) Protein gel showing CsUb-pre cleavage by CsRpn11Δ149-202 in presence of the indicated metals.

(B) Michaelis-Menten plot of CsUb-pre cleavage by Co2+-bound CsRpn11Δ149-202 in comparison with the Zn2+-bound enzyme (grey, see Figure 26C). All values represent mean values of three independent experiments, the error bar reflects the standard deviation.

(C) The absorption maximum of the metal chelating reagent 4-(2-pyridylazo)resorcinol (PAR) shifts from 415 nm to 500 nm upon Zn2+-binding (black curve) and to 514 nm upon Co2+-binding (blue). This characteristic was used to quantify the metal binding capabilities of EDTA-treated CsRpn11, that was dialyzed against buffers containing either 1 µM Co2+ or Zn2+. Similar absorption values were detected with 8.3 µM Co2+ (blue), 10 µM Zn2+ (black), 10.4 µM Zn2+-treated CsRpn11 (grey) and 16.4 µM Co2+-treated CsRpn11 (green), indicating that Zn2+ but not Co2+ binds stoichiometrically under these conditions. While the spectrum for Zn2+-treated CsRpn11 is virtually identical to that of Zn2+-standard, the curve of Co2+-treated CsRpn11 indicates a mixture of Co2+ and Zn2+, suggesting that the affinity for for Zn2+ is high enough to bind even traces in the ultra-pure buffers.

indicating that the highly conserved Gly-Gly motif is not absolutely required for deubiquitination (Figure 29A). However, larger residues in this position, such as valine (G78V) or glutamate (G78E) prevent cleavage, probably through repulsion by CsRpn11 Thr-101. By contrast, CsUb Gly-77 appears more restricted in our co-structure and consequently, a G77A mutant is processed much slower, while a G77V mutation prevents the reaction. Vice versa, the reaction speed can also be increased by reducing the spatial constraints, as demonstrated for the C79G mutant.

Interestingly, when we analyzed processing of this mutant via mass spectrometry, we observed cleavage both after CsUb G78 and G79 (data not shown). Apparently, the flexibility in this region allows a "wobbling" to occur, that is prevented by a large residue in position 79. Although residues C-terminal of the Gly-Gly motif can be anticipated to protrude from the CsRpn11 structure, they still influence the reaction speed severely, possibly by displacing CsUb Gly-78 through interaction with other elements in CsRpn11. For example, C79E or G80W mutations prevent CsUb-pre processing, while the E81W mutation increases the reaction speed. Therefore, it is conceivable that CsRpn11 mediated deubiquitination is also affected by the characteristics of ubiquitinated substrates in vicinity of the ubiquitination site. In this case, deubiquitination could be prevented until the cleavage site is more accessible, for example through unfolding of the substrate, as it occurs in the proteasome (Ripstein et al., 2017).

To investigate this possibility, we fused CsUb linearly to dihydrofolate reductase (DHFR). Apo-DHFR unfolds at moderate temperatures, but can be stabilized by the interactors NADPH and methotrexate (Clark and Frieden, 1999). Thus, at the melting temperature of the Apo-form, ligand-free DHFR exists in part as molten globule, while the Holo-form remains folded (Figure 29E). Interestingly, CsRpn11 could only process the Apo-form under these conditions, indicating that partial protein unfolding is required for deubiquitination of this substrate (Figure 29B). However, when we introduced a short linker region between CsUb and DHFR, CsRpn11 did no longer discriminate between folded and unfolded substrate, indicating that spatial constraints were responsible for this distinction.

In a related experiment, we produced a similar CsUb fusion protein with lactalbumin. Lactalbumin forms a Proteinase K sensitive molten globule under reducing conditions, but is stabilized through 4 disulfide bridges under oxidizing conditions, rendering the protein largely Proteinase K resistant (Figure 29E; (Ewbank and Creighton, 1993)). By contrast, CsRpn11 deubiquitinated both lactalbumin forms with comparable efficiency, indicating that the substrate folding state is of minor importance in this case. Consequently, unfolding could be required for CsRpn11 mediated deubiquitination of some, inaccessible substrates, but is probably not required for the majority of conceivable CsUb fusions.

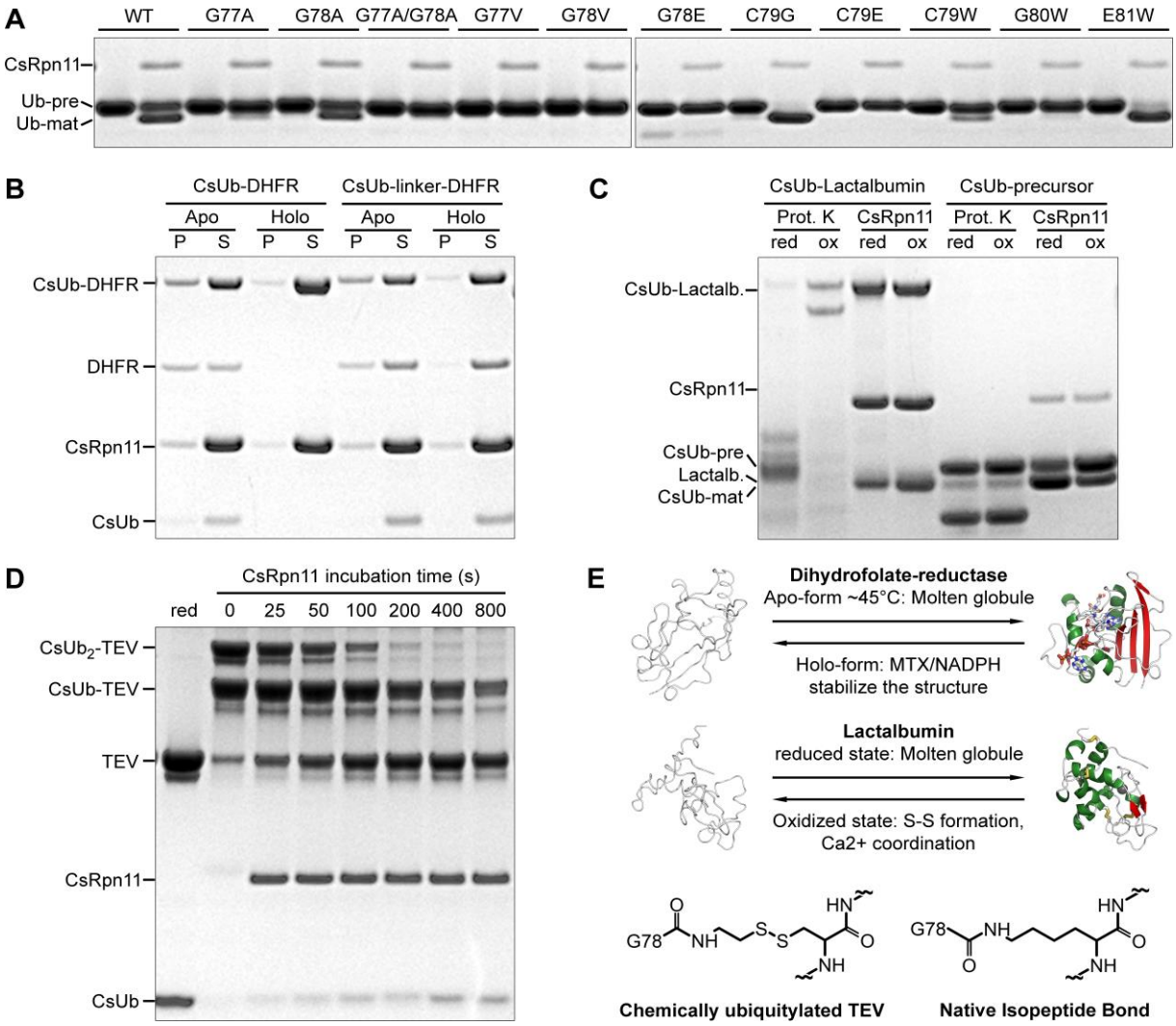


Figure 29: Investigation of determinants in vicinity of the deubiquitination site for CsRpn11 mediated cleavage

(A) Protein gels showing CsUb-pre mutants, each in one lane before and in one lane after CsRpn11Δ149-202 treatment.

(B) Processing of a linear CsUb-DHFR fusion by CsRpn11Δ149-202 in presence (Holo) or absence (Apo) of the stabilizing DHFR interactors MTX and NADPH (Clark and Frieden, 1999). At the assay temperature, Apo-DHFR unfolds and is found partially in the pellet (P), while Holo-DHFR is detected only in the supernatant (S).

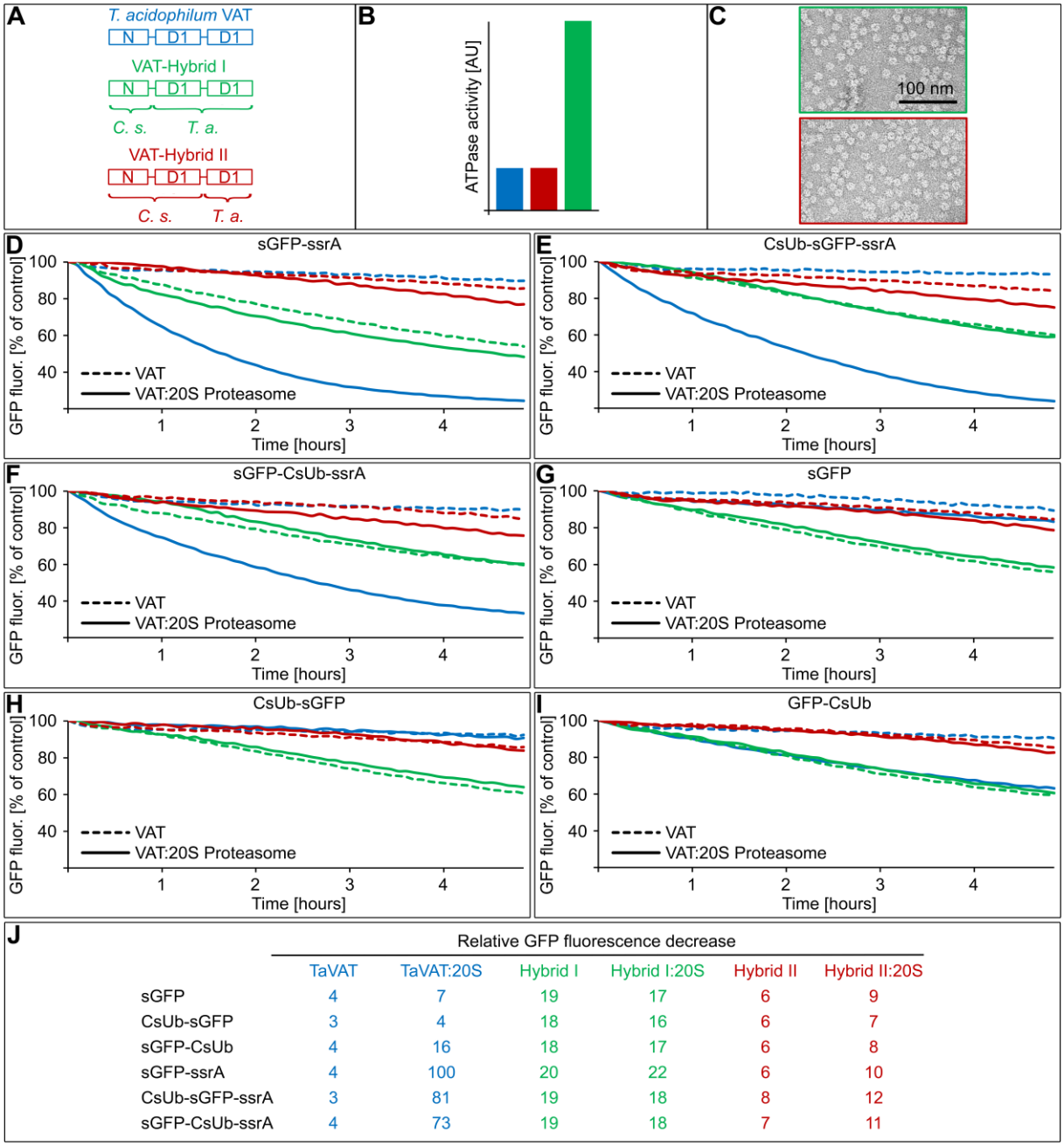
(C) Digestion of a linear CsUb-Lactalbumin fusion and CsUb-pre by CsRpn11Δ149-202 and proteinase K under reducing (red) and oxidizing (ox) conditions. These conditions have only minor effects on the general activity of both proteases, as shown by their effect on CsUb-pre digestion. Lactalbumin did not precipitate throughout the experiment.

(D) The time course of CsRpn11Δ149-202 mediated cleavage of chemically ubiquitylated TEV protease. The disulfide bridge that links CsUb to TEV is disrupted under reducing (red) conditions (lane 1).

(E) Schematical representation of the structural transitions in DHFR (B; PDB 4M6J) and Lactalbumin (C; PDB 1HFZ) substrates and comparison of the cleavage site in chemically ubiquitylated TEV (D) with classical isopeptide bonds (Chen et al., 2010).

Since most physiological CsRpn11 substrates can be expected not to be linked linearly, but via isopeptide bonds, we produced a comparable substrate via an intein fusion technique (Figure 29E; (Chatterjee et al., 2010)). To this end, CsUb was genetically fused to a C-terminal intein tag, that could be cleaved with cysteamine. CsUb-Cysteamine was then ligated with exposed cysteine residues in our chosen model substrate, the TEV protease. The resulting bond differs from native isopeptide linkages only by substitution of the lysine γ-methylene group with a disulfide bridge. This procedure produced mono- and diubiquitinylated TEV proteases (CsUb-TEV and $CsUb_2$-TEV), possibly linked via the well accessible Cys-130 and the partially buried Cys-151 (Nunn et al., 2005). We treated this mixture with CsRpn11 and visualized the time course of the reaction via SDS-PAGE (Figure 29D). The gel shows, that unconjugated CsUb and TEV protease concentrations increased over time, indicating that CsRpn11 also cleaves isopeptide-linked CsUb conjugates. The band corresponding to the size of $CsUb_2$-TEV disappears rapidly, indicating that at least one Ub linkage type is removed efficiently. This produces CsUb-TEV, explaining why the corresponding band appears to remain constant at first and then is reduced slowly. Therefore, it appears that one well accessible CsUb, possibly linked to TEV Cys-130, is cleaved off quickly by Rpn11, while a second less accessible CsUb, possibly linked to TEV Cys-151, is removed less efficiently due to steric hindrance. In agreement with our previous results (Figure 26B), these observations support the notion that CsRpn11 activity is dependent on the substrate characteristics in vicinity of the deubiquitination site.

Given that CsRpn11 appears to deubiquitinate most folded substrates, we wondered how its activity is regulated. Isolated eukaryotic Rpn11 might actually share this characteristic, yet through organization within the proteasomal lid its activity remains



J

	Relative GFP fluorescence decrease					
	TaVAT	TaVAT:20S	Hybrid I	Hybrid I:20S	Hybrid II	Hybrid II:20S
sGFP	4	7	19	17	6	9
CsUb-sGFP	3	4	18	16	6	7
sGFP-CsUb	4	16	18	17	6	8
sGFP-ssrA	4	100	20	22	6	10
CsUb-sGFP-ssrA	3	81	19	18	8	12
sGFP-CsUb-ssrA	4	73	19	18	7	11

Figure 30: Investigation of CsUb-mediated proteasome targeting and the requirement of CsUb removal for proteasomal degradation

(A) Outline of the domain structure in the designed *C. subterraneum* VAT hybrids.

(B) In activity assays, Hybrid II (red) hydrolyzes ATP with similar efficiency as *Thermoplasma acidophilum* VAT (blue), but Hybrid I (green) shows elevated ATPase activity. Since a similar phenotype had been described for a *T. acidophilum* VAT ΔN mutant (Gerega et al., 2005), this effect can possibly be ascribed to an abnormal conformation of the N-terminal domain in respect to the first ATPase module or an inefficient interaction between both.

(C) Electron micrographs show typical ring structures for Hybrid I (green) and II (red) complexes, reminiscent of those formed by *T. acidophilum* VAT (Rockel et al., 2002).

(D-I) Plots showing the effect of VAT (solid lines) or VAT:20S Proteasome complexes (dashed lines) on the fluorescence levels of the indicated superfolder GFP (sGFP) substrates (Pedelacq et al., 2006). The ssrA-tag is a sequence of ten amino acids that has been shown to increase *T. acidophilum* VAT activity towards GFP (Gerega et al., 2005). Each reaction was assayed in presence or absence (control) of ATP and the shown curves represent the ratio between both values. The color scheme is the same as in (A-C): Hybrid I curves are shown in red, Hybrid II in green and *T. acidophilum* VAT curves in blue.

(J) Table summarizing the relative fluorescence decrease velocities of the experiments shown in (D-I). All tested VAT proteins decrease fluorescence levels in each experiment, indicating that they unfold each substrate, although with varying efficiency. The presence of *T. acidophilum* 20S proteasome causes fluorescence to decrease faster for the *T. acidophilum* VAT (TaVAT) and Hybrid II samples, indicating that these unfoldases can catalyze GFP degradation in complex with the proteasome. Not surprisingly, the native *T. acidophilum* VAT:20S complex is more efficient in this respect than the complex with the artificial Hybrid II. By contrast, Hybrid I alone appears more active than the other ATPases (see also (B)), but cannot form a productive complex with the proteasome. Consequently, the Hybrid II:20S proteasome complex can serve as a model to study the interaction between CsUb and *C. subterraneum* VAT:20S. In analogy to eukaryotic Ub, we anticipated two potential effects for CsUb: (1) CsUb could target substrates to the proteasome (Hochstrasser, 2009), possibly by interacting with the VAT N-terminal domain and potentially leading to elevated unfoldase rates. (2) Proteasomal degradation could be stalled by the unfoldase-resistant Ub (Verma et al., 2002) and only continue upon addition of CsRpn11. However, in our experiments with Hybrid II, CsUb fusions alter neither unfolding nor degradation rates significantly. Thus, it remains unclear, whether the CsRpn11-CsUb is associated with the proteasome system or used in different processes.

tightly regulated (Dambacher et al., 2016). With the absence of other lid subunits from archaea, an alternative interaction could be formed between the N-terminal domain of the proteasomal AAA unfoldase and CsRpn11 or CsUb. Since the putative proteasomal *C. subterraneum* AAA, Csub_C0080 and Csub_C0800, could not be obtained in soluble form, we assayed this interaction with hybrid AAA, in which one or both ATPase modules in each subunit were substituted with the corresponding domain of the *T. acidophilum* ortholog (Figure 30A). These hybrid proteins assemble to hexameric rings, that are indistinguishable from the wild type *T. acidophilum* protein in electron microscope (EM) images and retain ATPase as well as GFP unfoldase activity

(Figure 30B and C; (Gerega et al., 2005)). Using these constructs, we assayed whether CsUb would target GFP to the proteasome and whether CsUb needs to be removed prior degradation, but could detect none of these eukaryote-type characteristics (Figure 30, D-I). Moreover, we could detect no interaction between hybrid AAA and CsRpn11 or CsUb via MST or in pulldown experiments (data not shown). However, as Ub interactions are often characterized by dissociation constants in the range of several hundred µM and given our artificial system, these results cannot rule out an involvement of CsRpn11-CsUb in proteasomal degradation ((Hurley et al., 2006)). Just like for the discovery of SAMP-directed proteasomal degradation, an accessible in-vivo system might be required to ultimately clarify this question (Humbard et al., 2010).

Conclusions

While Ub tagging was long thought to be restricted to eukaryotes and to have evolved from distant prokaryotic homologs, such as the ThiS or MoaD conjugation system (Hochstrasser, 2009), our sequence analysis demonstrates that a far more eukaryote-like Ub modifier operon was assembled in archaea by co-evolution with Rpn11-like JAMM proteases. However, whereas the involved proteins radiated in eukaryotes, this archaeal operon is usually found just as a single copy. Possibly, the fundamental function of this system, protein tagging, was already established in archaea, but not used in such a vast amount of processes as in eukaryotes. The physiological relevance of Ub-tagging in archaea remains unresolved, since organisms with the complete system, *C. subterraneum* and Asgard archaea, are currently not culturable. With only two E3 proteins compared to hundreds in eukaryotes, the range of functions could be assumed to be very limited. However, proteins in related archaeal systems have been shown to be remarkably versatile. For example, *H. volcanii* E1 activates SAMPs for modification of protein functions, for sulfur transfer and for use as a degradation tag (Hepowit et al., 2016). In a similar fashion, CsUb and CsRpn11 could present versatile precursors with functions, that are distributed among more specialized proteins in modern eukaryotes (Figure 22). To pursue these functions, CsRpn11 might not be stably associated with a single complex, as eukaryotic Rpn11

is in the proteasome, but reversibly interact with different partners, possibly through its C-terminus.

Overall, the *C. subterraneum* Ub system bridges the gap between archaeal cofactor synthesis and modern eukaryotic Ub tagging. Since the various eukaryotic JAMMs appear to have radiated from a CsRpn11-like precursor (Figure 22), CsRpn11 can serve as a model for this protein family. Indeed, our structural analysis provides new insights in the general Ub binding mode of eukaryotic deubiquitinases by revealing induced-fit structural transitions of the ins-1 region, contrasting the previously anticipated autoinhibitory function of this element.

4 Concluding remarks

The results presented in this work suggest that prokaryotic protein degradation is more multifaceted than previously thought and delineate the links between prokaryotic and eukaryotic systems. Through a combination of structural and bioinformatic analyses, we propose a scenario, in which Anbu's open-ring assembly represents an evolutionary intermediate at the origin of self-compartmentalization. The characterization of BPH revealed a second unprecedented structure and suggested an ATP-independent proteolytic function, that is distinct from known self-compartmentalizing proteases. Finally, we identified *C. subterraneum* Rpn11 and Ub as precursors, from which their eukaryotic homologs radiated and revealed common traits in their respective interactions.

While prokaryotes were originally thought to harbor either proteasome (Archaea and most Actinobacteria) or HslV (found in all other phyla, except cyanobacteria), the discovery of Anbu and BPH expand this picture. Given that the vast majority of prokaryotes are considered microbial dark matter (Rinke et al., 2013), it appears likely that our updated view on the proteasome family will need to be revised again with future discoveries. With a drastically growing number of genomes becoming available, the evolutionary links between proteins, pathways and organisms are unraveled. As one emerging insight from this process, it appears that many proteins are more widespread and possibly more ancient than previously thought. A great number of proteins originally thought to be restricted to eukaryotes were found over the last years in archaea and sometimes even in bacteria (Forterre, 2015; Spang et al., 2015). For example, the recently discovered Asgard archaea feature eukaryotic trafficking machinery components, key components of vesicle biogenesis as well as the here studied Ub system (Zaremba-Niedzwiedzka et al., 2017). While the distribution of some proteins may be influenced through horizontal gene transfer, sequence and phylogenetic analyses suggest that in a growing number of cases, they are more likely to be a result of linear inheritance (Danchin, 2016; Salzberg et al., 2001; Staley and Fuerst, 2017). In this work, we suggest such a model for the occurrence of proteasomes in bacteria and support the finding that ubiquitination emerged in archaea. It will be interesting to observe future findings in this context, e.g. with the

other bacterial Anbu operon proteins (see chapter 1.2) present in *Sulfolobales*, one could anticipate the occurrence of Anbu itself in other yet unstudied archaea - a finding, that would support our evolutionary scenario (Figure 5C).

The proteasome probably did not evolve for targeted protein degradation in the first place and possibly did not even interact with ATP-dependent unfoldases. Instead, it is much more likely that it originated as simple homomeric assembly with functions that required no specific interactors. Remarkably, we can anticipate such functions even today: a considerable portion of the eukaryotic proteasomes act on their own by recognizing intrinsic features of their substrates, for example in oxidatively denatured proteins (Pickering and Davies, 2012a). In these cases, (partially) unfolded substrates are presented by other means, for example through Hsp70, a molecular chaperone, that prevents denatured proteins from aggregating (Reeg et al., 2016). In a similar manner, proteasomal unfoldases have been found to possess chaperone function on their own and the Ub pathway is used for a multitude of non-proteolytic processes (Benaroudj and Goldberg, 2000). Consequently, these components could have originally evolved independent from the proteasome and might have only eventually become associated with it. Nevertheless, a potential function of Anbu as a general protease was originally ruled out based on the finding, that it is not genomically associated with a AAA+ ATPase (Iyer et al., 2009). While the archetypical ATP-dependent protein degradation is indeed exerted by all so-far characterized bacterial self-compartmentalizing proteases and is certainly one main function of the proteasome, a recent study found that only 20% of the cellular proteasomes are associated with the energy-dependent 19S lid, while 50% are "uncapped" and the rest interacts with alternative interactors, such as PA200, PA28 or PI31 (Ben-Nissan and Sharon, 2014; Fabre et al., 2013). Even though the exact numbers may be subject of debate, ATP-independent proteasome functions play an important role in eukaryotic physiology and recent studies could also identify ATP-independent proteasome interactors in bacteria and archaea (Jastrab et al., 2015; Kumoi et al., 2013). In this context, it would appear rather surprising if this principle was not also realized in bacterial proteasome homologs. Indeed, our structural characterization shows that Anbu and BPH are good candidates to adopt such functions. Though we could not elucidate their physiological role, one could speculate, that Anbu and BPH are part of

a bacterial degradation network, in which several specialized proteasome homologs perform functions that are united in the eukaryotic proteasome as a single complex, that is highly adaptable through its various regulators.

5 Contributions

Jörg Martin proposed the projects presented in this work and their experimental investigation would not have been possible without his continuous, substantial support. In a many of discussions, we exchanged ideas, designed experiments and analyzed the results. He also played a major role in the preparation of this thesis by proof-reading the initial drafts and by proposing own ideas, many of which found their way into the manuscript.

Marcus Hartmann solved the crystal structures and analyzed the SAXS data presented in this work. He was also involved in the preparation of the Anbu manuscript, that was used in modified form for the preparation of this thesis. Furthermore, Marcus proposed the concept, that the Anbu open-ring structure presents a precursor of self-compartmentalizing proteases.

Vikram Alva was involved in the interpretation of the bioinformatic analyses presented in the Anbu chapter and in the preparation of the Anbu manuscript; Moritz Ammelburg contributed the original idea to study Anbu; Andrei Lupas proposed to study the Anbu consensus protein and in discussions, he contributed his ideas on proteasome evolution; Lorena Maldoner cloned and purified some of the discussed proteins, studied the BPH pore loop mutants, and performed the Anbu crosslinking experiments; Andriko von Kügelgen purified the *T. acidophilum* proteasome; Matthias Wojtynek prepared the *C. subterraneum* Ub, that was used for structure determination; Shirley Tremel cloned and purified Td-BPH and putative BPH AAA interactors and screened their effect on the catalytic activity of BPH; Reinhard Albrecht and Kerstin Bär set-up the crystallization screens and Reinhard Albrecht loop-mounted the resulting crystals; Matthias Flötenmeyer and Katharina Hipp prepared the electron micrographs; Stefan Grüner performed some of the light scattering experiments presented in the Anbu chapter; Mirita Franz-Wachtel, Boris Macek (Uni Tübingen), Elisabeth Weyher and Victoria Sanchez (MPI Martinsried) performed the mass spectrometry experiments.

I performed all other experiments, wrote this manuscript and prepared all figures in it, for which I received valuable advice from Jörg Martin, Marcus Hartmann and Vikram Alva and for which I made use of the already published Anbu paper.

6 References

Akopian, T.N., Kisselev, A.F., and Goldberg, A.L. (1997). Processive degradation of proteins and other catalytic properties of the proteasome from Thermoplasma acidophilum. J Biol Chem *272*, 1791-1798.

Altschul, S.F., and Koonin, E.V. (1998). Iterated profile searches with PSI-BLAST--a tool for discovery in protein databases. Trends Biochem Sci *23*, 444-447.

Alva, V., Nam, S.Z., Soding, J., and Lupas, A.N. (2016). The MPI bioinformatics Toolkit as an integrative platform for advanced protein sequence and structure analysis. Nucleic Acids Res *44*, W410-415.

Ambroggio, X.I., Rees, D.C., and Deshaies, R.J. (2004). JAMM: a metalloprotease-like zinc site in the proteasome and signalosome. PLoS Biol *2*, E2.

Ammelburg, M., Frickey, T., and Lupas, A.N. (2006). Classification of AAA+ proteins. J Struct Biol *156*, 2-11.

Azim, M.K., Goehring, W., Song, H.K., Ramachandran, R., Bochtler, M., and Goettig, P. (2005). Characterization of the HslU chaperone affinity for HslV protease. Protein Sci *14*, 1357-1362.

Baker, N.A., Sept, D., Joseph, S., Holst, M.J., and McCammon, J.A. (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. Proc Natl Acad Sci U S A *98*, 10037-10041.

Bar-Nun, S., and Glickman, M.H. (2012). Proteasomal AAA-ATPases: structure and function. Biochim Biophys Acta *1823*, 67-82.

Barrett, T., Wilhite, S.E., Ledoux, P., Evangelista, C., Kim, I.F., Tomashevsky, M., Marshall, K.A., Phillippy, K.H., Sherman, P.M., Holko, M.*, et al.* (2013). NCBI GEO: archive for functional genomics data sets--update. Nucleic Acids Res *41*, D991-995.

Beller, H.R., Chain, P.S., Letain, T.E., Chakicherla, A., Larimer, F.W., Richardson, P.M., Coleman, M.A., Wood, A.P., and Kelly, D.P. (2006). The genome sequence of

the obligately chemolithoautotrophic, facultatively anaerobic bacterium Thiobacillus denitrificans. J Bacteriol *188*, 1473-1488.

Ben-Nissan, G., and Sharon, M. (2014). Regulating the 20S proteasome ubiquitin-independent degradation pathway. Biomolecules *4*, 862-884.

Benaroudj, N., and Goldberg, A.L. (2000). PAN, the proteasome-activating nucleotidase from archaebacteria, is a protein-unfolding molecular chaperone. Nat Cell Biol *2*, 833-839.

Bochtler, M., Ditzel, L., Groll, M., Hartmann, C., and Huber, R. (1999). The proteasome. Annu Rev Biophys Biomol Struct *28*, 295-317.

Bochtler, M., Ditzel, L., Groll, M., and Huber, R. (1997). Crystal structure of heat shock locus V (HslV) from Escherichia coli. Proc Natl Acad Sci U S A *94*, 6070-6074.

Burroughs, A.M., Iyer, L.M., and Aravind, L. (2009). Natural history of the E1-like superfamily: implication for adenylation, sulfur transfer, and ubiquitin conjugation. Proteins *75*, 895-910.

Burroughs, A.M., Iyer, L.M., and Aravind, L. (2012). The natural history of ubiquitin and ubiquitin-related domains. Front Biosci (Landmark Ed) *17*, 1433-1460.

Campbell, M.G., Veesler, D., Cheng, A., Potter, C.S., and Carragher, B. (2015). 2.8 A resolution reconstruction of the Thermoplasma acidophilum 20S proteasome using cryo-electron microscopy. Elife *4*.

Cao, S., Engilberge, S., Girard, E., Gabel, F., Franzetti, B., and Maupin-Furlow, J.A. (2017). Structural Insight into Ubiquitin-Like Protein Recognition and Oligomeric States of JAMM/MPN+ Proteases. Structure.

Cavadini, S., Fischer, E.S., Bunker, R.D., Potenza, A., Lingaraju, G.M., Goldie, K.N., Mohamed, W.I., Faty, M., Petzold, G., Beckwith, R.E.*, et al.* (2016). Cullin-RING ubiquitin E3 ligase regulation by the COP9 signalosome. Nature *531*, 598-603.

Cavalier-Smith, T. (2006). Rooting the tree of life by transition analyses. Biol Direct *1*, 19.

Chatterjee, C., McGinty, R.K., Fierz, B., and Muir, T.W. (2010). Disulfide-directed histone ubiquitylation reveals plasticity in hDot1L activation. Nat Chem Biol *6*, 267-269.

Clark, A.C., and Frieden, C. (1999). Native Escherichia coli and murine dihydrofolate reductases contain late-folding non-native structures. J Mol Biol *285*, 1765-1776.

Cowtan, K. (2006). The Buccaneer software for automated model building. 1. Tracing protein chains. Acta Crystallogr D Biol Crystallogr *62*, 1002-1011.

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol *26*, 1367-1372.

Cox, J., Neuhauser, N., Michalski, A., Scheltema, R.A., Olsen, J.V., and Mann, M. (2011). Andromeda: a peptide search engine integrated into the MaxQuant environment. J Proteome Res *10*, 1794-1805.

Dambacher, C.M., Worden, E.J., Herzik, M.A., Martin, A., and Lander, G.C. (2016). Atomic structure of the 26S proteasome lid reveals the mechanism of deubiquitinase inhibition. Elife *5*, e13027.

Danchin, E.G. (2016). Lateral gene transfer in eukaryotes: tip of the iceberg or of the ice cube? BMC Biol *14*, 101.

Dantuluri, S., Wu, Y., Hepowit, N.L., Chen, H., Chen, S., and Maupin-Furlow, J.A. (2016). Proteome targets of ubiquitin-like samp1ylation are associated with sulfur metabolism and oxidative stress in Haloferax volcanii. Proteomics *16*, 1100-1110.

Darwin, K.H., Ehrt, S., Gutierrez-Ramos, J.C., Weich, N., and Nathan, C.F. (2003). The proteasome of Mycobacterium tuberculosis is required for resistance to nitric oxide. Science *302*, 1963-1966.

De Mot, R. (2007). Actinomycete-like proteasomes in a Gram-negative bacterium. Trends Microbiol *15*, 335-338.

De Mot, R., Nagy, I., Walz, J., and Baumeister, W. (1999). Proteasomes and other self-compartmentalizing proteases in prokaryotes. Trends Microbiol *7*, 88-92.

des Georges, A., Dhote, V., Kuhn, L., Hellen, C.U., Pestova, T.V., Frank, J., and Hashem, Y. (2015). Structure of mammalian eIF3 in the context of the 43S preinitiation complex. Nature *525*, 491-495.

Dolinsky, T.J., Czodrowski, P., Li, H., Nielsen, J.E., Jensen, J.H., Klebe, G., and Baker, N.A. (2007). PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. Nucleic Acids Res *35*, W522-525.

Dotsch, A., Schniederjans, M., Khaledi, A., Hornischer, K., Schulz, S., Bielecka, A., Eckweiler, D., Pohl, S., and Haussler, S. (2015). The Pseudomonas aeruginosa Transcriptional Landscape Is Shaped by Environmental Heterogeneity and Genetic Variation. MBio *6*, e00749.

Elias, J.E., and Gygi, S.P. (2007). Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. Nat Methods *4*, 207-214.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr *60*, 2126-2132.

Ewbank, J.J., and Creighton, T.E. (1993). Pathway of disulfide-coupled unfolding and refolding of bovine alpha-lactalbumin. Biochemistry *32*, 3677-3693.

Fabre, B., Lambour, T., Delobel, J., Amalric, F., Monsarrat, B., Burlet-Schiltz, O., and Bousquet-Dubouch, M.P. (2013). Subcellular distribution and dynamics of active proteasome complexes unraveled by a workflow combining in vivo complex cross-linking and quantitative proteomics. Mol Cell Proteomics *12*, 687-699.

Firoved, A.M., Wood, S.R., Ornatowski, W., Deretic, V., and Timmins, G.S. (2004). Microarray analysis and functional characterization of the nitrosative stress response in nonmucoid and mucoid Pseudomonas aeruginosa. J Bacteriol *186*, 4046-4050.

Forouzan, D., Ammelburg, M., Hobel, C.F., Stroh, L.J., Sessler, N., Martin, J., and Lupas, A.N. (2012). The archaeal proteasome is regulated by a network of AAA ATPases. J Biol Chem *287*, 39254-39262.

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

Forterre, P. (2015). The universal tree of life: an update. Front Microbiol *6*, 717.

Fredriksson, A., Ballesteros, M., Dukan, S., and Nystrom, T. (2005). Defense against protein carbonylation by DnaK/DnaJ and proteases of the heat shock regulon. J Bacteriol *187*, 4207-4213.

Frickey, T., and Lupas, A. (2004). CLANS: a Java application for visualizing protein families based on pairwise similarity. Bioinformatics *20*, 3702-3704.

Fu, X., Liu, R., Sanchez, I., Silva-Sanchez, C., Hepowit, N.L., Cao, S., Chen, S., and Maupin-Furlow, J. (2016). Ubiquitin-Like Proteasome System Represents a Eukaryotic-Like Pathway for Targeted Proteolysis in Archaea. MBio *7*.

Fuchs, A.C.D., Alva, V., Maldoner, L., Albrecht, R., Hartmann, M.D., and Martin, J. (2017). The Architecture of the Anbu Complex Reflects an Evolutionary Intermediate at the Origin of the Proteasome System. Structure *25*, 834-845 e835.

Fukasawa, K.M., Hata, T., Ono, Y., and Hirose, J. (2011). Metal preferences of zinc-binding motif on metalloproteases. J Amino Acids *2011*, 574816.

Gerega, A., Rockel, B., Peters, J., Tamura, T., Baumeister, W., and Zwickl, P. (2005). VAT, the thermoplasma homolog of mammalian p97/VCP, is an N domain-regulated protein unfoldase. J Biol Chem *280*, 42856-42862.

Gersch, M., Stahl, M., Poreba, M., Dahmen, M., Dziedzic, A., Drag, M., and Sieber, S.A. (2016). Barrel-shaped ClpP Proteases Display Attenuated Cleavage Specificities. Acs Chem Biol *11*, 389-399.

Gille, C., Goede, A., Schloetelburg, C., Preissner, R., Kloetzel, P.M., Gobel, U.B., and Frommel, C. (2003). A comprehensive view on proteasomal sequences: implications for the evolution of the proteasome. J Mol Biol *326*, 1437-1448.

Glynn, S.E., and Chien, P. (2016). Sending protein aggregates into a downward spiral. Nat Struct Mol Biol *23*, 769-770.

Grana, M., Bellinzoni, M., Miras, I., Fiez-Vandal, C., Haouz, A., Shepard, W., Buschiazzo, A., and Alzari, P.M. (2009). Structure of Mycobacterium tuberculosis

Rv2714, a representative of a duplicated gene family in Actinobacteria. Acta Crystallogr Sect F Struct Biol Cryst Commun *65*, 972-977.

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

Groll, M., Kim, K.B., Kairies, N., Huber, R., and Crews, C.M. (2000). Crystal structure of epoxomicin : 20S proteasome reveals a molecular basis for selectivity of alpha ',beta '-epoxyketone proteasome inhibitors. J Am Chem Soc *122*, 1237-1238.

Gur, E., Biran, D., and Ron, E.Z. (2011). Regulated proteolysis in Gram-negative bacteria--how and when? Nat Rev Microbiol *9*, 839-848.

Gusarov, I., and Nudler, E. (2005). NO-mediated cytoprotection: instant adaptation to oxidative stress in bacteria. Proc Natl Acad Sci U S A *102*, 13855-13860.

Helmann, J.D., Wu, M.F., Kobel, P.A., Gamo, F.J., Wilson, M., Morshedi, M.M., Navre, M., and Paddon, C. (2001). Global transcriptional response of Bacillus subtilis to heat shock. J Bacteriol *183*, 7318-7328.

Hemsley, A., Arnheim, N., Toney, M.D., Cortopassi, G., and Galas, D.J. (1989). A simple method for site-directed mutagenesis using the polymerase chain reaction. Nucleic Acids Res *17*, 6545-6551.

Hepowit, N.L., de Vera, I.M., Cao, S., Fu, X., Wu, Y., Uthandi, S., Chavarria, N.E., Englert, M., Su, D., Sll, D.*, et al.* (2016). Mechanistic insight into protein modification and sulfur mobilization activities of noncanonical E1 and associated ubiquitin-like proteins of Archaea. FEBS J *283*, 3567-3586.

Hervas, A.B., Canosa, I., and Santero, E. (2008). Transcriptome analysis of Pseudomonas putida in response to nitrogen availability. J Bacteriol *190*, 416-420.

Hirano, H., Kimura, Y., and Kimura, A. (2016). Biological significance of co- and post-translational modifications of the yeast 26S proteasome. J Proteomics *134*, 37-46.

Hochstrasser, M. (2009). Origin and function of ubiquitin-like proteins. Nature *458*, 422-429.

Holm, L., Kaariainen, S., Rosenstrom, P., and Schenkel, A. (2008). Searching protein structure databases with DaliLite v.3. Bioinformatics *24*, 2780-2781.

Huber, E.M., Basler, M., Schwab, R., Heinemeyer, W., Kirk, C.J., Groettrup, M., and Groll, M. (2012). Immuno- and Constitutive Proteasome Crystal Structures Reveal Differences in Substrate and Inhibitor Specificity. Cell *148*, 727-738.

Huber, E.M., de Bruin, G., Heinemeyer, W., Soriano, G.P., Overkleeft, H.S., and Groll, M. (2015). Systematic Analyses of Substrate Preferences of 20S Proteasomes Using Peptidic Epoxyketone Inhibitors. J Am Chem Soc *137*, 7835-7842.

Huber, E.M., Heinemeyer, W., Li, X., Arendt, C.S., Hochstrasser, M., and Groll, M. (2016). A unified mechanism for proteolysis and autocatalytic activation in the 20S proteasome. Nat Commun *7*, 10900.

Hug, L.A., Baker, B.J., Anantharaman, K., Brown, C.T., Probst, A.J., Castelle, C.J., Butterfield, C.N., Hernsdorf, A.W., Amano, Y., Ise, K.*, et al.* (2016). A new view of the tree of life. Nature Microbiology *1*, 16048.

Humbard, M.A., and Maupin-Furlow, J.A. (2013). Prokaryotic proteasomes: nanocompartments of degradation. J Mol Microbiol Biotechnol *23*, 321-334.

Humbard, M.A., Miranda, H.V., Lim, J.M., Krause, D.J., Pritz, J.R., Zhou, G., Chen, S., Wells, L., and Maupin-Furlow, J.A. (2010). Ubiquitin-like small archaeal modifier proteins (SAMPs) in Haloferax volcanii. Nature *463*, 54-60.

Hurley, J.H., Lee, S., and Prag, G. (2006). Ubiquitin-binding domains. Biochem J *399*, 361-372.

Iyer, L.M., Abhiman, S., Maxwell Burroughs, A., and Aravind, L. (2009). Amidoligases with ATP-grasp, glutamine synthetase-like and acetyltransferase-like domains: synthesis of novel metabolites and peptide modifications of proteins. Mol Biosyst *5*, 1636-1660.

Janssen, P.J., Van Houdt, R., Moors, H., Monsieurs, P., Morin, N., Michaux, A., Benotmane, M.A., Leys, N., Vallaeys, T., Lapidus, A.*, et al.* (2010). The complete genome sequence of Cupriavidus metallidurans strain CH34, a master survivalist in harsh and anthropogenic environments. PLoS One *5*, e10433.

Jastrab, J.B., Wang, T., Murphy, J.P., Bai, L., Hu, K., Merkx, R., Huang, J., Chatterjee, C., Ovaa, H., Gygi, S.P.*, et al.* (2015). An adenosine triphosphate-independent proteasome activator contributes to the virulence of Mycobacterium tuberculosis. Proc Natl Acad Sci U S A *112*, E1763-1772.

Jones, J., Wu, K., Yang, Y., Guerrero, C., Nillegoda, N., Pan, Z.Q., and Huang, L. (2008). A targeted proteomic analysis of the ubiquitin-like modifier nedd8 and associated proteins. J Proteome Res *7*, 1274-1287.

Kanemori, M., Nishihara, K., Yanagi, H., and Yura, T. (1997). Synergistic roles of HslVU and other ATP-dependent proteases in controlling in vivo turnover of sigma32 and abnormal proteins in Escherichia coli. J Bacteriol *179*, 7219-7225.

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

Koshland, D.E., Jr. (2002). Special essay. The seven pillars of life. Science *295*, 2215-2216.

Kumoi, K., Satoh, T., Murata, K., Hiromoto, T., Mizushima, T., Kamiya, Y., Noda, M., Uchiyama, S., Yagi, H., and Kato, K. (2013). An archaeal homolog of proteasome assembly factor functions as a proteasome activator. PLoS One *8*, e60294.

Kunjappu, M.J., and Hochstrasser, M. (2014). Assembly of the 20S proteasome. Biochim Biophys Acta *1843*, 2-12.

Levine, R.L., Garland, D., Oliver, C.N., Amici, A., Climent, I., Lenz, A.G., Ahn, B.W., Shaltiel, S., and Stadtman, E.R. (1990). Determination of carbonyl content in oxidatively modified proteins. Methods Enzymol *186*, 464-478.

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

Lupas, A., Zwickl, P., and Baumeister, W. (1994). Proteasome sequences in eubacteria. Trends Biochem Sci *19*, 533-534.

Marques, A.J., Palanimurugan, R., Matias, A.C., Ramos, P.C., and Dohmen, R.J. (2009). Catalytic Mechanism and Assembly of the Proteasome. Chem Rev *109*, 1509-1536.

Matic, I., van Hagen, M., Schimmel, J., Macek, B., Ogg, S.C., Tatham, M.H., Hay, R.T., Lamond, A.I., Mann, M., and Vertegaal, A.C. (2008). In vivo identification of human small ubiquitin-like modifier polymerization sites by high accuracy mass spectrometry and an in vitro to in vivo strategy. Mol Cell Proteomics *7*, 132-144.

Matyskiela, M.E., Lander, G.C., and Martin, A. (2013). Conformational switching of the 26S proteasome enables substrate degradation. Nat Struct Mol Biol *20*, 781-788.

Maupin-Furlow, J. (2012). Proteasomes and protein conjugation across domains of life. Nat Rev Microbiol *10*, 100-111.

McCall, K.A., and Fierke, C.A. (2000). Colorimetric and fluorimetric assays to quantitate micromolar concentrations of transition metals. Anal Biochem *284*, 307-315.

McLean, S., Bowman, L.A., Sanguinetti, G., Read, R.C., and Poole, R.K. (2010). Peroxynitrite toxicity in Escherichia coli K12 elicits expression of oxidative stress responses and protein nitration and nitrosylation. J Biol Chem *285*, 20724-20731.

Molina-Henares, M.A., de la Torre, J., Garcia-Salamanca, A., Molina-Henares, A.J., Herrera, M.C., Ramos, J.L., and Duque, E. (2010). Identification of conditionally essential genes for growth of Pseudomonas putida KT2440 on minimal medium through the screening of a genome-wide mutant library. Environ Microbiol *12*, 1468-1485.

Murshudov, G.N., Vagin, A.A., Lebedev, A., Wilson, K.S., and Dodson, E.J. (1999). Efficient anisotropic refinement of macromolecular structures using FFT. Acta Crystallogr D Biol Crystallogr *55*, 247-255.

Nunn, C.M., Jeeves, M., Cliff, M.J., Urquhart, G.T., George, R.R., Chao, L.H., Tscuchia, Y., and Djordjevic, S. (2005). Crystal structure of tobacco etch virus protease shows the protein C terminus bound within the active site. J Mol Biol *350*, 145-155.

Nunoura, T., Takaki, Y., Kakuta, J., Nishi, S., Sugahara, J., Kazama, H., Chee, G.J., Hattori, M., Kanai, A., Atomi, H.*, et al.* (2011). Insights into the evolution of Archaea and eukaryotic protein modifier systems revealed by the genome of a novel archaeal group. Nucleic Acids Res *39*, 3204-3223.

Osbourn, A.E., and Field, B. (2009). Operons. Cell Mol Life Sci *66*, 3755-3775.

Park, E., Lee, J.W., Yoo, H.M., Ha, B.H., An, J.Y., Jeon, Y.J., Seol, J.H., Eom, S.H., and Chung, C.H. (2013). Structural alteration in the pore motif of the bacterial 20S proteasome homolog HslV leads to uncontrolled protein degradation. J Mol Biol *425*, 2940-2954.

Pathare, G.R., Nagy, I., Sledz, P., Anderson, D.J., Zhou, H.J., Pardon, E., Steyaert, J., Forster, F., Bracher, A., and Baumeister, W. (2014). Crystal structure of the proteasomal deubiquitylation module Rpn8-Rpn11. Proc Natl Acad Sci U S A *111*, 2984-2989.

Pedelacq, J.D., Cabantous, S., Tran, T., Terwilliger, T.C., and Waldo, G.S. (2006). Engineering and characterization of a superfolder green fluorescent protein. Nat Biotechnol *24*, 79-88.

Pei, J., Kim, B.H., and Grishin, N.V. (2008). PROMALS3D: a tool for multiple protein sequence and structure alignments. Nucleic Acids Res *36*, 2295-2300.

Perrakis, A., Morris, R., and Lamzin, V.S. (1999). Automated protein model building combined with iterative structure refinement. Nat Struct Biol *6*, 458-463.

Petoukhov, M.V., and Svergun, D.I. (2005). Global rigid body modeling of macromolecular complexes against small-angle scattering data. Biophys J *89*, 1237-1250.

Pickering, A.M., and Davies, K.J. (2012a). Degradation of damaged proteins: the main function of the 20S proteasome. Prog Mol Biol Transl Sci *109*, 227-248.

Pickering, A.M., and Davies, K.J. (2012b). Differential roles of proteasome and immunoproteasome regulators Pa28alphabeta, Pa28gamma and Pa200 in the degradation of oxidized proteins. Arch Biochem Biophys *523*, 181-190.

Porebski, B.T., and Buckle, A.M. (2016). Consensus protein design. Protein Eng Des Sel *29*, 245-251.

Qian, M.X., Pang, Y., Liu, C.H., Haratake, K., Du, B.Y., Ji, D.Y., Wang, G.F., Zhu, Q.Q., Song, W., Yu, Y.*, et al.* (2013). Acetylation-mediated proteasomal degradation of core histones during DNA repair and spermatogenesis. Cell *153*, 1012-1024.

Ramachandran, R., Hartmann, C., Song, H.K., Huber, R., and Bochtler, M. (2002). Functional interactions of HslV (ClpQ) with the ATPase HslU (ClpY). Proc Natl Acad Sci U S A *99*, 7396-7401.

Reeg, S., Jung, T., Castro, J.P., Davies, K.J., Henze, A., and Grune, T. (2016). The molecular chaperone Hsp70 promotes the proteolytic removal of oxidatively damaged proteins by the proteasome. Free Radic Biol Med *99*, 153-166.

Rho, S.H., Park, H.H., Kang, G.B., Im, Y.J., Kang, M.S., Lim, B.K., Seong, I.S., Seol, J., Chung, C.H., Wang, J.*, et al.* (2008a). Crystal structure of Bacillus subtilis CodW, a noncanonical HslV-like peptidase with an impaired catalytic apparatus. Proteins *71*, 1020-1026.

Rho, S.H., Park, H.H., Kang, G.B., Im, Y.J., Kang, M.S., Lim, B.K., Seong, I.S., Seol, J., Chung, C.H., Wang, J.M.*, et al.* (2008b). Crystal structure of Bacillus Subtilis CodW, a noncanonical HsIV-like peptidase with an impaired catalytic apparatus. Proteins *71*, 1020-1026.

Rinke, C., Schwientek, P., Sczyrba, A., Ivanova, N.N., Anderson, I.J., Cheng, J.F., Darling, A., Malfatti, S., Swan, B.K., Gies, E.A.*, et al.* (2013). Insights into the phylogeny and coding potential of microbial dark matter. Nature *499*, 431-437.

Ripstein, Z.A., Huang, R., Augustyniak, R., Kay, L.E., and Rubinstein, J.L. (2017). Structure of a AAA+ unfoldase in the process of unfolding substrate. Elife *6*.

Rockel, B., Jakana, J., Chiu, W., and Baumeister, W. (2002). Electron cryo-microscopy of VAT, the archaeal p97/CDC48 homologue from Thermoplasma acidophilum. J Mol Biol *317*, 673-681.

Ruepp, A., Eckerskorn, C., Bogyo, M., and Baumeister, W. (1998). Proteasome function is dispensable under normal but not under heat shock conditions in Thermoplasma acidophilum. FEBS Lett *425*, 87-90.

Sali, A., Potterton, L., Yuan, F., van Vlijmen, H., and Karplus, M. (1995). Evaluation of comparative protein modeling by MODELLER. Proteins *23*, 318-326.

Salzberg, S.L., White, O., Peterson, J., and Eisen, J.A. (2001). Microbial genes in the human genome: lateral transfer or gene loss? Science *292*, 1903-1906.

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

Schmidt, R., Bukau, B., and Mogk, A. (2009). Principles of general and regulatory proteolysis by AAA+ proteases in Escherichia coli. Res Microbiol *160*, 629-636.

Schneidman-Duhovny, D., Hammel, M., Tainer, J.A., and Sali, A. (2016). FoXS, FoXSDock and MultiFoXS: Single-state and multi-state structural modeling of proteins and their complexes based on SAXS profiles. Nucleic Acids Res *44*, W424-429.

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

Seong, I.S., Oh, J.Y., Lee, J.W., Tanaka, K., and Chung, C.H. (2000). The HslU ATPase acts as a molecular chaperone in prevention of aggregation of SulA, an inhibitor of cell division in Escherichia coli. FEBS Lett *477*, 224-229.

Sheldrick, G.M. (2008). A short history of SHELX. Acta Crystallogr A *64*, 112-122.

Shibatani, T., and Ward, W.F. (1995). Sodium dodecyl sulfate (SDS) activation of the 20S proteasome in rat liver. Arch Biochem Biophys *321*, 160-166.

Shrestha, R.K., Ronau, J.A., Davies, C.W., Guenette, R.G., Strieter, E.R., Paul, L.N., and Das, C. (2014). Insights into the mechanism of deubiquitination by JAMM deubiquitinases from cocrystal structures of the enzyme with the substrate and product. Biochemistry *53*, 3199-3217.

Soding, J. (2005). Protein homology detection by HMM-HMM comparison. Bioinformatics *21*, 951-960.

Sousa, M.C., Trame, C.B., Tsuruta, H., Wilbanks, S.M., Reddy, V.S., and McKay, D.B. (2000). Crystal and solution structures of an HslUV protease-chaperone complex. Cell *103*, 633-643.

Spang, A., Saw, J.H., Jorgensen, S.L., Zaremba-Niedzwiedzka, K., Martijn, J., Lind, A.E., van Eijk, R., Schleper, C., Guy, L., and Ettema, T.J. (2015). Complex archaea that bridge the gap between prokaryotes and eukaryotes. Nature *521*, 173-179.

Staley, J.T., and Fuerst, J.A. (2017). Ancient, highly conserved proteins from a LUCA with complex cell biology provide evidence in support of the nuclear compartment commonality (NuCom) hypothesis. Res Microbiol *168*, 395-412.

Suryadinata, R., Roesley, S.N., Yang, G., and Sarcevic, B. (2014). Mechanisms of generating polyubiquitin chains of different topology. Cells *3*, 674-689.

Tang, G., Peng, L., Baldwin, P.R., Mann, D.S., Jiang, W., Rees, I., and Ludtke, S.J. (2007). EMAN2: an extensible image processing suite for electron microscopy. J Struct Biol *157*, 38-46.

Unni, S., Huang, Y., Hanson, R.M., Tobias, M., Krishnan, S., Li, W.W., Nielsen, J.E., and Baker, N.A. (2011). Web servers and services for electrostatics calculations with APBS and PDB2PQR. J Comput Chem *32*, 1488-1491.

Vagin, A., and Teplyakov, A. (2000). An approach to multi-copy search in molecular replacement. Acta Crystallogr D Biol Crystallogr *56*, 1622-1624.

Valas, R.E., and Bourne, P.E. (2008). Rethinking proteasome evolution: two novel bacterial proteasomes. J Mol Evol *66*, 494-504.

Verma, R., Aravind, L., Oania, R., McDonald, W.H., Yates, J.R., 3rd, Koonin, E.V., and Deshaies, R.J. (2002). Role of Rpn11 metalloprotease in deubiquitination and degradation by the 26S proteasome. Science *298*, 611-615.

Williams, T.A., Foster, P.G., Cox, C.J., and Embley, T.M. (2013). An archaeal origin of eukaryotes supports only two primary domains of life. Nature *504*, 231-236.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G., McCoy, A.*, et al.* (2011). Overview of the CCP4 suite and current developments. Acta Crystallogr D Biol Crystallogr *67*, 235-242.

Winter, J., Linke, K., Jatzek, A., and Jakob, U. (2005). Severe oxidative stress causes inactivation of DnaK and activation of the redox-regulated chaperone Hsp33. Mol Cell *17*, 381-392.

Woese, C.R., Kandler, O., and Wheelis, M.L. (1990). Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. Proc Natl Acad Sci U S A *87*, 4576-4579.

Worden, E.J., Padovani, C., and Martin, A. (2014). Structure of the Rpn11-Rpn8 dimer reveals mechanisms of substrate deubiquitination during proteasomal degradation. Nat Struct Mol Biol *21*, 220-227.

Xiao, Q., Zhang, F., Nacev, B.A., Liu, J.O., and Pei, D. (2010). Protein N-terminal processing: substrate specificity of Escherichia coli and human methionine aminopeptidases. Biochemistry *49*, 5588-5599.

Yamada, S., Hojo, K., Yoshimura, H., and Ishikawa, K. (1995). Reaction of 20S proteasome: shift of SDS-dependent activation profile by divalent cations. J Biochem *117*, 1162-1169.

Yokom, A.L., Gates, S.N., Jackrel, M.E., Mack, K.L., Su, M., Shorter, J., and Southworth, D.R. (2016). Spiral architecture of the Hsp104 disaggregase reveals the basis for polypeptide translocation. Nat Struct Mol Biol *23*, 830-837.

Yoo, S.J., Seol, J.H., Shin, D.H., Rohrwild, M., Kang, M.S., Tanaka, K., Goldberg, A.L., and Chung, C.H. (1996). Purification and characterization of the heat shock proteins HslV and HslU that form a new ATP-dependent protease in Escherichia coli. J Biol Chem *271*, 14035-14040.

Yoo, S.J., Shim, Y.K., Seong, I.S., Seol, J.H., Kang, M.S., and Chung, C.H. (1997). Mutagenesis of two N-terminal Thr and five Ser residues in HslV, the proteolytic component of the ATP-dependent HslVU protease. FEBS Lett *412*, 57-60.

Yu, Y., Smith, D.M., Kim, H.M., Rodriguez, V., Goldberg, A.L., and Cheng, Y. (2010). Interactions of PAN's C-termini with archaeal 20S proteasome and implications for the eukaryotic proteasome-ATPase interactions. EMBO J *29*, 692-702.

Zaremba-Niedzwiedzka, K., Caceres, E.F., Saw, J.H., Backstrom, D., Juzokaite, L., Vancaester, E., Seitz, K.W., Anantharaman, K., Starnawski, P., Kjeldsen, K.U.*, et al.* (2017). Asgard archaea illuminate the origin of eukaryotic cellular complexity. Nature *541*, 353-358.

Zhang, W., Culley, D.E., Hogan, M., Vitiritti, L., and Brockman, F.J. (2006). Oxidative stress and heat-shock responses in Desulfovibrio vulgaris by genome-wide transcriptomic analysis. Antonie Van Leeuwenhoek *90*, 41-55.

Zusammenfassung

Umwelteinflüsse verlangen eine ständige Anpassung des zellulären Proteinbestands. Eine Schlüsselkomponente für diesen Prozess stellt das Proteasom dar, ein selbstkompartmentalisierender Komplex, der über das Ubiquitin (Ub)-System für den Großteil des Proteinabbaus verantwortlich ist. Lange wurde angenommen, dieser Stoffwechselweg sei in Eukaryoten entstanden, während in Prokaryoten eine Vielzahl anderer, weniger charakterisierter Systeme gefunden wurden. Mit der zunehmenden Anzahl sequenzierter Genome und ausgefeilteren bioinformatischen Methoden entdeckte man jedoch immer weitere evolutionäre Zusammenhänge und gemeinsame Konzepte in prokaryotischen und eukaryotischen Systemen. In der vorliegenden Arbeit werden diese Verbindungen verdeutlicht durch die Charakterisierung der neuen Proteasomhomologe Anbu und BPH und des ersten archaealen (De-) Ubiquitinylierungssystems.

Anbu-Sequenzen (Ancestral beta subunit, engl. für urtümliche beta Untereinheit) sind unter phylogenetisch entfernten Organismen weit verbreitet und daher bedeutete ihre Entdeckung, dass das weit verbreitete Konzept, nach dem das Proteasom von seinem einzig anderen bekannten Homolog HslV abstammt, überdacht werden muss. In Kapitel 3.1 analysieren wir diese Verwandtschaftsverhältnisse durch Clusteranalyse und durch die erste biophysikalische Charakterisierung von Anbu. Anbu bildet einen zwölfgliedrigen Komplex einzigartiger Architektur, welche nur durch die Kombination von Röntgenkristallographie und Kleinwinkelstreuexperimenten bestimmt werden konnte. Während gleichförmige Helices in Kristallstrukturen und Elektronenmikroskopiepräparationen beobachtet wurden, enthüllte eine Korrektur der Kristallstruktur durch die Kleinwinkelstreudaten eine offene Ringstruktur, welche im Kontrast zu den geschlossenen Ringstrukturen von Proteasom und HslV steht. Auf Grundlage dieser urtümlichen Architektur und ausführlichen Sequenzanalysen, schlagen wir Anbu als Vorläufer moderner selbstkompartmentalisierender Proteasen vor.

In diesem Zusammenhang stellt BPH (Beta-proteobacterial proteasome homolog, engl. für beta-proteobakterielles Proteasomhomolog) ein vergleichsweise junges

Proteasomhomolog dar, welches durch Genduplikation aus HslV entstand. In Kapitel 3.2 präsentieren wir die erste Charakterisierung von BPH und enthüllen Kristallstrukturen zweier solcher Komplexe aus *Thiobacillus denitrificans* und *Cupriavidus metallidurans*. Interessanterweise bilden beide Proteine siebengliedrige Doppelringe, im Gegensatz zu den sechsgliedrigen Ringen, zu denen sich HslV Untereinheiten zusammenlagern. Die Relevanz dieser unerwarteten Strukturen wird durch Kleinwinkelstreumessungen, analytische Ultrazentrifugation und durch Mittelung von Elektronenmikroskopieaufnahmen untermauert. Die proteolytische Aktivität des BPH-Komplexes wird durch Massenspektrometrie und eine Kristallstruktur, in der der Proteasominhibitor Epoxomicin das aktive Zentrum modifiziert, gezeigt. Im Gegensatz zu HslV scheint BPH unabhängig von Entfaltungsenzymen des AAA+ Typs zu fungieren und könnte daher eine unterschiedliche Rolle im Vergleich zu bisher bekannten selbstkompartmentalisierenden Proteasen annehmen, die möglicherweise analog zu den ATP-unabhängigen Funktionen des Proteasoms ist.

In Kapitel 3.3 zeigen wir durch die Charakterisierung des ersten sequenzierten archaealen Ubiquitinylierungsoperons aus *Caldiarchaeum subterraneum*, dass Proteinubiquitinylierung nicht nur in Eukaryoten vorkommt. Durch Clusteranalyse enthüllen wir die evolutionäre Geschichte dieses Operons und zeigen, dass seine schrittweise Komplettierung von einer Koevolution mit zunehmend Rpn11-ähnlichen Deubiquitinasen begleitet wurde, ähnlich derer im eukaryotischen Proteasom. Der *C. subterraneum* Vertreter CsRpn11 markiert den Höhepunkt dieses Prozesses. CsRpn11 zeigt Aktivität gegenüber Ubiquitin-Vorläufern (CsUb; K_D = 14.6 µM, K_M = 24.2 µM) und gegenüber isopeptid-verknüpften ubiquitinylierten Proteinen. Wir zeigen eine CsUb-CsRpn11 Kostruktur, welche den ersten katalytisch aktiven Deubiquitinase-Ubiquitin Komplex mit zur Proteolyse orientiertem Wassermolekül im aktiven Zentrum darstellt. Ein Vergleich mit der CsRpn11 Struktur zeigt einen ubiquitininduzierten Übergang zwischen zwei Konformationen, die ursprünglich mit unterschiedlichen Ubiquitinbindungsstrategien in eukaryotischem Rpn11 und dessen Homolog AMSH assoziiert wurden. Insofern stellt die CsRpn11-CsUb-Interaktion ein nützliches Modell für homologe eukaryotische Systeme dar und weißt auf einen archaealen Ursprung der Ubiquitinmodifikation hin.

Danksagung

Zuallererst möchte ich mich bei Jörg Martin bedanken, der mir ein super Betreuer seit meiner Diplomarbeit war. Danke für all die Zeit, die Du in Diskussionen und mit mir verwendet hast und für die Unterstützung, die Du mir gegeben hast! Ich glaube, wenige Doktoranden bekommen gleichzeitig eine solche Betreuung und so viel Mitspracherecht und ich bin froh, es so gut getroffen zu haben.

Besonders bedanke ich mich auch bei Marcus Hartmann, der viel Zeit in die Lösung sämtlicher Kristallstrukturen und in die Interpretation von SAXS-Daten gesteckt hat und wie auch Jörg eine entscheidende Rolle in der Bearbeitung des Anbu-Manuskripts gespielt hat.

Des weiteren will ich dem Boss Andrei Lupas danken, weil er eine Abteilung mit angenehmer Atmosphäre geschaffen hat und weil er mir ermöglicht hat dort zu arbeiten.

Ich möchte mich auch bei meinen anderen TAC-Mitgliedern, Fulvia Bono und Thilo Stehle, für ihre Zeit und Unterstützung bedanken. Im Voraus bedanke ich bei den Mitgliedern des Prüfungsausschusses, Volkmar Braun, Andrei Lupas, Fulvia Bono und Thilo Stehle.

Ich danke Vikram Alva, für die Hilfe bei der Interpretation bioinformatischer Analysen und für seine Hilfe bei deren Implementierung in das Anbu-Manuskript; Lorena Maldoner, für großartige Unterstützung bei meinen Experimenten; Reinhard Albrecht und Kerstin Bär für die Hilfe bei der Kristallisation von Proteinen und Reinhard nochmals im Besonderen für das Gefrieren unzähliger Kristalle; Matthias Flötenmeyer und Katharina Hipp für die EM-Bilder; Stefan Grüner für Hilfe mit LS-Experimenten; Marc Claassen, Mirita Franz-Wachtel und Boris Macek (Uni Tübingen), Elisabeth Weyher und Victoria Sanchez (MPI Martinsried) für MS-Messungen; Andriko von Kügelgen, Matthias Wojtynek und Shirley Tremel für experimentelle Zusammenarbeit.

Schlussendlich möchte ich mich auch bei allen anderen bedanken, die das Arbeiten in unserer Abteilung schöner gemacht haben: Ioanna Karamichali, Harshul Arora, Andriko von Kügelgen, Friday, Moritz Ammelburg, Dara Forouzan, Silvia

Wuertenberger, Edgardo Sepulveda, Maria Logotheti, Mateusz Korycinski, Laura Weidmann, Matthias Wojtynek, Shirley Tremel, Birte Hernandez-Alvarez, Tjeerd Dijkstra, Jens Baßler, Claire Bedez, Hongbo Zhu, Eva Sulz, Silvia Deiss, Karin Lehmann, Mohammad ElGamacy, Anja Heubach, Astrid Ursinus, Murray Coles, Lukas Zimmermann, Jyoti Adlakha, Eugen Netz, Manish Chaubey, Emilija Basina. Und natürlich auch allen, die ich hier vergessen habe.