Doctoral Thesis

Characterization of a new AAA+ protein from A. fulgidus and functional studies on chaperone-proteasome complexes in Archaea and Eubacteria

Author(s): Summer, Heike

Publication Date: 2006

Permanent Link: https://doi.org/10.3929/ethz-a-005207512

Rights / License: In Copyright - Non-Commercial Use Permitted
Characterization of a new AAA+ protein from *A. fulgidus*
and Functional Studies on
Chaperone-Proteasome Complexes in Archaea and Eubacteria

HEIKE SUMMER

Institute of Molecular Biology and Biophysics
ETH Zurich

2006

DISS. ETH NO. 16678
Characterization of a new AAA+ protein from *A. fulgidus*
and Functional Studies on
Chaperone-Proteasome Complexes in Archaea and Eubacteria

A dissertation submitted to the

SWISS FEDERAL INSTITUTE OF TECHNOLOGY ZURICH

for the degree of

Doctor of Natural Sciences

presented by
HEIKE SUMMER
Mag. rer. nat., University of Innsbruck
Born 04.10.1978
Austrian

Accepted on the recommendation of:

Prof. Eilika Weber-Ban, examiner
Prof. Rudolf Glockshuber, co-examiner

2006
Table of contents

Summary 3
Zusammenfassung 4
Aim of this thesis 5

Introduction
1. AAA+ protein family 6
2. Molecular Chaperones 9
3. Structure and Function of proteasomes 13
   3.1 Eukaryotic system 13
   3.2 Archaeal PAN-proteasome 17
   3.3 Eubacterial ARC-proteasome 20
4. Overview—the phylogeny of the living world 22
   Archaea 22
   Eubacteria 25

Materials and Methods 26

Results and Discussion
I Cloning, expression and reconstitution of functional PAN-proteasome systems of A. fulgidus, P. furiosus and M. acetivorans 39
1. Genomic gene location, sequence analysis and comparison of PAN and proteasomal α- and β-subunits from A. fulgidus, P. furiosus and M. acetivorans 39
2. Cloning and expression of the proteins of interest 41
3. Core particle assembly and reconstitution 42
4. The core particles from A. fulgidus, P. furiosus and M. acetivorans display peptidase activity 46
5. Assembly states of the regulatory particle PAN from A. fulgidus, P. furiosus and M. acetivorans 47
6. The regulatory PAN complex displays ATPase activity 49
7. Formation of PAN-CP complexes 51
8. AfuPAN interacts with the AfuCP to degrade GFP-SsrA 53
II Cloning of the *M. tuberculosis* proteasome and its regulatory AAA partner ARC and characterization of ARC  
1. Sequence analysis and similarities with other AAA+ proteins  
2. Construction of the over-expressing plasmids and purification of MtuARC  
3. ATPase activity measurements of ARC  
4. Analysis of the assembly state of ARC  
5. MtuARC does not recognize GFP-SsrA

III Research Article: Characterization of a new AAA+ protein from Archaea

IV Ongoing work with AfuATA  
1. Crystallization trials with AfuATA  
2. Test for possible interaction with the proteasomal core particle  
3. Identification of binding partners of the AfuATA complex

Prospects

Curriculum Vitae

References
Summary

This thesis is divided into two parts. The first part describes the cloning, purification and characterization of three prokaryotic chaperone-proteasome complexes, whereas the second part represents the result of the characterization of a novel AAA+ protein from *Archaeoglobus fulgidus*.

Prokaryotic Chaperone-Proteasome Complexes

We have cloned four members of ATP-dependent protease complexes from three archaea and one eubacterial source. Chaperone proteasome complexes contain a central protease part and an associated chaperone part for substrate recognition, unfolding and translocation to the protease part. The archaeal chaperone part PAN (Proteasome Activating Nucleotidase) is a member of the AAA+ superfamily (ATPases associated with various cellular activities), forms homohexameric rings, has chaperone activity and is responsible for recognition and unfolding of substrate proteins that are then translocated into the protease part. The prokaryotic proteasomes consist of two different subunits, alpha and beta, with each subunit forming homoheptameric rings. The rings stack on top of each other in an αββα order with two inner beta-rings and two outer alpha-rings. The proteins of the three archaea members were over-expressed in *E. coli* and purified to homogeneity. The proteasome core complexes could be reconstituted and basic measurements were performed to proof their function ability in vitro. Interaction between PAN and the proteasome could be shown by SsrA-tagged green fluorescent protein (GFP) degradation and peptidase activity of the proteasome was measured separately with fluorogenic peptides. As the interaction between the chaperone and proteasome is transient, no convincing results were obtained by electron microscopy or analytical size exclusion experiments.

A novel AAA+ protein from *Archaeoglobus fulgidus*

A new member of the AAA+ protein family was examined from *Archaeoglobus fulgidus*. This protein is also found in all methanogenic archaea. We named the protein ATA for archaeal triple-A ATPase. It has a monomeric molecular weight of 40'000 Dalton. We over-expressed this protein in *E. coli* and purified it to homogeneity. Like other AAA+ proteins ATA forms hexameric rings with 125 Å diameters, even in absence of ATP. The hexameric assembly state was confirmed by using sedimentation equilibrium analysis over a wide concentration range. Sequence alignments revealed homology to the AAA-domain of FtsH from bacteria, while the N-terminal domain shows homology to members of the CDC48 family of AAA+ proteins. The complex is capable to bind and hydrolyze ATP. Michaelis-Menten analysis revealed a $k_{cat}$ of 118 min$^{-1}$ and a $K_M$ of 1.4 mM at 78 °C. As *Archaeoglobus*
fulgidus is a hyperthermophilic sulphate-reducing archaea this AAA-ATPase is stable to 86 °C and the ATPase activity is maximal at this temperature.

Zusammenfassung

Diese Doktorarbeit ist in zwei Teile aufgeliedert. Der erste Teil behandelt das Klonieren, Aufreinigen und die Beschreibung von prokaryontischen Chaperon-Proteasom Komplexen, während der zweite Teil die Charakterisierung eines neuen AAA+ Proteins aus Archaeoglobus fulgidus präsentiert.

Prokaryontische Chaperon-Proteasom Komplexe


Ein neues AAA+ Protein aus Archaeoglobus fulgidus

Ein neues Mitglied der AAA+ Proteinfamilie aus Archaeoglobus fulgidus wurde untersucht. Dieses Protein kommt ausserdem in allen methanogenen Archaeen vor. Wir benannten das Protein ATA fuer archaeale triple-A ATPase. Es hat ein monomerer Molekulargewicht von 40 kD. Die Ueberexpression in E. coli und

**Aim of this thesis**

When this work was initiated, nothing was known about the herein characterized AAA+ protein from *Archaeoglobus fulgidus* and methanogenic archaea. In order to answer some questions about the function on chaperone proteasome complexes, we cloned, purified and performed first functional studies. In this thesis I will describe and discuss the experiments that were performed in order to achieve these goals. I first give an introduction about this topic and will then describe how the experiments were done. Next, I describe the functional studies which were done on the archaeal chaperone proteasome complexes plus the eubacterial ARC complex. The final chapter presents the published work on the new described AAA+ protein. The thesis concludes with the description of future prospects of the work.
Introduction

1. AAA+ protein family

As ATP is the main cellular energy unit, a lot of processes depend on ATP hydrolysis to convert the chemically stored energy into biological activity. There are many different types of ATPases in the cell and one of the largest classes is the AAA+ superfamily (‘triple-A family’, ATPases associated with various cellular activities)(Patel and Latterich, 1998). This family is characterized by the presence of one or two copies of a conserved region of 220-250 amino acids (AAA cassette) (Confalonieri and Duguet, 1995; Kunau et al., 1993) that include an ATP-binding consensus sequence. Members of this family are found in all organisms.

The classical AAA family was defined as containing a specific motif, namely the second region of homology (SRH) or also called the AAA minimum consensus, in addition to the highly conserved Walker A and B motifs, which are responsible for NTP-binding and hydrolysis (Walker et al., 1982). AAA+ proteins are involved in a great variety of processes (examples given in the overview figure 1) ranging from DNA repair and replication, to organelle biogenesis, cellular housekeeping, cell cycle functions, vesicular transport, transcriptional regulation and protein quality control.

Figure 1: Schematic overview of a hypothetical eukaryotic cell demonstrating the diverse functions of AAA+ proteins including prokaryotic AAA members and AAA+ members (green). Adapted from Ogura T., 2001.
AAA+ proteins are found in multimeric states and are all Mg$^{2+}$-dependent ATPases. There are at least six major clades of AAA domains (metalloproteases, meiotic proteins, D1 and D2 domains of ATPases with two AAA domains, proteasome subunits, and BSC1), as well as several minor clades, some of which consist of hypothetical proteins. The domain organization of AAA ATPases consists of a non-ATPase N-terminal domain that acts in substrate recognition, followed by one or two AAA domains (D1 and D2), of which one may be degenerate. The P-loop NTPase fold is characterized by parallel β-sheets in which α-helices connect the β-strands in a right handed manner. NTPs bind close to the C-terminal ends of the β-strands. The Walker A motif sequence (G-X-X-G-X-G-K-T/S) is contained in a large loop between a β-strand of the sheet and an α-helix. The loop forms a giant anion hole accommodating the phosphates of the NTP (P-loop).

The AAA domain is divided into two subdomains: the P-loop domain has an α/β fold and harbors the nucleotide binding pocket, and the second, less structurally conserved, C-terminal α-helical domain is a bundle of four helical segments (see also figure 2), which varies in size. The P-loop domain forms a core of parallel β-sheets arranged in a β5-β1-β4-β3-β2 topology, where strand 1 harbors the Walker A and strand 3 the Walker B motif (Ogura and Wilkinson, 2001). The β4 insertion between β1 and β3 discriminates the AAA+ proteins from other NTP binding motifs. This β4 sheet is part of the SRH and contains the sensor-1 motif, having a conserved polar residue (Iyer et al., 2004) involved in ATP hydrolysis. The not strictly conserved SRH is formed by parts of β4, alpha helix 4 and the loop to β5, but comparable parts are always present. The SRH contains two other structural elements: the previously mentioned sensor-1 and arginine fingers. The arginine fingers got their name because of the related arrangement of arginines in GTPase-activator proteins. There, they assist nucleotide hydrolysis (Scheffzek et al., 1997). From oligomeric structures it is known that the arginine finger from one subunit forms a part of the nucleotide binding site of the neighboring subunit. Mutation of arginine fingers affects ATP hydrolysis and oligomerization (Ogura et al., 2004).

Figure 2: Monomer structure of *T. thermophilus* ADP bound FtsH (Niwa et al., 2002) with the α-helical domain in yellow:

**Walker A motif:** GxxxxGKT
- K ionic interactions with β- and γ-phosphate oxygens, T provides ligand for Mg$^{2+}$ coordination

**Walker B motif:** hhhhDExxx
- D provides ligand for Mg$^{2+}$ coordination, h= hydrophobic residues, E is catalytic base to activate H$_2$O

**Second Region of Homology**
The last structural element of the AAA+ domain at the C-terminal end is the sensor-2, located in the alpha helical subdomain (α7 helix) with a mildly conserved arginine, which is often exchanged by an alanine. The sensor-2 interacts directly with the γ-phosphate of ATP. This helix is next to an adjacent subunit and is therefore involved in conformational changes between subunits similar to the SRH (Ogura and Wilkinson, 2001). The importance of sensor-2 for function between neighboring subunits has been reported by several examples (Liu et al., 2000; Song et al., 2000).

The mechanism of ATP hydrolysis starts with the nucleophilic attack of an activated H₂O molecule at the γ-phosphorus of the NTP and the formation of a 5-coordinate transition state. The negative charge which is accumulating at the γ-phosphate is stabilized by a magnesium ion and by positively charged groups or hydrogen bond donors. The active sites contain in addition one or more sensor domains, which interact with Walker elements and the γ-phosphate and sense, if the γ-phosphate is bound and transmit this information through conformational changes. The nucleotide binding pocket is formed by the interface of two neighboring subunits. AAA+ enzymes assemble into higher oligomers, generally hexamers, which seem to be the biologically active form. Several structures of AAA+ protein oligomers have been solved, e.g. NSF, PA 700 in complex with the proteasome or p97 (DeLaBarre and Brunger, 2003; Forster et al., 2005; Lenzen et al., 1998). ClpA hexamer is formed by a dimer-tetramer-hexamer pathway within 60-90 seconds and then binds to ClpP (Kress WB, personal communication).

AAA+ members having two AAA modules forming two distinct homotypic ring structures stacking on top of one other (Guo et al., 2002) but the two AAA modules differ from each other in terms of function, structure and sequence.

For the cycle of ATP hydrolysis in the hexameric ATPases several possible models have been suggested (Ogura and Wilkinson, 2001). The first proposed model refers to a synchronized model where all six subunits bind and hydrolyze ATP and the six-fold symmetry is kept throughout the reaction. In the second rotational model, which was suggested for the F₀F₁ ATP Synthetase, ATP binds to three of the six subunits, which are then active in ATP hydrolysis, while the other free subunits stay inactive. Each active subunit is in a different state of the reaction cycle (ATP bound, ADP and inorganic Phosphate are also bound and the nucleotide binding site is unoccupied). A sequential model has been suggested as an advancement of the previously described rotational model, where all subunits are active. Pairs of subunits on the opposite sides of the hexameric ring are in the same state of ATP hydrolysis with one another, but are in different ATP hydrolyzing states than the rest of the subunits. Recently a stochastic model has been proposed for ClpX (Hersch et al., 2005).

For substrate binding, substrates often contain recognition tags or motifs on their N- or C-terminus, but it is also known that recognition sequences can occur internally or at multiple sites (Wickner et al., 1999). Two examples are the polyubiquitin tag in eukaryotes or the SsrA tag in prokaryotes (Deveraux et al., 1994; Keiler et al., 1996). It is not clear which part of the AAA+ protein is responsible for substrate recognition, because two possibilities can be found also here. For the AAA protease Yme1p the
N-domain binds to denatured proteins (Leonhard et al., 1999). In other AAA+ proteins substrate binding occurs via the C-domain of the AAA+ module including the α-helices and the sensor-2 and this is referred to sensor- and substrate-discrimination (SSD) domain (Smith et al., 1999). The substrate binding itself is not ATP-dependent but for some AAA+ proteins ATP is required to form the active oligomeric state, e.g. ClpA. The working mechanism of AAA+ proteins is best described by a threading mechanism. This means that a polypeptide or polynucleotide chain is directed through the central pore of an AAA+ oligomeric ring. Because of the narrow pore opening of the protease channels, only unfolded substrates are able to enter the cavity. This unfoldase activity was nicely shown for the unfolding of SsrA-tagged green fluorescent protein (GFP) by ClpA and ClpX (Kim et al., 2000; Weber-Ban et al., 1999). The substrate unfolding mechanism is a processive and directional one and is ATP dependent (Reid et al., 2001).

Another function of AAA+ proteins is the ability to break up or prevent protein aggregation and to act as molecular chaperones (Leonhard et al., 1999; Wickner et al., 1994). There are many described additional functions of AAA+ proteins and still a lot of proteins are uncharacterized and therefore, the distinct function is hypothetical or unknown. Since this huge family of proteins evolved in many different directions and members are important in several essential cell processes, it is not surprising that some of the human proteins are associated with genetic diseases. As the AAA+ family is large and involved in a variety of functions, there are many open questions, such as the dynamics of ATP hydrolysis, the occurring conformational changes and the substrate recognition, binding and translocation in the various systems.

2. Molecular chaperones

Proteins are the basic modules of a cell as they fulfill different and heterogeneous functions. Not only are they involved in structural parts in giving cells a shape, but more importantly they are the molecular machines. As enzymes they execute numerous chemical reactions, or work as motor- or transport-proteins. The mature and active conformation of a polypeptide is encoded within its primary amino acid sequence. This information is sufficient for protein folding in vitro (Anfinsen, 1973). But the folding in vivo is even more sophisticated. For proper functionality, it is fundamental that the proteins are in their correct native conformation. The proteome of a living cell is huge and diverse. In the crowded environment of a cell it is very important that the protein folding occurs efficiently and without aggregation as a major side reaction (the disordered and non-specific association of polypeptide chains).

Cells did evolve strategies to assure productive folding of synthesized polypeptides into active proteins. But, folding of a nascent chain is not the only situation when non-native proteins occur in the cell. Undesirable misfolding or aggregation can be
induced by several stressors like non-physiological high temperatures or in the presence of certain chemicals or other environmental conditions.

The principle function of molecular chaperones can be summarized as follows: they prevent misfolding and aggregation of folding proteins and promote proliferous active folding in an ATP-dependent manner (Buchner, 1996; Gething and Sambrook, 1992) like folding machines or passive folding in an ATP-independent mechanism (holders and folding catalysts) (Fig 3). The helper proteins are divided, according to their principle function, into molecular chaperones (efficiency enhancer, unaffected by the folding rate of the substrate protein) and folding catalysts (responsible for acceleration of rate limiting folding steps).

Molecular chaperones are a group of structurally divers and mechanistically distinct proteins. A cell contains a set of these folding helpers, eg yeast contains 14 different variants of the heat shock protein 70 (Hsp70) (Craig, 1999). Within the same group of chaperones there are homologies in sequence and structure, but less consilences between chaperones from different families, although common functional features are always shared. Molecular chaperones induce conformational changes in their substrate protein during the ability of binding unfolded or partially folded polypeptides.

An example for ATP-independent folding is, when the nascent polypeptide chain exits the ribosome. Hydrophobic residues are partially solvent accessible and therefore do tend to aggregate if this residues are unprotected from the surrounding solvent. The *E. coli* Trigger Factor (Fig 3) provides a shielded folding environment for the newly synthesized polypeptide chain at the ribosomal exit site. In the hydrophobic cradle co-translational domain-wise folding of nascent peptides is promoted (Ferbitz et al., 2004).

ATP-dependent folding is performed by GroEL-GroES (Fig 3), which is found to be essential for eubacterial growth and in mitochondria and is a group I chaperonin, while group II chaperonins are found in the cytosol of eukarya and archaea, lacking a co-chaperone. GroEL is a large cylindrical complex and the two heptameric rings are stacked back to back. During the reaction cycle each ring is in the opposite state than the other (*cis* and *trans*). The substrate polypeptide binds through hydrophobic interactions to a free ring of GroEL. ATP and GroES (co-chaperone) bind to the same ring and form a folding-active *cis* stage. Under the half-dome structure of GroES, an enlarged cavity is formed and at the same time conformational changes rotate the hydrophobic particles away from the inner cavity and the polypeptide is released into the cavity.

Within 15-30 seconds the *cis*-ring-bound ATP is hydrolyzed and a second conformational change leads GroEL to release GroES by binding ATP to the *trans* ring (Bukau and Horwich, 1998). Then, the polypeptide is released from the cage, regardless of its folding state. This nucleotide binding to GroEL induces the transition from the high to the low affinity state and therefore controls the substrate binding and release.
Figure 3: Nomenclature of the major heat shock proteins (Hsp), which are categorized on their molecular weight and divided into ATP-dependent groups (adapted from Lee and Tsai, 2005). Structures: yeast Hsp90 (PDB: 2CG9), *E. coli* DnaK (PDB: 1DKG), *E. coli* GroEL (PDB: 1PCQ), *M. jannaschii* Hsp16.5 (PDB: 1SHS) and *E. coli* Trigger Factor (PDB: 1W26). Prokaryotic members are often known by their specific designations, while eukaryotic members refer to by their Hsp nomenclature. ATP-dependent members often require co-chaperones or nucleotide exchange factors for function (some are listed).

Molecular chaperones play a fundamental role in protein quality control. Not only do they support efficient folding and maintenance of the folded state under stress conditions, but in addition they are involved in disposing damaged or degrading tagged proteins. These control machineries are energy-dependent proteases. Examples are the 26S Proteasome and the eubacterial homologs ClpAP or ClpXP (caseinolytic protease). The ClpP cylinder is formed by two homoheptameric rings of ClpP subunits and the protease components itself are unable to hydrolyze substrates directly. Only small peptides of < 5 amino acids (Thompson et al., 1994; Woo et al., 1989) are able to access the catalytic sites. Therefore the proteases associate with regulatory particles (RP). ClpP forms complexes with AAA+ ATPases such as ClpA or ClpX (see figure 4).
The homohexameric regulatory part recognizes, binds and unfolds substrate proteins under ATP consumption. One of the chaperone activities of ClpA for example is to monomerize the RepA dimer, a DNA binding protein which is only active in P1 origin binding as a monomer. This group of chaperones is also named Hsp100 based on their molecular weight. ClpA also protects firefly luciferase from heat inactivation by preventing its aggregation but it is unable to reactivate the heat treated luciferase (Wickner et al., 1994).

The archaeal precursor of the 26S proteasome from eukaryotes is of simpler architecture. PAN (proteasome activating nucleotidase) is equivalent to ClpA with the main difference in carrying only one AAA ATPase. PAN reduces the aggregation of heat denatured proteins like rhodanese and enhances the refolding of guanidine hydrochloride denatured glucose dehydrogenase. It was also shown that PAN was able to recognize the SsrA-tag and to unfold SsrA-tagged GFP (Benaroudj and Goldberg, 2000). Summarized PAN helps to reduce aggregation, folding and promotes the unfoldase part in the ATP-dependent protein degradation pathway. One characterized recognition motif is the SsrA-tag, composed of 11 amino acids encoded by the small 10S RNA and is added to a polypeptide when the ribosome is blocked up (Keiler et al., 1996). The unfolded substrate is sequestered into the protease, where it is cleaved into small peptides (Hoskins et al., 1998).
3. Structure and function of proteasomes

3.1 Eukaryotic system

One quintessential function in a cell is protein degradation. The posttranslational protein quality control depends on this activity, since aberrant nonfunctional proteins, misfolded proteins as well as unnatural proteins generated from heat or oxidative stress have to be removed to avoid damage to the cell (Gottesman, 1996). In addition the degradation of regulatory proteins, which are used at very specific times for only short duration and often indirectly control the cell cycle, like transcription factors or components of the signal transduction pathway (Coux et al., 1996; Varshavsky, 1997b) is of key importance to the viability of a cell. It is obvious that degradation has to be controlled to avoid any unwanted protein degradation. One control function is the specificity of recognition. A well known example is the recognition of certain N-terminal amino acids of a protein itself, which determines its life time, referred to as the N-end rule (Varshavsky, 1997a). Another control function is that protease active sites, which perform the breakdown of unwanted proteins, are compartmentalized in protein complexes and not easily accessible for bulk proteins. But the most complex and diverse regulation is adding a recognition signal to a protein. One of these recognition tags in eukaryotes is ubiquitin. Ubiquitin is a 76 amino acid regulatory polypeptide with a relative molecular weight of 8.5 kDa and is ubiquitous to eukaryotes. The sequence is highly conserved. Yeast differs from human ubiquitin in only three amino acids. The native conformation is globular, only the last four C-terminal residues point outside. The function of ubiquitin is to highlight proteins determined for degradation, if a protein becomes old, denatured or was wrongly synthesized. Another example shows that in plants the circadian clock and phytohormon signaling pathways are regulated in this way (Smalle and Vierstra, 2004). But there are many other functions of ubiquitin depending on how many ubiquitin chains are put together. A protein can also be monoubiquitinated or polyubiquitinated through alternative linkages. These modifications control the protein activity or its localization (Hicke et al., 2005). The following is the ubiquitin domain signature motif from prosite (www.expasy.ch)

**Consensus pattern:** K - x(2) - [LIVM] - x - [DESAK] - x(3) - [LIVM] - [PAQ] - x(3) - Q - x - [LIVM] - [LIVMC] - [LIVMFY] - x - G - x(4) - [DE] (Prosite search, Accession number: PS 00299).

The human ubiquitin sequence is:

MQIFVKTLTG KTITLEEPS DTIEVKAKI QDKEGIPPDQ QRLIFAGKQL #50
EDGRILSLDY I QKESTLHLV LRLR GG #76

The ubiquitin C-terminus forms an isopeptide bond with the amino group of a lysine side chain in the target protein. In this way proteins are covalently modified by
adding ubiquitin and this changes the target protein’s state. As ubiquitin itself contains seven lysines, multiple molecules of ubiquitin can become linked to each other to form polyubiquitin chains. This modification targets the protein for degradation by the large 26S proteasome complex (Fig. 5).

Figure 5: The ubiquitin-mediated protein degradation (adapted from Miller & Gordon 2005; Nobelprize.org).

The 26S proteasome is an ATP-dependent protease complex, consisting of the 20S core particle (CP) and the two outer 19S regulatory particles. The ubiquitin-proteasome pathway is complex, involving about 100 known components, with only 30 distinct subunits in the proteasome. The CP consists of four stacked rings, which give a barrel shape structure with three large cavities and narrow constrictions between them and have a molecular weight of about 700 kDa. The outer chambers are formed by alpha and beta rings, while the central chamber, containing the proteolytic threonines, is formed by the two inner β-rings. The eukaryotic proteasome is built of 14 different subunits which show a similar fold, even though their amino acid sequences show low homology. A subunit is folded in a five stranded β-sandwich, flanked by α-helices on the top and bottom (see also Fig. 9 and 10). In the eukaryotic proteasome only three (β1, β2 and β5) of the seven β-subunits are proteolytically active. The beta subunits contain, in their inactive form, a
prosequence of various lengths. In the path of maturation, the autocatalytic cleavage of the propeptides is the last step (Fig. 6) in forming the functional complex.

Figure 6: Proposed model for the proteasome maturation. Adapted from Groll et al. 2005. The N-termini of the alpha-subunits are colored orange. The beta propeptide is colored black. The α-rings assemble spontaneously and serve as templates for the β-precursors building the half proteasomes (αβ). Upon autocatalytic cleavage of the β-propeptide, when both half proteasomes come together, the catalytic active sites are formed and the mature core particle assembles.

This assembly pathway represents a kind of auto regulation within the cell to prevent uncontrolled protein degradation. Another function of the prosequence is to bring the half proteasomes into the right positions before the complex formation. In addition it is also known, that proteasome maturation factors exist (Ramos et al., 1998; Tone et al., 2000). The 20S proteasomes were found to be a novel protease family: the threonine proteases. The eukaryotic CP can cleave after almost every amino acid, but five cleavage preferences were determined:

* chymotryptic-like (β5 subunit),
* tryptic-like (β2 subunit),
* peptidyl-glutamyl-peptide hydrolyzing (PGPH) (β1 subunit),
* branched chain amino acid preferring (BrAAP) (β1, β5 subunits) and
* small neutral amino acid preferring (SNAAP) (β5 subunit) activity (Arendt and Hochstrasser, 1997; Groll et al., 2005; Groll et al., 1997; Orlowski and Wilk, 2000).

The proteolytic active centers produce cleavage products in the length of 3-25 amino acids, with an average length distribution of 7-8 amino acids. The control mechanism for peptide product length is due to specific channels, in which the substrate is bound. These pockets exhibit binding sites for peptides of 7-9 amino acids. The N-terminal threonine is exposed to reveal a γ-oxygen, which acts as the nucleophile and an α-amino group that is the likely proton acceptor in the peptide bond hydrolysis. Additionally, a salt bridge is formed by Lys33 and Asp/Glu17. The Thr1 N atom represents the proton acceptor. The threonine N-terminus is hydrogen bridged to Ser129Oγ, Asp168O and Ser169γ. A nucleophilic water molecule (NUK) is essential for proteolysis.
Figure 7: Schematic overview of the 26S proteasome (adapted from Viestra 2003). a,b) Organization and composition of events that lead to the degradation of an ubiquitinated protein by the 26S holoprotease. The N-terminal threonine residues that form the protease active sites in β, β2 and β5 subunits are indicated in dark red. c) organization of the regulatory particle (RP) with its lid and base subparticles. Abbreviations: N= non ATPase subunits, T= AAA-ATPases subunits.

The 20S CP interacts with different caps, including the 19S regulatory particle (RP, Fig. 7), PA28 (also named proteasome activator, REG, 11S complex or PA26, see figure 8) or PA 200 (200 kDa protein, involved in DNA repair (Ustrell et al., 2002)). These different caps modify the proteasome activity. The 26S proteasome consists of the CP and two 19S caps and is responsible for protein degradation. The 19S cap includes six regulatory particle triple-A (Rpt1-6) type I proteins and 11-12 regulatory particle non-ATPases (Rpn1-12) (Finley et al., 1998) as indicated in figure 7c. The regulatory particle functions as a gate opener and is responsible for substrate recognition, unfolding and translocation (Groll and Huber, 2003). The heptameric 11S regulatory complex in association with the proteasome enhances the production of antigenic peptides for presentation by MHC class 1 molecules (Dick et al., 1996; Groettrup et al., 1996).
Figure 8: Crystal structure of yeast 20S CP in complex with the heptameric 11S RP of *T. brucei* (Whitby et al., PDB: 1FNT).

From structural analysis of proteasome core particle structures it is known that the N-termini of the α-rings keep the entrance channel closed in absence of a regulatory complex (Groll et al., 1997). This was also nicely shown for another proteasome activator namely the 11S RP in complex with the yeast proteasome, where the whole complex structure could be solved (Fig. 8, Whitby et al., 2000). Recent investigations on *T. acidophilum* proteasome showed that the antechambers act in substrate storage before the degradation, when the translocation speed from the RP into the alpha ring is slower than proteolysis (Sharon et al., 2006).

### 3.2 Archaeal PAN-proteasome

The archaeal equivalent to the eukaryotic 26S proteasome is described as PAN-Proteasome (PAN = Proteasome Activating Nucleotidase, Rpt-like). These genes are not located within an operon. PAN is a member of the AAA+ protein family with only one AAA motif. Most of the sequenced archaea have only one type of α and β subunit but some, as for example *P. horikoshii*, have two different β subunits. The interaction between PAN and the proteasome is weak and not very stable and could only be demonstrated using *T. acidophilum* CP and *M. jannaschii* PAN, in presence of ATP or ATPγS (Smith et al., 2005). But also in this case it is reported that the interaction is short-lived, but long enough for substrate translocation. Seen on EM images, assembled PAN has a two ring appearance. A larger inner ring, bound to the proteasome α-ring and a smaller outer ring (Smith et al., 2005). The small outer ring corresponds probably to the PAN N-terminal domain, similar to the two ring appearance in HsIU (Song et al., 2000).
The ubiquitin tagging system is absent in prokaryotes and the SsrA-tagging system exists also in bacteria. Nevertheless it was shown that the SsrA-tag on globular substrates is recognized by PAN in vitro (Benaroudj and Goldberg, 2000), probably due to its unstructured nature. PAN shares 41-45% sequence homology to the six 19S ATPases (Zwickl et al., 2000) and is most homologous to the Rpt2 subunit, which is responsible for gate opening in the 20S CP (Kohler et al., 2001). Recently two groups of archaea (haloarchaea and methanosarcinales) were found encoding for two different PAN homologs (PAN A & B) that are differentially regulated (Reuter et al., 2004). Before that, all other known archaea were thought to encode for only one type of regulatory PAN protein. The functional role of the 20S proteasome in vivo has not been identified yet and no substrate proteins have been found.

The archaeal propeptides are in the range of ten amino acids and the maturation of the proteasome follows the previously described model (Fig. 6). The crystal structure of the A. fulgidus β Thr1Gly proteolytically inactive mutant represents one of the last CP assembly intermediates in archaea (figure 11 and 10 on the right: Ribbon drawing of three β-subunits, illustrating that there is no structural order at the β–β interface). The N-terminal residues of the α-subunits are unstructured and not defined, even though the conserved YDR motif exists, which forms the motif of the gate in eukaryotes (Groll and Huber, 2003). A closed gate structure, like in eukaryotic proteasomes, could not be observed by structural analysis. This might be due to the lack of sequence asymmetry, which seems to be important for a closed pore in eukaryotic α-rings. But from biochemical analysis, a gating mechanism is proposed, as only small four-residue peptides are able to diffuse into the archaeal proteasome to be cleaved. Another explanation could be that, through crystallization conditions, the closed gate structure is disrupted (see also figure 9).

Figure 9: On the left side one A. fulgidus α-subunit and as a complete heptameric ring on the right (Groll et al., 2003) the pore size is 1.3 nm.
Similar to the eukaryotic α/β subunits, both subunits show a similar overall fold: a central beta sandwich surrounded by α-helices (figure 9 and 10). The archaeal α-subunits are able to assemble spontaneously into single heptameric rings or double stacked rings (Zwickl et al., 1994).

Like the eukaryotic relative, the proteasome active site in the mature complex is Thr1.

Figure 10: *A. fulgidus* single beta subunit on the left side and the β-ring is shown on the right side, PDB: 1j2p.

Figure 11: Ribbon drawing of the 20S CP from *Archaeoglobus fulgidus* (PDB: 1J2Q, Groll et al.).
The cleavage is processive, the product size ranges from 6 to 10 residues and the velocity of hydrolysis increases with the substrate length (Kisselev et al., 1998). The main ATP requiring step in protein degradation is the unfolding of globular proteins. Translocation is also possible by carrier-mediated diffusion from PAN to the CP (Smith et al., 2005). A possible biological role of the 20S CP and PAN could be involvement in mediating the cellular stress response.

3.3 The eubacterial ARC-proteasome

Actinomycetes are the only known bacteria that contain in the addition to the Clp-proteases proteasome genes (Bochtler et al., 1997; Wang et al., 1997). The predicted but not yet demonstrated interacting partner of the proteasome was first described from *Rhodococcus erythropolis* and referred to as ARC (AAA ATPase forming ring-shaped complexes) (Wolf et al., 1998) and the same hexamer complex was identified from *Mycobacterium tuberculosis* but no interaction partner or distinct function could be shown *in vitro*, as the association constant might be too low to detect interactions outside the cell (Darwin et al., 2005). The ARC protein could be purified as a complex of single and double stacked homohexameric rings, able to hydrolyze ATP, ADP and CTP (Wolf et al., 1998). The *M. tuberculosis* ARC (Mpa, mycobacterium proteasome ATPase) forms hexameric rings and shows ATPase activity *in vitro* and is 82% identical to the *R. erythropolis* ARC (Darwin et al., 2005). The genes for encoding the proteasome subunits and the proposed interaction partner ARC are typically formed in linked operons. *R. erythropolis* CP is the only known bacterial proteasome built of two different α- and β-subunits (Tamura et al., 1995; Zuhl et al., 1997), the other actinomycetes have a CP architecture similar to the archaeal ancestor with only one type of each α- and β-subunits. The maturation pathway differs from the previously described model (Fig. 6). Other than for archaea, α/β dimers are the early intermediates that then form the half proteasomes. The β propeptide is even longer than in archaeal subunits, up to 65 amino acids longer. In *Rhodococcus* this propeptide acts as a molecular chaperone helping β subunit folding and in the maturation of the complete proteasome (Zuhl et al., 1997).

In comparison to archaeal or eukaryotic CP, the proteasomes of actinomycetes have smaller contact regions, which could explain the fact that the α-subunits do not assemble into rings by themselves. From heat shock experiments no changes in expression levels of 20S proteasome subunits could be detected, even in knockout mutations, no phenotypic changes could be detected. Reasonable explanations could be that other proteases compensate the lack of proteasome function (Knipfer and Shrader, 1997). This year the structure of the *M. tuberculosis* CP was solved and crystals soaked with proteasome inhibitor indicated that all 14 Thr1 residues in the β subunits are proteolytically active (Fig. 12) showing tryptic, chymotryptic and PGPH (petidyl-glutamyl-peptide-hydrolysing) activities.
Figure 12: *M. tuberculosis* proteasome structure, with α-subunits in blue and β-subunits in red (PDB: 2fhh).
4. Overview – the phylogeny of the living world

The relevant archaea and bacteria used as sources for DNA or proteins in this thesis are briefly described in the following sections. To get a general idea about their classification and phylogeny an overview of the three kingdoms of life is mapped in figure 13.

Figure 13: The phylogenetic tree of life – adapted from www.uccs.edu/~rmelamed/MicroFall2000/Chapter 9/Origins of Life.html.

4.1 Archaea

The Kingdom Archaea consists of three phylogenetically-distinct groups: Crenarchaeota, Euryarchaeota and Korarchaeota (Fig. 14).
These interesting organisms show extreme adaptations in their physiology, to be able to live in almost every environment under highly extreme conditions (hence they are often named extremophiles). Based on this ability, they are divided into methanogens (methane producing prokaryotes), extreme (hyper) thermophiles (living at very high temperatures, above 50 °C) and extreme halophiles (living at very high salt concentrations).

**Archaeoglobus fulgidus**


*Archaeoglobus fulgidus* is a thermophilic, sulphate-reducing archaea and is much related to methanogens, not only based on ssrRNA analysis but this organism also encodes enzymes for methanogenesis (Brock). Growth occurs between 60 to 95 °C and the optimum is at 83 °C. It grows organoheterotrophically (using different carbon and energy sources) as well as lithoautotrophically on hydrogen, thiosulphate and carbon dioxide. The cells are irregular shaped spheres containing a glycoprotein envelope and flagella at one end.

A gene has been found in the sequence of *A. fulgidus* with similarity to the AAA-domain (ATPases Associated with diverse cellular Activities) of the FtsH Protein from *E. coli* (membrane bound ATP dependent protease). This protein is described in
the result part of this thesis. The DNA sequence used for our analysis was determined by Klenk et al (Klenk et al., 1997) encoding for 4881 predicted proteins.

**Pyrococcus furiosus DSM 3638**


*P. furiosus* is a strictly anaerobic, heterotrophic, sulfate reducing archaea, found in the marine sand surrounding of sulfurous volcanoes. Growth occurs between 70-103 °C at pH from 5-9, with an optimal growth temperature of 98 °C. The cells form motile cocci, have about 50 flagella at one end and are often found as pairs. Even at 103 °C their chromosomal integrity is maintained and the DNA shows little accumulation of breaks and additionally it is also resistant to radiation.

*P. furiosus* was sequenced in 2001 by Robb et al (Robb et al., 2001), has a GC-content of 40% and 4300 predicted proteins.

**Methanosarcina acetivorans C2A**


*Methanosarcina* species live in oil wells, sewage lagoons, trash dumps, decaying leaves, stream sediments, and the stomach of cows and other places.

![Figure 15: Light microscopy image of *M. acetivorans.*](image)

*M. acetivorans* is one of the most versatile methane-producing microbes (or methanogens). It subsists on a diverse menu of energy sources, including acetate. It
is the only known organism, able to possess all three pathways for methanogenesis using nine different substrates.

The methane production plays an important role in the global carbon cycle and a large amount is made from biological sources. As seen in figure 15, these organisms are able to form complex multicellular structures, during different growth phases and also dependent on the environment. Single cell growth, with and without a cell envelopes as well as multicellular packets are reported (Macario and Conway De Macario, 2001) with a cell size of 1.7 µm (Fig 15.). The pH ranges from 6.9-7.3 for growth at 37 °C. *M. acetivorans* contains the largest archaeal genome, with 9080 predicted proteins and was sequenced in 2002 (Galagan et al., 2002).

### 4.2. Eubacteria

Bacteria are a huge group of prokaryotic microorganisms that shared a distant common ancestor. In terms of pure cultures fourteen phyla of Bacteria are known. Based on ssrRNA many more phyla are known to exist (Madigan, 2000).

*Mycobacterium tuberculosis*


*M. tuberculosis* was first described by Robert Koch in 1882, as an infection leads to formation of tubercles in the lung and further tissue damage. The name *Mycobacterium* originated from the fact that this bacteria show mold-like (Greek mykes = fungus) surface growing in liquid cultures. It is also stomach acid resistant and insensitive to low temperatures as well as to temperatures up to 65 °C and contains waxes in its cell wall. The genome contains 8192 predicted proteins and was sequenced in 2002 (Camus et al., 2002; Fleischmann et al., 2002). Approximately one third of the world’s population is infected with *Mycobacterium tuberculosis* (http://www.who.int/gtb/). Not every infected person shows symptoms of illness, but many develop an active infection that is often lethal when not treated. The drug therapy is prolonged and toxic and another problem is the increase in multidrug resistant strains of *M. tuberculosis* (Frieden et al., 2003).
Materials and Methods

Materials
Genomic DNA from *Methanosarcina acetivorans* was purchased from the „Deutsche Sammlung für Mikroorganismen und Zellkulturen (DMSZ)”, genomic DNA of *Pyrococcus furiosus* and *Archaeoglobus fulgidus* was kindly provided by Karl Stetter, genomic DNA of *Mycobacterium tuberculosis* was a gift from Roland Brosch, Institut Pasteur. Primers to amplify the genes encoding PAN, the alpha- and beta-subunits of the core particle, a new AAA+ protein and the *M. tuberculosis* Rv2111c (named here MtuMyst) genes were obtained from Microsynth. The restriction enzymes *Nde*I, *Bgl*II, *Hind*III, *Bam*HI, *Nhe*I, *Spe*I, *EcoR*V as well as dNTPs, CIAP and T4 Ligase were from MBI Fermentas. *Nco*I and *Bfa*I were from New England BioLabs. The plasmids pETDuet, pET20b(+), pet11a, pET16b, pET21d(+) as well as the *E. coli* cells BL21 (DE3) and Rosetta™ were from Novagen. DMSO, Pfu Turbo polymerase and Herculase were from Stratagene. IPTG, MgCl₂, chloramphenicol, ampicillin, EDTA, NADH, DNase and DTT were from AppliChem. Citrate synthase, MDH, Oxalacetat, Rhodanese, ATP and ATPγS were purchased from Roche Molecular Biochemicals. Glycerol was from Schweizerhall Chemie AG. Trizma base, MES, PMSF, protease inhibitor cocktail, Pepstatin A, HEPES and PNPass were from Sigma. NaCl and K₂HPO₄ were obtained from Merck. Ammonium sulfate, Imidazol and KH₂PO₄ were from Fluka. The HMW Gel Filtration Calibration Kit was from Pharmacia as well as all the chromatography resins. PVDF sequencing membranes and membranes for stirred cells were purchased from Milipore. The QIAquick Gel Extraction Kit and the MinElute PCR Purification Kit were from QIAGEN. The CoIP kit was from Pierce.

Methods

Cloning
All genes from *Methanosarcina acetivorans*, *Archeoglobus fulgidus*, *Pyrococcus furiosus* and *Mycobacterium tuberculosis* were amplified from genomic DNA using PCR and all results were verified by sequencing.

Used Primers:

M.ac. Alpha sense 5’-ACTCCATGGCACCACAGATGGG-3’
M.ac. Anti sense 5’-CCTGACGGATCCCTATTGCCTGCTTTCC-3’
M.ac. Beta sense 5’-ATGGATGATGACAAATATCTAAAGGGC-3’
M.ac. Anti sense 5’-GCACGAAGATCTTAGTTTAATGAATCTCTTCTTCTGG-3’
M.ac. PAN sense 5’-GCACGAAGATCTTAGTTTAATGAATCTCTTCTTCTGG-3’
M.ac. PAN antisense 5’-TAATGTGCTTGTAACAACAGAGACCTTCCAACGG-3’
M.ac. ATA sense 5’-AAAACATATGGCTGTGCGACCGGTCTCACAAAGTG AAC-3’
M.ac. ATA antisense 5’-AAGTAAAGCTTATACATAAAGATGGGTCGGATCCG-3’
A.fu. Alpha sense 5’-TAGAATCATATGCATTTACCGCAAATGGGA-3’
A.fu. Anti sense 5’-TAGAATCATATGCATTTACCGCAAATGGGA-3’
A.fu. Beta sense 5'-AGTTAACATATGAGCATAATAGAGGAGAAG-3'
A.fu. Beta antisense 5'-AGTATTGGATCCTTATTTCCTGAACTTGGCCA-3'
A.fu. PAN sense 5'-TAGTTACATATGGGCGATAGCGAAATAC-3'
A.fu. PAN antisense 5'-ATAAGTGGATCCTTAAACGAACATCACTCCCTT-3'
P.fu. Alpha sense 5'-ATACGTCCATGGCATTTGTTCCACCTCAGGCTGGG-3'
P.fu. Alpha antisense 5'-ATTAAGGGATCCTCAGTAGTTGCTATCCAGTCCGAAT-3'
P.fu. Beta sense 5'-AGTATACATATGGAAAAAAGACTGGAACG-3'
P.fu. Beta antisense 5'-TCAGATGGATCCTCATACAAACTCTTCCTTCATGC-3'
P.fu. PAN sense 5'-ATAGAACATATGAGTGAGGACGAAGCTC-3'
P.fu. PAN antisense 5'-ATACTAGGATCCTCAGCCGTAAATGACTTCATG-3'
M.tu. Alpha sense 5'-AGATGGCCCATGGCTTTTCCTGATTTCATCTCCGC-3'
M.tu. Alpha antisense 5'-ATCCATGAATTCTCAGCCGACGATTCGCC-3'
M.tu. Beta sense 5'-TCTTCACATATGGACCTGGCCTTTGCCCG-3'
M.tu. Beta antisense 5'-AGTTCTGGATCCTCACTTCTCACCGCCATCGG-3'
M.tu. ARC sense 5'-TGTAGGCCATATGGGTGAGTCAGAGCGTTCTGAGGGACGCTCCGCATC-3'
M.tu. ARC antisense 5'-GCAGTAGGATCCTCAGGTAATGGCAGGCTCGGTATC-3'
M.tu. Rv2111c (Myst) sense 5'-TGTAGGCCATATGGGCGAAGACGACCAAGCGTGCGGCGGCAAGTGATC-3'
M.tu. Rv2111c (Myst) antisense 5'-GCAGTAGGATCCTCAGGTAATGGCAGGCTCGGTATC-3'
Sequencing Primer: M.tu. ARC middle 5'-CTGGCTGATCCCCTGATC-3'
pET Upstream Primer 5'-ATGCGTCCGGCGTAGA-3'
DuetDown1 Primer 5'-CTAATACGCCGGACATGTT-3'
DuetUP2Primer 5'-TTGTACACGGCCGCATAATC-3'
T7Promotor 5'-TAATACGACTTCAGATATAGG-3'
T7 terminator 5'-CGATCAATACGACTTCAGATATAGG-3'

**Construction of the alpha-beta plasmids**

MA3873 (beta subunit) is encoded by nucleotides 4757421 to 4758053, MA1779 (alpha subunit) from 2184888 to 2185631 of the *M. acetivorans* genome. Both inserts were cloned into the pETDuet vector and were each under control of the T7 promoter. The Alpha subunit was cloned into multiple cloning site one and Beta in multiple cloning site two. This gene construct will be called AB. Rv2109c (PrcA, α-subunit) from *M. tuberculosis* is encoded by nucleotides 2368983-2369729 (pETDuet, MCS I) and Rv2110c (PrcB, β-subunit) from 2369726-2380601 (pET20b+). *P. furiosus* α-subunit encoded by nucleotides 1467425-1468207 (pET21d+) and the β-subunit from 131728-1318448 (pET20b+). The genes for AfuAlpha (AF0490, 448081-448821) and AfuBeta (AF0481, 439929-440570) from *A. fulgidus* were cloned in pET20b+ separately.

**Construction of the PAN plasmids**

The gene MA 4268 (PAN) is encoded from nucleotide 5190315 to nucleotide 5191640 on the *M. acetivorans* genome and was cloned into pET20b(+) separately.
(119316-120506) from *P. furiosus* was cloned into pET20b+. The gene AF1976 (1776883-1778079) from *A. fulgidus* was cloned into pET20b+.

**Construction of the ATA plasmid**
The gene MA3029 (ata) is encoded by nucleotides 3'763'307 to 3'764'425 on the *M. acetivorans* genome and was cloned into pET20b(+).

**Construction of the MtuARC plasmid**
MtuARC (*M. tuberculosis*, Rv2115c) is encoded by nucleotides 237461-2376290 and was cloned into pET20b+ vector.

**Expression Tests**
To test the best conditions for expression of the various proteins plasmid DNA was transformed into *E. coli* BL21 (DE3) and *E. coli* BL21 (DE3) Rosetta™ cells. Several clones were grown in LB medium containing ampicillin and/or chloramphenicol to an OD_{600} of 0.7 at 25, 30 and 37 °C. IPTG was added to final concentrations of 0.1 and 0.5 mM. Samples were taken before induction and after 2, 3 and 4 hours of induction. The samples were centrifuged (5 min, 13.2 x 10^3 g, RT), resuspended in water and sonicated. The lysate was centrifuged (5 min, 13.2 x 10^3 g, RT) and the soluble fraction was collected. After washing the pellet twice with water, aliquots of the fractions were applied to SDS-PAGE for detection of protein expression. The identity of the protein was verified by N-terminal sequencing according to the Edman technique.

**Protein purification**
All proteins were purified from 9 L LB culture containing ampicillin (80 µg/ml) and/or chloramphenicol (100 µg/ml). The *E. coli* BL21 (DE3)/Rosetta cells were grown at 37 or 25 °C and induced at OD_{600} = 0.8 with a final concentration of 0.1 mM IPTG. The cells were harvested after 4 hours of induction. The cell pellets were shock frozen in liquid nitrogen. The cells were gently thawed at 4 °C and resuspended in 30 ml equilibration buffer (50 mM Hepes/NaOH pH 7.5 at 4 °C, 2 mM EDTA/NaOH, 20 mM NaCl, +/- 10% Glycerol) containing protease inhibitors (1 mM PMSF, 200 µl SIGMA protease inhibitor cocktail, 12.5 µM Pepstatin A), DNase (4 µl of 50 U/µl stock per 1L culture grown) and 2 mM MgCl$_2$. No protease inhibitors were added when expressing the proteasome genes. The cells were lysed by passing them three times through an EmulsiFlex-C5 High Pressure Homogenizer at 10^5-1.5x10^5 kPa. The cell fragments were pelleted by ultracentrifugation at 45'000 rpm for 60 min. The supernatant was further cleared by passing it through a 0.45 µm sterile filter. For proteins from a thermophilic source, the first purification step was heating of the lysate to 55 or 78 °C for 30 minutes. The next purification step was always an anion exchange column (FastFlowQ Sepharose).
**MacAB**
The cleared lysate was applied to a Fast Flow Q (Sepharose Q) column (60 ml, Amersham Biosciences), that had previously been washed with equilibration buffer (EB) identical to the resuspension buffer (containing no protease inhibitors). Unbound protein was washed out with 2 CV EB. The most proteins eluted at ~350 mM NaCl in a linear gradient (8 CV) from 20 mM–1 M NaCl. The pooled fractions were precipitated with 50-70% saturated ammonium sulfate (AS) solution on ice overnight. The precipitated protein was centrifuged at 18'000 rpm for one hour and solubilized in 5 ml of a second equilibration buffer (50 mM Hepes/NaOH, 50 mM NaCl, 2 mM EDTA/NaOH, 10% Glycerol, pH 8 at 4 °C). After dialyzing against 5 liter of the same buffer for 4 hours the protein was loaded onto a Source 30 Q column (23 ml, Amersham Biosciences) that had been washed and equilibrated with the buffer. The protein eluted at ~260 mM NaCl in a linear gradient from 50 mM to 1 M NaCl.

**MacPAN**
The lysate was applied to a Fast Flow Q (Sepharose Q) column (60 ml, Amersham Biosciences) that had been washed with equilibration buffer (50 mM Tris/HCl pH 7.5 at 4 °C, 2 mM EDTA/NaOH, 20 mM NaCl, 10% Glycerol). Unbound protein was washed out with EB. MacPAN eluted at ~265 mM NaCl in a linear gradient from 20 mM to 1 M NaCl. The pooled fractions were precipitated with 45% ammonium sulfate on ice overnight. The precipitated protein was centrifuged at 18'000 rpm for 60 minutes and solubilized in a second equilibration buffer (50 mM Tris/HCl, 50 mM NaCl, 2 mM EDTA/NaOH, 10% Glycerol, pH 8 at 4 °C). After dialysis against 5 liter of the same buffer for at least 4 hours the protein was loaded onto a Source 30 Q column (23 ml, Amersham Biosciences) that had been washed and equilibrated with the same buffer. MacPAN eluted at ~250 mM NaCl in a linear gradient from 50 mM to 1 M NaCl.

**MacATA**
The lysate was applied to a Fast Flow Q (Sepharose Q) column (60 ml, Amersham Biosciences) that had been washed with equilibration buffer (50 mM Tris/HCl pH 7.5 at 4 °C, 2 mM EDTA/NaOH, 20 mM NaCl, 10% Glycerol). Unbound protein was washed out with this buffer. MacATA eluted at ~150 mM NaCl in a linear gradient. The pooled fractions were diluted and loaded onto a Source 30 Q column (buffer: 50 mM Tris/HCl, 50 mM NaCl, 2 mM EDTA/NaOH, 10% Glycerol, pH 8 at 4 °C). The protein eluted at ~100-230 mM NaCl in a linear gradient. The pooled and concentrated fractions were applied to a HiLoad Superdex200 gel filtration column (buffer: 50 mM Tris/HCl, 300 mM NaCl, 2 mM EDTA, 10% Glycerol, pH 7.5 at 4 °C).

**AfuAlpha**
The protein was expressed in BL21 cells at 28 °C and induced with 0.1 mM IPTG. The cleared lysate was heated for 30 minutes to 78 °C and precipitate was spun down again. The first anion exchange column was a FastFlowQ column (buffers: 50 mM
Tris/HCl pH 7.5, 2 mM EDTA/NaOH, 10 % glycerol, 20 mM NaCl–A or 1 M NaCl-B). A linear gradient from 0-1 M NaCl was run over 800 ml. AfuAlpha elutes at 0.55-0.65 M NaCl and the collected fractions were precipitated with AS to a 70% AS concentration on ice over night. The precipitated protein was centrifuged at 18'000 rpm for 60 min and solubilized in solubilization buffer (20 mM Tris/HCl, 10 ml NaCl, 1 mM EDTA/NaOH, pH 7.5 at 4 °C). After dialysis against 5 liter of the same buffer for at least 4 hours the protein was loaded onto a Superose 6 gel filtration column (CV=60 ml) and eluted at 40-50 ml.

Concentration: $\varepsilon_{280nm} = 12.8 \text{ mM}^{-1}\text{cm}^{-1}$ in 6M GdmCl with 20 mM NaPi pH 6.5

The yield for a wt preparation is about 10.99 mg (i.e. about 1.22 mg/L culture).

AfuBeta
The protein was expressed in BL21 cells at 27 °C and induced with 0.1 mM IPTG. The lysate was applied to a Fast Flow Q (Sepharose Q) column (60 ml, Amersham Biosciences) that had been washed and equilibrated with buffer (20 mM Tris/HCl pH 7.5 at 4 °C, 2 mM EDTA, 50 mM NaCl, buffer B same with 1 M NaCl). Unbound protein was washed out with the same buffer. AfuBeta eluted at 0.3-0.6 M NaCl the collected fractions were precipitated with AS to a 60% AS concentration on ice over night. The precipitated protein was centrifuged at 18'000 rpm for 60 min and solubilized in buffer (50 mM Tris/HCl pH 7.5, 50 ml NaCl, 2 mM EDTA, 10% Glycerol, at 4 °C). After dialysis against 5 liter of the same buffer for at least 4 hours the protein was loaded onto a Superdex 75 gel filtration column (CV= 330 ml). AfuBeta eluted at 40-60 ml. The collected fractions were concentrated with a stirred cell.

Extinction coefficient $\varepsilon_{280nm} = 12.8 \text{ mM}^{-1}\text{cm}^{-1}$ in 6M GdmCl with 20 mM NaPi pH 6.5

The yield for a wt preparation is about 45.23 mg (i.e. about 5 mg/L culture).

AfuPAN
The cleared lysate was heated for 30 minutes to 78 °C and precipitate was spun down again. The first anion exchange column was a FastFlowQ column with buffers (50 mM Tris/HCl pH 7.5, 2 mM EDTA/NaOH, 10 %Glycerol, 20 mM NaCl –A or 1 M NaCl-B). After washing with two column volumes, a linear gradient from 0-1 M NaCl was run over 800 ml. AfuPAN elutes at 0.3-0.4 M NaCl. Pooled fractions were precipitated with saturated ammonium sulfate to get a final AS concentration of 70%. The over night precipitated solution is then centrifuged for 30 minutes in a SS34 rotor at 18 Krpm. Each pellet is solved in MES buffer (25 mM MES pH 6.3, 2 mM EDTA, 12% Glycerol, 30 mM NaCl) and dialyzed against five liter of MES buffer. PAN will then precipitate. Then a second AS precipitation was carried out like before and the resolubilized protein was dialyzed against five liters of Tris/HCl buffer pH 7.5, 300 mM NaCl, 2 mM EDTA, 10% glycerol. The protein was then loaded on a Superose 6 preparative gel filtration column (equilibrated with the same buffer, CV= 60 ml). PAN elutes at 39-60 ml. The pooled fractions were then dialyzed against storage buffer (20 mM TrisCl pH 7.5, 20 mM NaCl, 10 mM MgCl₂, 10% glycerol).
**Extinction coefficient:** $\varepsilon_{280\text{nm}} = 12.8 \text{ mM}^{-1}\text{cm}^{-1}$ in 6M GdmCl with 20 mM NaPi pH 6.5

The yield for a wt preparation is about 5 mg (i.e. about 0.83 mg/L culture).

**PfuAlpha**

The cleared lysate was heated for 30 minutes to 70 °C and precipitate was spun down again. The first anion exchange column was a FastFlowQ column with buffers (20 mM Tris/HCl pH 7.5, 2 mM EDTA/NaOH, 50 mM NaCl –A or 1 M NaCl-B). After washing with two column volumes a linear gradient from 0-1 M NaCl was run over 300 ml. PfuAlpha elutes at 0.3-0.8 M NaCl. Pooled fractions, as seen from gel, were precipitated with saturated ammonium sulfate (AS) to get a final AS concentration of 60%. The overnight precipitated solution is then centrifuged for 30 minutes in a SS34 rotor at 18 Krpm. A second anion exchange column was equilibrated with buffer (50 mM Tris/HCl pH 7, 50 mM NaCl, 1 mM EDTA, buffer B same with 1 M NaCl). This was also the dialysis buffer. The protein was applied to the column and eluted at 0.25-0.45 M NaCl. The protein containing fractions were pooled and concentrated via Amicon stirred cells. For the last purification step the protein was loaded on a 330 ml (CV) Superdex 200 gel filtration column (equilibrated with buffer 50 mM Tris/HCl pH 7.5, 2 mM EDTA, 200 mM NaCl). PfuAlpha eluted at 100-150 ml.

**Extinction coefficient:** $\varepsilon_{280\text{nm}} = 22.9 \text{ mM}^{-1}\text{cm}^{-1}$ in 6M GdmCl with 20 mM NaPi pH 6.5

The yield for a wt preparation is about 77 mg (i.e. about 8.5 mg/L culture).

**PfuBeta**

The cleared lysate was heated for 20 minutes to 70 °C and precipitate was spun down again. The first anion exchange column was a FastFlowQ column (buffers: 50 mM Tris/HCl pH 7.5, 2 mM EDTA/NaOH, 50 mM NaCl –A or 1 M NaCl-B). After washing with two column volumes a linear gradient from 0-1 M NaCl was run over 300 ml. PfuBeta elutes at 0.25-0.7 M NaCl. Pooled fractions as seen from gel were precipitated with saturated ammonium sulfate (AS) to get a final AS concentration of 60%. The overnight precipitated solution is then centrifuged for 30 minutes in a SS34 rotor at 18 Krpm. A second anion exchange column was equilibrated with buffer (50 mM Tris/HCl pH 7, 50 mM NaCl, 1 mM EDTA, buffer B same with 1 M NaCl). This was also the dialysis buffer. The protein was applied to the column and eluted at 0.15-0.25 M NaCl. The protein containing fractions were pooled and concentrated via Amicon stirred cells. For the last purification step the protein was loaded on a Superdex 75 gel filtration column (equilibrated with buffer 50 mM Tris/HCl pH 7.5, 2 mM EDTA, 200 mM NaCl, CV= 330 ml). PfuBeta eluted at 210-350 ml.

**Extinction coefficient:** $\varepsilon_{280\text{nm}} = 21.05 \text{ mM}^{-1}\text{cm}^{-1}$ in 6M GdmCl with 20 mM NaPi pH 6.5

The yield for a wt preparation is about 85.38 mg (i.e. about 9.4 mg/L culture).

**PfuPAN**

The cleared lysate was heated for 20 minutes to 55 °C and precipitate was spun down again. The first anion exchange column was a FastFlowQ column (buffers: 50 mM Tris/HCl pH 8, 2 mM EDTA/NaOH, 50 mM NaCl –A or 1 M NaCl-B). After
washing with two column volumes a linear gradient from 0-1 M NaCl was run over 300 ml. PfuPAN elutes at 0.20-0.54 M NaCl. Pooled fractions as seen from gel were precipitated with saturated ammonium sulfate (AS) to get a final AS concentration of 60%. The over night precipitated solution was then centrifuged for 30 minutes in a SS34 rotor at 18 Krpm. For the last purification step the protein was loaded on a 330 ml (CV) Superdex 200 gel filtration column (equilibrated with buffer 50 mM Tris/HCl pH 7, 300 mM NaCl, 2 mM EDTA, 10% glycerol). PAN elutes at 110-210 ml. 

Extinction coefficient: ε_{280nm} = 18.91 mM^{-1}cm^{-1} in 6M GdmCl with 20 mM NaPi pH 6.5

The yield for a wt preparation is about 29.12 mg (i.e. about 9.7 mg/L culture).

**MtuARC**

The protein was expressed in Rosetta competent cells at 37 °C and induced with 0.1 mM IPTG. The lysate was applied to an anion exchange FastFlowQ column (buffers: 50 mM Tris/HCl pH 8, 2 mM EDTA, 10% glycerol, 1 mM DTT, 50 mM NaCl or with 1 M NaCl). A linear gradient was run from 0-1 M NaCl over 800 ml. MtuARC elutes at 0.3-0.4 M NaCl. The fractions as seen from SDS-PAGE were pooled and an AS precipitation to reach 60% AS final concentration was done. The over night precipitated solution is then centrifuged for 30 minutes in a SS34 rotor at 18 Krpm. The pellet was solved in MES buffer (25 mM MES pH 6.3, 2 mM EDTA, 12% Glycerol, 30 mM NaC, and 1 mM DTT) and dialyzed against 5 liter of the same buffer. The protein was loaded on a, with the same buffer equilibrated, cation exchange column (SourceS Sepharose). The applied sample was washed with five column volumes (CV) of MES buffer and a linear gradient to MES buffer containing 1 M NaCl over 20 CV was run. MtuARC eluted at 0.3 M NaCl. With the collected fractions a second AS precipitation was done like before. The resulting pellets were solved in and dialyzed against 50 mM Tris/HCl Buffer pH 7.5 containing 300 mM NaCl, 2 mM EDTA, 10% glycerol, 1 mM DTT. The final purification step was gel filtration using a Superose 6 preparative column (60 ml). MtuARC elutes at 39-60 ml. After pooling the fractions together the sample was dialyzed against storing buffer (20 mM Tris/HCl pH 7.5, 20 mM NaCl, 10 mM MgCl₂, 10% glycerol, 1 mM DTT). Extinction coefficient: ε_{280nm} = 29.3 mM^{-1}cm^{-1} in 6M GdmCl with 20 mM NaPi pH 6.5

The yield for a wt preparation is about 3.99 mg (i.e. about 0.44 mg/L culture).

**Analytical Size Exclusion**

Analytical size exclusion experiments were done on a Superose 6 analytical column (24 ml) with a flow rate of 0.5 ml/min. A UV signal was detected at 227 and 280 nm. The buffer used (pH 7.5 at RT) contained 50 mM Tris/HCl, 10 mM MgCl₂, 50 mM NaCl, 10% glycerol, 1 mM ATP and 1 mM DTT. For the calibration Thyroglobulin (667 kDa; 2 mg/ml), Dextran blue (2 MDa; 0.5 mg/ml), Ferritin (440 kDa; 0.5 mg/ml) and Aldolase (158 kDa; 2 mg/ml) from the HMW Gel Filtration Calibration Kit were used. 100 µl each were applied onto the column. For analysis the molecular weight of the standards was plotted against the respective elution volume and the slope of the fitted curve was determined.
140 µl of 35.6 µM MacAB was loaded and in a second run 80 µl of 20 µM MacPAN hexamer. In order to check for assembly of the complex, 200 µl of MacAB with 80 µl of MacPAN in 0.5 mM ATP were applied. According to the elution volume the molecular weight and the assembly state of the proteins were determined using the fitted curve of the standards.

Core particle reconstitution tests by analytical size exclusion

α- and β-subunits were mixed in a 1:1 stoichiometry in buffer containing 50 mM Tris/HCl pH 7.5, 2 mM EDTA/NaOH, 200 mM NaCl at RT with a calculated final concentration of the CP of 10 µM. The protein mixture was incubated at 90 °C for 20 minutes, cooled to RT and loaded on the column. For each subunit the same was also repeated separately to have appropriate controls.

For the AfuCP a similar setup was chosen, with the main difference, that the protein mixture was less concentrated (the alpha subunit was in the first trial in a 1.4 excess over beta, later a 1:1 ratio was chosen) and the incubation temperature was lower (78 and 65°C).

A variation of the reconstitution for the AfuCP was to unfold the β-subunit in 5 M guanidine hydrochloride (GdmCl) and allowed it to refold in the presence of the α-subunit. Folding was initiated by fast dilution (1:21) into refolding buffer (700 µl beta unfolded into 15 ml refolding buffer 20 mM Tris/HCl pH 7.5, 200 mM NaCl, 2 mM EDTA). This results in a final protein concentration of 0.25 mg/ml beta and half alpha subunit (0.125 mg/ml) and a final GdmCl concentration of 0.233 M. This solution is then concentrated and loaded on a superose 6 preparative gel filtration column.

Core particle binding tests by analytical size exclusion

First tests to prove an interaction between PAN and the CP were performed using AfuAlpha-rings, as it is known from literature that these subunits are capable to assemble into stable rings. On an analytical Superose 6 column equilibrated with buffer (50 mM Tris/HCl pH 7.5, 300 mM NaCl, 10% glycerol) 200 µl sample mixture containing 5 µM PAN hexamer and 2.5 µM alpha 14-mer were loaded. The reaction mix contained 1 µM ATP and absorbance was detected at 228, 278 and 280 nm. Similar tests were performed using whole assembled CP.

ATPase Assay (Malachite green assay)

This assay measures the formation of Pi upon hydrolysis of ATP according to the method of Lanzetta (Lanzetta et al., 1979).

The assay is based on the formation of a complex between phosphomolybdenum and malachite green, which shows a shift in the absorbance maximum upon formation of the complex (ε650 = 80 mM⁻¹cm⁻¹). This is measured by monitoring the increase in the absorbance signal at 650 nm. Quenching of color development was obtained by addition of 34% Na-citrate after five minutes. Measurement was performed when 10 min of equilibration have elapsed.
The concentration of the ATPase was always 30 µM in buffer (0.05 M Hepes/NaOH pH 7.5 (23 °C), 0.3 M NaCl, 20 mM MgCl₂, 10% (v/v) glycerol). The ATP concentration varied from 100 µM to 30 mM. Michaelis-Menten assays were performed at 37, 60, 65 and 78 °C for the different AAA+ ATPases. Activity profile was measured between 35 °C and 90 °C using 5 mM ATP.

**Components of the color reagent:**

**Malachite Green**

**Hexammonium molybdate (NH₄)₆Mo₇O₂₄**

---

**Edman Degradation**

Proteins were separated on 15% SDS-gels and transferred to a PVDF membrane using Biometra Semidry Blotting System (10 W, 1.5 h). The membranes were stained with Amidoblack. Edman degradation reactions were carried out and analyzed on an Applied Biosystems Procise Sequencer.

**MALDI-MS Analysis**

Mass spectra were recorded using an Ultraflex II MALDI TOF/TOF mass spectrometer from Bruker. Sinapinic acid was used as a matrix in the linear mode. The protein concentration was 6 µM.

**Determination of Protein Concentration**

All protein concentrations were measured using a Cary E3 UV-Visible spectrophotometer (Varian). The extinction coefficients at 280 nm for monomeric subunits are calculated from www.expasy.org using the protparam tool (Gasteiger E., 2005). The native extinction coefficient was calculated in the following way:

\[ \varepsilon_{280\text{nm}} \text{(native protein)} = (A_{280\text{nm, native}}/A_{280\text{nm, denatured}}) \cdot \varepsilon_{280\text{nm}} \text{(calculated)} \]  

Measurement of denatured protein was done in buffer containing 6 M guanidinium chloride.
Sequence alignments
The multiple sequence alignment of CP, ATA, PAN and ARC were carried out with Multalign (http://prodes.toulouse.inra.fr/multalin/) (Corpet, 1988).

Crystallization trials
Crystallisation trials were done with AfuPAN, ATA and His-ATA using the sitting drop method. The protein was dissolved in 10 mM Hepes/NaOH pH 7.5, 10 mM NaCl, 10% glycerol (v/v) at a concentration of 7.8 mg/ml. Needle clusters were obtained after four days at 20 °C form two conditions for AfuATA: from 1.8 M lithium sulfate and 100 mM sodium cacodylate pH 6.5 at 20 °C (conditions from Clear Strategy Screen™ by Molecular Dimensions Ltd) and from 2 M ammonium sulfate, 5% 2-propanol. The drops contained two volumes of protein and one volume of reservoir with a total drop volume of 300 nl equilibrated against 450 µl stock reservoir scaled up later to 6 µl drops equilibrated against 750 µl reservoir.

Electron micrographs
Samples of AfuCP, AfuPAN, AfuATA, PfuCP, PfuPAN, MtuARC were adsorbed to glow discharged carbon-coated copper grids for 30 seconds. The liquid was wicked away with filter paper, negatively stained with 1% (w/v) uranyl acetate and air dried after removal of the excess staining solution. The specimens were examined in a Zeiss EM12 transmission electron microscope at 100 kV or in a Fei Morgani 286 electron microscope at 80 kV. Images were recorded with a CCD camera.

Circular dichroism spectroscopy - Temperature transition measurements of PAN
Far-UV circular dichroism (CD) spectra were measured on a Jasco 710 CD spectropolarimeter at 20 °C. Far-UV spectra were recorded in 0.02 cm quartz cuvettes and averages of 10 measurements were accumulated. Used buffers for measuring native and unfolded CP subunits was 10 mM Tris/HCl pH 7.5 containing 1 mg/ml subunit. The spectra were measured at 25 °C and were corrected for buffer background.
The protein concentration for ATA or PAN hexamer was 5 µM. The spectra were measured at 25 °C in 0.01 M HEPES, pH 7.5 (RT), 0.3 M NaCl and 10% glycerol in a 0.1 cm quartz cuvette. The recorded spectra were averages of 20 measurements and were corrected for buffer background. The temperature transition was recorded from 25 °C to 95 °C at 1 °C/min in buffer in a 0.1 cm quartz cuvette. The temperature dependent unfolding curve was fitted as described in Pace et al. (1998). As the process of unfolding was not reversible for the PAN and ATA complex, because it precipitated after heating to 95 °C, the enthalpy values are not meaningful and only the inflexion point is reported.
Substrate degradation assays

a) Peptidase Assay

The fluorogenic substrate N-Succinyl-Leu-Leu-Val-Tyr-AMC (7-amino-4-methylcoumarin) was used to measure the peptidylglutamyl peptide-hydrolyzing activity of the 20S proteasome. If the peptide is cleaved by the proteasome, the measured fluorescence signal at 410 nm decreases. Fluorescence was measured in a Hitachi F-4500 Fluorescence Spectrophotometer with excitation wavelength at 380 nm and emission wavelength at 410 nm. The sample volume was 1000 µl containing 50 nM CP and 100 µM substrate peptide in buffer (50 mM Hepes/NaOH pH 7.5, 50 mM NaCl). 100 µl samples were taken after different time points and the reaction was stopped by adding 200 µl of ice cold pure ethanol.

b) GFP unfolding experiments

Unfolding of GFP-SsrA by ATA in presence of GroEL (mutant D87K) trap or core particle. Unfolding of GFP is measured at different temperatures. Folded GFP shows a strong fluorescence signal, unfolded GFP shows no fluorescence signal. The measurements were done in a quantamaster (QM-7/2003) fluorimeter from PTI. For GFP-SsrA unfolding test with ARC the sample volume was 100 µl containing 6.45 µM hexamer ARC, 1 µM GFP, 2.5 mM ATP and 1 mM DTT. The reaction was measured at 25 and 37 °C for 60 minutes. For GFP-SsrA unfolding tests with PAN or ATA the sample volume was 100 µl containing 1 µM chaperone, 0.5 µM CP and 1 µM GFP. Measurements were done with or without 5 mM ATP. PAN activity was also measured with 2 mM ATPγS at 65 °C. Excitation wavelength was 400 and emission wavelength 510 nm. Measurements were performed at different temperatures (35, 45, 50, 60 and 65 °C).

Lysate binding tests (CoImmunoPrecipitation)

Co-Immunoprecipitation (Co-IP) has been carried out with His-tagged AfuATA protein as the bait and archaeal (A. fulgidus) whole cell lysate as prey. The Co-IP kit from Pierce was used. This kit is based on coupling an antibody to the gel and the recombinant protein is then bound to the antibody. The kit contains an amine-reactive gel, reagents for direct covalent immobilization of the primary antibody and Handee™ Spin Cup Columns. The entire procedure is performed in the spin columns. Monoclonal antibody to Histidine TAG Mouse IgG1 (hexa His) [Acris Antibodies] was used. The prey lysate was prepared from 1 g of frozen archaeal cell pellet. The pellet was resuspended in 1 ml PBS buffer (1 L buffer pH 7.4, 8 g NaCl, 0.2 g KCl, 1.44 g Na₂HPO₄, 0.25 g KH₂PO₄, pH adjusted with 1 M HCl, autoclaved) spun down again and resuspended a second time in 0.5 ml PBS buffer. 5 µM His-tagged protein is then added to the cell lysate and incubated at 80 °C in presence of 2 mM ATPγS for 10 minutes. The kit procedure was then followed to couple the bait protein from the lysate mixture to the antibody-charged spin column resin and to wash and elute any bait-prey complexes. Eluted samples were tested by SDS-PAGE analysis and Edman degradation.
**Binding test with *A. fulgidus* lysate and His-ATA (pull down assay)**

To identify potential binding partners of AfuATA His-tagged ATA was used as bait protein and whole cell lysate as prey. Ni-chelating affinity chromatography was then used to isolate bait-prey complexes. 1 g *A. fulgidus* cells were solved in 2 ml buffer (50 mM HEPES pH 7.5, 0.5 M NaCl, 10% glycerol, 1 mM DTT), five minutes of sonication and ultracentrifugation at 45 krpm for 60 minutes followed. One ml cleared lysate was mixed with 1 mM ATPγS and 4 µM ATA (hexamer) and incubated for 30 minutes at 60 °C. The mixtures was loaded on HiTrap Ni-chelate affinity column (CV= 1 ml), previously equilibrated with buffer. The unbound sample was washed with 10 CV with the same buffer containing 0.5 mM ATPγS and 20 mM imidazole. Elution was carried out in two steps:

1. step: elution with 2 CV of buffer containing 700 mM imidazole (~ concentration where ATA elutes).
2. step: elution with 2 CV of buffer containing 1 M imidazole.

Finally the collected 1 ml fractions were concentrated by chloroform-methanol precipitation and loaded on denaturing SDS-PAGE, semidy blotting on PVDC-sequencing membrane for N-terminal sequencing.

**Growth of *Methanosarcina acetivorans* cells**

The cell material we had from *Archeaglobus fulgidus* was limiting in our testings and the fact that possible interaction between complexes is only occurring at high temperatures similar experiments were planed to be done with a mesophilic organism. For further lysate binding tests with MacATA whole cell lysate was needed. For the circumstance that *Methanosarcina* cells are not available to by in great amounts, growth in our laboratory was successfully established.

For further lysate testing a culture of *M. acetivorans* was ordered at the DSMZ (*M. acetivorans* C2A, DSM 2834) and grown as described. The medium was prepared as described from ATCC (protocol ATCC Medium 1355) with differences in filtering the media instead of autoclaving it and installing only a nitrogen flow instead of a CO2-nitrogen gas mixture. Another variation was the growth under shaking conditions (130 rpm at 37 °C) and the transfer of the cells without an anaerobic glove box. 1 M DTT was used to protect this strictly anaerobic organism from oxygen. Growth was detected measuring turbidity at OD 600 nm and by visualizing the cells using light microscopy.

The cells were grown while shaking at 130 rpm at 37 °C in a HT Infors Lab shaker (Infors AG, Bottingen/Switzerland). The bacterial growth was determined measuring optical density at 600 nm with a Novaspec II visible Spectrophotometer. Samples were taken at different time points. Glycerol stocks of *M. acetivorans* cells were prepared as described (Sowers et al., 1993). At different OD<sub>600</sub> values gaseous samples were collected in a balloon and the volume was measured by putting the balloon into a water filled beaker to determine the volume change through water displaced. At OD<sub>600</sub> above 0.5 cells were harvested by centrifugation (6'000 rpm for 15 minutes) and the resulting pellet was weighed to determine the wet and dry
weight amount of cells. Under these growing conditions a doubling time of 18 hours was calculated. Growth conditions can still be optimized following the protocol described previously (Rother et al., 2005) and also by increasing concentrations of methanol. The observed growth rates could be due to the shaking of the bottles and also to remaining nitrogen bound to media contents, as media preparation was performed different from existing protocols. In addition, the large volume in one serum bottle might not be optimal. From one liter of media the wet weight amount of cells ranged from 1.1 -1.96 grams and resulted in 0.26 gram dry weight of one liter high density grown culture. The fact, that we use whole lysate for proteomic studies, methods are limited for monitoring viable cells. A sample of cells was loaded on a 15 % reducing SDS-Gel displaying the proteome of *M. acetivorans*. As a proof that cells were growing properly the amount of produced gas was measured and shown to correlate with increasing OD₆₀₀ values. Additionally a gaseous sample was analyzed and methane was detected (all data not shown).

The used device was a Trace MC, Thermo Electron Corporation with a GS-Gaspro column (J&W Scientific, length: 60 m, inner diameter 0.32 mm). The gas flow was 0.8 ml/min helium and the split ratio was 1:400 with used temperature of 10 minutes at 40 °C isotherm and a Quadrupol mass spectrometer detector. The collected and analyzed gas was composed of 70-75% nitrogen, 22-30% methane and 0.4-1% carbon dioxide. This used method was only half quantitative but good enough to separate methane from air.
Results and Discussion

I Cloning, expression and reconstitution of functional PAN-proteasome systems of *A. fulgidus*, *P. furiosus* and *M. acetivorans*

1. Genomic gene location, sequence analysis and comparison of PAN and proteasomal α- and β-subunits from *A. fulgidus*, *P. furiosus* and *M. acetivorans*

Archaea contain in their genomes genes encoding the individual components of the proteasome degradation machinery. These encompass the two subunits of the proteasomal core (α and β) as well as a regulatory AAA-ATPase referred to as PAN. For all three archaeal chaperone proteasome complexes no clustering of the according genes in a specific area of the genome could be found.

**P. furiosus**

```
PF0115 (PAN) -\-\- PF0159 (beta2) -\-\- PF1404 (beta1) -\-\- PF1571 (alpha)
```

**A. fulgidus**

```
AF0481 (beta) -\-\- AF0490 (alpha) -\-\- AF1976 (PAN)
```

**M. acetivorans**

```
MA1779 (alpha) -\-\- MA3873 (beta) -\-\- MA4123 (PAN) -\-\- MA4268 (PAN2)
```

Figure 16: Genomic location of the chaperone proteasome genes. Right hand arrows indicate the + coding strand and left handed arrows −coding strand.

Figure 16 shows the genomic location and gene numbers of the proteasomal genes in the studied organisms.

![Sequence alignment for PAN proteins](image)

Figure 17: Sequence alignment for PAN proteins (*A. fu*. *Archaeoglobus fulgidus*, *M. ac*. *Methanosarcina acetivorans*, *P. fu*. *Pyrococcus furiosus*.)
Multiple sequence alignment for the proteins from the three organisms was done for PAN (Fig. 17), α (Fig. 18) and for β (Fig. 19). The conserved regions are highlighted and boxed. The predicted N-domain of the PAN protein, with its assumed function in substrate recognition and initial binding is marked in the sequence alignment. The Walker A and B motifs contain the ATP binding sites and the second region of homology (SRH) contains the sensor-1 motif which is involved in ATP hydrolysis. The sequence alignments for all three archaeal α-subunits show the conserved YDR-motif which is probably involved in the gating mechanism (Groll, et. al, 2003). The contact surfaces from one subunit (n) and from the neighboring (n+1) subunit of the proteasome interact with the chaperone part (Fig. 18).

A sequence alignment of all three β-subunits with propeptides is displayed in figure 19. The prosequences of *M. acetivorans*, *A. fulgidus* and *P. furiosus* β-subunits have all approximately the same length (8-10 amino acids) and the catalytic Thr1 is followed by two additional Thr. The prosequence is auto catalytically cleaved off and the catalytic residues are indicated. The catalytic important pockets P1 (a small unpolar substrate binding pocket) and P3 are indicated as well as the hydrophobic S1 substrate binding pocket.
2. Cloning and expression of the proteins of interest.

To study the archaeal PAN-proteasome system \textit{in vitro}, the individual components were cloned from genomic DNA of three organisms, the hyperthermophilic \textit{A. fulgidus} and \textit{P. furiosus} and the mesophilic \textit{M. acetivorans}. The two subunits making up the proteasome core particle (CP) were cloned separately into expression vectors in case of the \textit{A. fulgidus} and \textit{P. furiosus} \(\alpha\)- and \(\beta\)-subunit genes. After separate expression and purification, the \(\alpha\)- and \(\beta\)-subunits were reconstituted \textit{in vitro} into the assembled core particle. For the \textit{M. acetivorans} \(\alpha\)- and \(\beta\)-subunits a different strategy was chosen. The two genes were cloned into the pETDuet vector, so that they could be co-expressed. This allows the formation of the fully assembled core particle already during the expression. Table 1 gives an overview of the generated constructs and the expected molecular weight of the over-expressed proteins.

<table>
<thead>
<tr>
<th>Name</th>
<th>Vector</th>
<th>Insert size (bp)</th>
<th>aa</th>
<th>MW (kDa)</th>
<th>Ext. coeff.</th>
<th>pI</th>
</tr>
</thead>
<tbody>
<tr>
<td>AfuA</td>
<td>pET 20 b (+)</td>
<td>741 Ndel/BamHI</td>
<td>214</td>
<td>23.7002</td>
<td>12.8</td>
<td>4.96</td>
</tr>
<tr>
<td>AfuB</td>
<td>pET 20 b (+)</td>
<td>641 Ndel/BamHI</td>
<td>213</td>
<td>23.4179</td>
<td>14.0800</td>
<td>5.28</td>
</tr>
<tr>
<td>processed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AfuPAN</td>
<td>pET 20 b (+)</td>
<td>1197 Ndel/BamHI</td>
<td>398</td>
<td>44.964</td>
<td>12.8</td>
<td>5.43</td>
</tr>
<tr>
<td>PfuA</td>
<td>pET 21 d (+)</td>
<td>783 Nco I/ BamHI I</td>
<td>260</td>
<td>29.0991</td>
<td>21.050</td>
<td>5.09</td>
</tr>
<tr>
<td>PfuB</td>
<td>pET 20 b (+)</td>
<td>590 Ndel/BamHI</td>
<td>196</td>
<td>21.6488</td>
<td>22.9</td>
<td>5.16</td>
</tr>
<tr>
<td>processed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PfuPAN</td>
<td>pET 20 b (+)</td>
<td>1190 Ndel/BamHI</td>
<td>396</td>
<td>44.8048</td>
<td>17.21</td>
<td>6.03</td>
</tr>
<tr>
<td>MacA</td>
<td>pETDuet MCS I</td>
<td>743- Ncol-BamHI I</td>
<td>247</td>
<td>27.0196</td>
<td>12.8</td>
<td>5.00</td>
</tr>
<tr>
<td>MacB</td>
<td>pETDuet MCS II</td>
<td>632-Nde I- Bgl II</td>
<td>210</td>
<td>22.8801</td>
<td>12.8801</td>
<td>5.5</td>
</tr>
<tr>
<td>processed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MacPAN</td>
<td>pET 20 b (+)</td>
<td>1326-Nde I-BamHI I</td>
<td>441</td>
<td>49.2449</td>
<td>8.96</td>
<td>5.52</td>
</tr>
</tbody>
</table>

Table 1: Overview of the cloned constructs.

Expression of the archaeal PAN and proteasome subunits was carried out in BL21(DE3) Rosetta cells harboring the p-RARE plasmid providing rare t-RNAs. The use of this expression strain increased the yield of the over-expressed proteins because all of the genes contain codons that are rarely used in \textit{E. coli}.

For the proteins of the thermophilic organisms heating of the cleared lysate was generally the first purification step. This was tested with a small amount of cell lysate and sequential heating for 10 minutes. The samples were analyzed via SDS-PAGE (data not shown). The temperatures started from 50 °C in 5 ° steps to 70 °C. From the SDS-PAGE analysis it was visible, which is the optimal temperature for a purification effect, where the protein of interest is still soluble (see methods). This gave also an indication about the stability of the protein. All proteins were further purified by two anion exchange chromatography steps on column material with different bead properties and at slightly different pH. Due to the rather low pI of all proteins used in this study cation exchange chromatography was not possible for these proteins, even when the protein of interest was in the flow through fraction, no purification effect was observed. As the last step, additional preparative size exclusion was
carried out. The purified proteins were subjected to mass spectroscopy and Edman degradation to confirm the correct identity and size. PfuA and AfuA subunit degradation gave the correct result, without the N-terminal methionine in case of the former. The PfuB and the AfuB subunits were shown to still possess their propeptides. For the M. acetivorans β-subunits N-terminal sequencing revealed processed N-termini lacking the propeptide. This was a good indication that the in vivo assembly of the M. acetivorans core particle had, in fact, taken place.

3. Core particle assembly and reconstitution

The proteasome core particle is made of the two types of proteasomal subunits α and β to form a stacked structure of four heptameric rings, the two inner β-rings carrying the catalytic residues and an outer α-ring on each end of the β core complete the fully assembled proteasome. In case of the expression of M. acetivorans proteasomal subunits from the pETDuet vector, we expect the fully assembled core particle to be formed during the expression. This was tested by analytical size exclusion chromatography on an analytical Superose 6 gel filtration column (Fig. 20).

The elution profile shows a single peak at 13.2 ml with a minor shoulder. Calibration with molecular weight standards showed that this elution volume roughly corresponds to the fully assembled 28mer.

In case of the two hyperthermophilic proteasomes from A. fulgidus and P. furiosus, the subunits were expressed separately and had to be reconstituted to generate the active core particle. The assembly route for archaeal CPs has been shown to proceed from spontaneously formed α heptamer rings acting as ‘templates’ for the β-subunits to form so called half-proteasomes (α7β7). The half proteasomes then loosely assemble into α7β7|β7α7 particles. Only at this stage, the propeptides can be cleaved off leading to the mature, active core particle (Groll et al., 2005).
The reconstitution for the hyperthermophilic core particles was done by mixing and incubating the purified α- and β-subunits in a 1:1 ratio at 85 °C. The elution profile of the analytical size exclusion run (Superose 6) in figure 21 shows that for the *Pyrococcus furiosus* proteins only a portion of the α- and β-subunits assembled into the core particle apparent as a shoulder at 11 ml retention volume.

![Figure 21](image)

The SDS-PAGE gel analysis of this fraction clearly shows bands for both the α- and the processed β-subunits. The pronounced peak eluting at 13-15 ml most likely corresponds to half-proteasomes. The bands seen on the gel for the fractions eluting at 13-15 ml represent α- and unprocessed β-subunits expected for half-proteasomes and also some processed β-subunits due to strong overlaps with the peak of the CP. Some of the α-subunits and β-subunits elute at the positions for α-heptamer rings (15 ml) and unassembled β-subunits (21 ml, gel not shown), respectively. The peak eluting at 21 ml likely also contains completely unassembled α-subunits, since the α-subunit alone also elutes in two peaks corresponding to α7 (15 ml) and unassembled α-subunits (21 ml, gel not shown).
Figure 22: Elution profile of AfuCP reconstitution using Superdex 200 gel filtration column. The corresponding gel with the collected fractions is displayed below.

The elution profile of the *A. fulgidus* CP reconstitution experiment shows two separated peaks. The first peak eluting at 11 ml corresponds to the fully assembled CP and shows bands with equal intensity for α- and processed β-subunits. The second peak eluting at 13 ml very likely corresponds to half-proteasomes, since no free β-subunit peak can be detected. The α-subunit band and the band of unprocessed β-subunit overlay on the SDS-PAGE analysis (Fig. 22). This explains the single band seen on the gel.

In a variation of the experiment α-subunits from *A. fulgidus* were mixed with β-subunits of *P. furiosus* in a 1:1 ratio to see, if heterologous complexes could be assembled. No fully assembled core particles could be detected in this case (data not shown). The conclusion is that the α|β interaction at the contact surface is very specific for each organism and cannot be exchanged by a subunit from a different organism. No mixed complexes can be obtained.

For further proof of the assembly state of the purified and reconstituted proteasomal core particles, the complexes were investigated by electron microscopy.
preparation 50 nM assembled core particles were used. Core particles are very well suited for investigations by negative stain EM, because their cylindrical shape is easily recognizable at even low resolution. Top views of the core particles present as rings with dark centers, side views present as four stacked striations. *P. furiosus* similarly reconstituted CP gave a few nicely visible side views where all four stacked rings could be easily counted, but the majority of the particles present as top views (Fig. 23). Why this is the case we cannot explain. This may be due to the higher affinity of the α-N-termini to the grid surface. Another possibility could be dissociation of the proteasomes into half proteasomes during preparation of the EM-probes.

The AfuCP gave compared to the PfuCP similar results with even less visible side views of the proteasome stacked rings. The assembled core particle alone mainly shows up in top views, when visualized this way (Fig. 24). As will be shown later, in presence of the chaperone particle more side views are detected.
4. The core particles from *A. fulgidus*, *P. furiosus* and *M. acetivorans* display peptidase activity

In absence of their chaperone ATPase partners, proteasomal core particles are unable to degrade proteins. However, small peptides can gain access to the proteolytic chamber in absence of the ATPase partner and are cleaved. The core particles alone, therefore, display peptidase activity. To test whether our purified and reconstituted core particles are catalytically active, we measured their peptidase activity using a fluorogenic peptide (Suc-Leu-Leu-Val-Tyr-AMC, (Stein et al., 1996)). As shown in figure 25 all three core particles cleave the fluorogenic model peptide leading to an increase in fluorescence intensity.

For the *Archaeoglobus* and *Pyrococcus* proteasomes a temperature profile of peptidase activity was measured. The slopes of the regression from each temperature gave the initial velocity of peptide hydrolysis and these were plotted against the temperature. Figure 26 shows that for PfuCP the activity profile is more or less linear, while for AfuCP there is a clear temperature optimum at 85 °C and the activity drops sharply above 85 °C.
5. Assembly states of the regulatory particle PAN from *A. fulgidus*, *P. furiosus* and *M. acetivorans*

The regulatory triple-A partners of the proteasomes usually assemble into ring-shaped oligomers and are functionally active in this form. In order to test if the various PAN proteins are in their assembled state, analytical gel filtration analysis was carried out. For the *M. acetivorans* PAN a first indication of higher oligomerisation already came from its late elution in the preparative size exclusion chromatography step used in the purification procedure. This was then confirmed by analytical size exclusion on an analytical Superose 6 gel filtration column. A 24 ml Superose 6 analytical column was calibrated with Dextranblue (2'000 kDa), Thyroglobulin (667 kDa), Ferritin (440 kDa) and Aldolase (158 kDa), each in a separate run. Their molecular weight was plotted against their elution volume. By
using the slope of the fitted linearization the molecular weight of the proteins could be calculated from their elution volume. MacPAN has a monomeric weight of 49.2 kDa and would, therefore, be expected to elute late from the column, if it is in monomeric form. MacPAN eluted at 15.7 ml, in between the calibration standard Ferritin (440 kDa) and Aldolase (158 kDa) (Fig. 27). The molecular weight of the MacPAN complex determined as described before was 345 kDa. This is certainly too large to represent monomers or dimers. The calculated weight of the expected hexamer is 295 kDa. It is possible that the molecular weight is overestimated by this method due to unspecific interaction with the column material. In principle this method is not the ideal method for molecular weight determination and gives only estimations. ATP was not included in this run. There was, however, no detectable difference upon repeating this run in presence of ATP. Thus MacPAN apparently assembles into oligomers in absence of nucleotide.

The two hyperthermophilic PAN complexes were also tested by analytical size exclusion and here shown to run in an assembled form both in absence and presence of nucleotide (data not shown). While comparison with the calibration standards slightly underestimates the molecular weight of the PfuPAN hexamer (235 kDa determined versus the calculated 269), the AfuPAN complex appeared again to be larger than the calculated hexamer, similar to what was found for MacPAN. The electron micrograph in figure 28 shows that PfuPAN assembles into ring-shaped complexes in absence of ATP that present in top or bottom views. The individual rings have diameters of about 120-140 Å.
6. The regulatory PAN complex displays ATPase activity

The chaperone partners of the proteasome belong to the AAA+ protein family. The hallmark of this family is the AAA module, containing the P-loop domain with the signature motifs (Walker A and B) for nucleotide binding and hydrolysis. To test if the formed PAN complexes are catalytically active, their ATPase activities were analyzed using the malachite green assay (Lanzetta et al., 1979). This ATPase assay is based on formation of a complex between inorganic phosphate with ammonium molybdate and malachite green and the complex shows a shift in the absorbance maximum to 650 nm (ε = 80 mM⁻¹ cm⁻¹). The inorganic phosphate concentration was determined using the increase of absorbance at 650 nm. The advantage of this test is that it can be carried out over a wide range of temperatures.

First the temperature-dependence of the ATPase activity was examined. The activity of the PfuPAN complex was measured from 25 to 90 °C at eight different temperatures. The data displayed in figure 29 have been corrected for spontaneous ATP hydrolysis. The PfuPAN complex shows highest ATPase activity at around 60 °C and at higher temperatures the activity drops sharply due to loss of structure. To confirm this, temperature denaturation curves for PAN were recorded to measure the transition from the folded to the unfolded state using the circular dichroism signal at 222 nm. The inflexion point for PfuPAN is at 65 °C (Fig. 29) and most of the protein is irreversibly aggregated, rendering the transition irreversible.

Full Michaelis-Menten analysis was carried out at 60 °C with ATP concentration varying between 200 µM to 20 mM. The concentration of PAN was kept constant at 3 µM for the hexamer. From this analysis the catalytic rate constant (kcat) for ATP hydrolysis at 60 °C was determined to be 4.07 ± 0.095 min⁻¹ (per hexamer) and the Michaelis constant (KM) was determined as 0.24 ± 0.03 mM.
The temperature transition for AfuPAN measured by far-UV CD-spectra showed a transition midpoint of 80 °C (Fig. 30). This is higher than compared to PfuPAN. For the AfuPAN ATPase measurements a temperature of 65 °C was chosen to ensure that the AfuPAN complex remained stable over the course of the measurement. From this analysis the catalytic constant for ATP hydrolysis at 65 °C was determined to be $20.55 \pm 0.46 \text{ min}^{-1}$ (per hexamer) and the apparent dissociation constant $K_M$ was determined as $1.72 \pm 0.15 \text{ mM}$. The corresponding plot is displayed in figure 30.

Figure 30: AfuPAN Michaelis Menten measurement of ATP hydrolysis (left side). Temperature transition of AfuPAN (right side).
7. Formation of PAN-CP complexes

From literature it is known that for other organisms the proteasome core particles interact with PAN to form the fully assembled chaperone proteasome (Benaroudj and Goldberg, 2000). In order to find out if the chaperone hexamer binds to the proteasome core particle, both protein complexes, PAN and the core particle were mixed in a ratio of 2:1 and applied to the analytical Superose 6 gel filtration column. Ideally one chaperone ring should bind to each end of the core particle. The calculated molecular weight of the fully assembled chaperone proteasome complex is 1'276 kDa. If only one chaperone was bound, the molecular weight should amount to around 981 kDa.

The proteins eluted in two separate peaks corresponding to the proteasome core particle and the chaperone complex (data not shown). Each binding partner eluted after exactly the same volume as when it was run separately. Not even partial assembly was detectable.

In transmission electron microscopy we were able to see AfuPAN ring complexes (data not shown). We wanted to test if we could also see the proposed AfuPAN-core particle complexes.

To test this, the proteins were incubated at 80 °C for 15 minutes and put on the grid immediately. The pictures were made at 25'000x magnification and 100 kV intensity. It is known from literature that the interaction between these huge complexes is
transient (Voges et al., 1999). Hence it is not surprising that only a few possible complexes between PAN and CP could be observed with this method (Fig. 31). The first publication reporting visualization of PAN-proteasome complexes by EM claimed they see this complex formation in less than 0.5 % of particles (Wilson et al., 2000). In our case the complexes which are visible are in the range of 3-5 %. We conclude that additional factors are necessary for stable complex formation, or maybe substrates need to be present.

After the publication of stable complex formation between *M. jannaschii* PAN and *T. acidophilum* proteasome (Smith et al., 2005), we thought that we could also try to make PAN-proteasome complexes with subunits from two different species, namely with *A. fulgidus* PAN and *P. furiosus* CP. As described by Smith et al., we used a 4:1 ratio of PAN over CP and incubated the sample mixture at 60 °C for 15 minutes in presence of 10 µM ATPγS or GTPγS. We could not see an increased amount of AfuPAN-PfuCP complexes in this experiment (Fig. 32) compared to the AfuPAN-AfuCP mixture, with about a less amount of possible formed complexes.

![Figure 32: Negative stain images of mixed complexes containing 100 nM AfuPAN and 25 nM PfuCP. The red arrow heads indicate possible formed complexes, magnification 25'000x at 80KV intensity. The samples were fixed with 2.5% glutaraldehyde and stained with 1% uranyl acetate.](image)

Next, we tried to make the MacPAN-CP complex. This mesophilic representative does not require high temperatures for activity and maybe for stable interaction between the complexes. As can be seen from the strong background of damaged protein in figure 33, the mixture of *M. acetivorans* PAN-proteasome did not endure the process for EM imaging. Only single side views of the CP and only a few top views of some complexes can be seen. With the used microscope resolution it is impossible to distinguish the hexameric PAN complex from the heptameric CP assembly and therefore from the visible top views it is hard to judge which top view corresponds to PAN and which to the CP. But there are still possibilities to improve
EM results for MacPAN images in the future, e.g. different stain or shorter staining time and faster transfer of the proteins onto the grid.

Figure 33: Two merged electron micrographs of MacCP. The white arrowheads point some selected top views, whereas the red boxes point some side views. Magnification 25'000x and 80 kV intensity.

8. AfuPAN interacts with the AfuCP to degrade GFP-SsrA

Degradation of protein substrate requires association and concerted action of the core particle and its chaperone partner PAN. PAN has been shown previously to recognize SsrA-tagged proteins *in vitro*, likely due to the unstructured nature of the SsrA-tail (Tu et al., 1995). The chaperone PAN unfolds the substrate proteins under expense of ATP, followed by translocation into the core particle, where the protein degradation takes place. We could not detect stable interaction between the core particles and PAN chaperone complexes, neither for the mesophilic *M. acetivorans* nor the hyperthermophilic *A. fulgidus* and *P. furiosus* by gel filtration analysis at room temperature (data not shown). For the hyperthermophilic proteins this could be due to the limitation of these experiments to low temperature with respect to their optimal growth temperatures. Furthermore, a more transient interaction might not be detectable with this method. To test for PAN-CP driven protein degradation we used the model substrate GFP-SsrA originally developed for the ClpAP chaperone-protease system (Weber-Ban et al., 1999). We measured the fluorescence change of SsrA tagged GFP after addition of either just PAN chaperone in presence of ATP or both the PAN chaperone and the *A. fulgidus* core particle and detected efficient unfolding of the substrate proteins (Fig. 34). This suggests that PAN and the CP interact, even if only transiently, to achieve the unfolding and degradation of GFP-SsrA. This activity is clearly ATP-dependent, since in absence of ATP no fluorescence decrease is detected upon incubation of GFP-SsrA with the AfuPAN-CP complex.
The MacPAN does not appear to interact with the core particle in the analogous assay as no significant decrease in fluorescence can be detected.

Figure 34: GFP unfolding experiments. Upper left: Results from AfuPAN unfolding experiments with or without CP. Upper right: MacPAN-CP GFP unfolding at 30 °C (indicated +/- ATP). PAN measurements (SV= 100 µl, 2 or 1.66 µM PAN, 2 µM GFP-SsrA, 1 mM DTT, 2 or 5 mM ATP or GTP, 50 mM Tris/HCl pH 7.5, 10 mM MgCl$_2$, baseline GFP + ATP in Buffer).
II Cloning of the *M. tuberculosis* proteasome and its regulatory AAA partner ARC and characterization of ARC

1. Sequence analysis and similarities with other AAA+ proteins

Certain *Actinobacteria* contain in their genomes in addition to the usual genes for Clp chaperone-proteases also genes for proteasomes acquired likely by horizontal gene transfer. *Mycobacteria* belong to this group of bacteria. We are studying the proteasome system of *M. tuberculosis*. The availability of the *M. tuberculosis* full genome sequence allows investigating the genome organization for the proteasomal genes and the putative associate ARC in this organism.

![Figure 35: The gene organization of *M. tuberculosis* proteasome genes and ARC gene.](image)

The genome organization differs highly from the previously described archaeal genomes. In *Mycobacteria* the proteasome encoding genes (prcB and prcA) are adjacent to each other and therefore an operon organization similar to *R. erythropolis* is proposed (Wolf et al., 1998). The VIMSS (Virtual Institute for Microbial Stress and Survival, http://www.microbesonline.org/operons/gnc2234.html#bOp) operon prediction database showed with a high score that at least the proteasomal genes plus three following genes are localized within an operon and the gene for the putative interacting AAA ATPase ARC lies upstream of the proteasomal subunit genes separated only by a few open reading frames. In fact, in all *Actinomycetal* genomes examined so far that contain the ARC and proteasomal genes ARC genes cluster together with genes encoding the α- and β-subunits of the proteasome (Zhang et. al. 2004)

The ARC protein seems more related to VAT than to PAN. ARC shows only weak homology with proteasomal ATPases in eukaryotes but it is the proposed proteasome partner. ARC contains also the second region of homology (SRH) (Fig. 36). Although ARC clearly belongs to the AAA+ protein family it has no detectable sequence similarity to other AAA+ proteins outside the AAA domain. It contains two insertions in the conserved domain, which are not observed in other members of the family. Unlike for the related PAN protein, ATP-dependent proteasome activation has not been demonstrated for ARC (Wolf et al., 1998) and remains to be investigated.
2. Construction of the over-expressing plasmids and purification of MtuARC

To study the eubacterial ARC-proteasome system in vitro, the genes for *M. tuberculosis* ARC (AAA ATPases forming Ring-shaped Complexes) and the two proteasomal subunits (α and β) were amplified from genomic DNA and cloned into the listed over-expression vectors. The results from cloning were verified by DNA-sequencing.

<table>
<thead>
<tr>
<th>Name</th>
<th>Vector</th>
<th>Insertsize (bp)</th>
<th>aa</th>
<th>MW (kDa)</th>
<th>Ext. coeff.</th>
<th>comments</th>
<th>pl</th>
</tr>
</thead>
<tbody>
<tr>
<td>M. tu. A</td>
<td>pETDuet</td>
<td>746-NcoI-EcoRI</td>
<td>248</td>
<td>26.881</td>
<td>14.08</td>
<td>internal BamHI</td>
<td>5.41</td>
</tr>
<tr>
<td>M. tu. B</td>
<td>pET 20 b(+)</td>
<td>875-NdeI-BamHI</td>
<td>291</td>
<td>30.305</td>
<td>22.9</td>
<td></td>
<td>4.65</td>
</tr>
<tr>
<td>processed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M. tu. ARC</td>
<td>pET 20 b(+)</td>
<td>1829-NdeI-BamHI</td>
<td>609</td>
<td>67.401</td>
<td>29.3</td>
<td></td>
<td>4.89</td>
</tr>
</tbody>
</table>

Table 2: Overview of the cloned and used constructs.

Expression of the eubacterial ARC protein was carried out in BL21(DE3) Rosetta competent cells harboring the p-RARE plasmid providing rare t-RNAs. The protein was purified by a first anion exchange chromatography, followed by cation exchange chromatography. The last purification step was preparative size exclusion chromatography. The purified MtuARC protein was verified by Maldi mass spectrometry analysis as well as N-terminal sequencing.
3. Analysis of the assembly state of ARC

ARC is an ATPase associated with various cellular activities (AAA) that forms hexameric rings resembling the eukaryotic complex p97/valosin-containing protein (VCP) (Darwin et al., 2005). To test if the MtuARC assembled into oligomers, we applied it onto analytical Superose 6 gel filtration column. ARC eluted at 12.2 ml and from comparison with the molecular weight standards this results in a determined molecular weight of about 603 kDa (data not shown). This result lies above the calculated molecular weight of 404 kDa and could be due to nonspecific interaction with the column material. To further test the assembly state we took negative stain electron microscopy images.

The used concentration for MtuARC electron microscopy was 50 nM hexamer. Images were obtained by using a magnification of 25’000x at 80 kV. The determined diameter of ARC is within 12-14 nm (Fig. 37). Compared to the archaeal chaperone complexes ARC seems to be more sensitive to this method displaying more background of crashed and damaged proteins upon binding to the EM grid. Darwin and co-workers showed by Cryo-EM that ARC has a dome shape and similarity in shape to p97 (Darwin et al., 2005), even though p97 has two AAA+ domains.

![Figure 37: Two electron micrographs of MtuARC complexes. The arrowheads point some selected top views.](image-url)
4. MtuARC displays ATPase activity

As ARC belongs to the AAA+ family containing the Walker A and B motif, as well as the second region of homology (SRH) we were interested if ARC has ATPase activity. Like for the archaeal PAN proteins the malachite green assay was used to determine the ATPase activity of this AAA+ protein. The results are shown in figure 38.

![Figure 38: Michaelis Menten curve of MtuARC measured at 37 °C and corrected for ATP auto-hydrolysis.](image)

The MtuARC measurements were performed at 37 °C to be under physiological conditions. From this analysis the catalytic constant for ATP hydrolysis at 37 °C was determined to be $5.6 \pm 0.2 \text{ min}^{-1}$ (per hexamer) and the apparent dissociation constant $K_M$ was determined as $0.19 \pm 0.04 \text{ mM}$.

5. MtuARC does not recognize GFP-SsrA

Degradation of protein substrate requires association and concerted action of the core particle and its proposed chaperone partner. As the SsrA-tag is recognized by the bacterial Clp-chaperone protease and also by the archaeal PAN-proteasome, we tested if ARC can recognize and unfold GFP-SsrA. As can be seen from figure 39 MtuARC does not show any unfolding activity and seems to not recognize the SsrA-tag.
Figure 39: ARC GFP unfolding measurements at 37 °C. ARC measurements (SV= 100 µl, 6.46 µM ARC₆, 1 µM GFP, 1 mM DTT, 2.5 mM ATP, 50 mM Tris pH 7.5, 10 mM MgCl₂)
III Research article

Characterization of a new AAA+ protein from archaea

Abstract

We investigated a new archaeal member of the AAA+ protein family (ATPases associated with various cellular activities) which is found in all methanogenic archaea and the sulphate-reducer Archaeoglobus fulgidus. These proteins cluster to COG1223 predicted to form a subgroup of the AAA+ ATPases. The gene from Archaeoglobus fulgidus codes for a protein of 40 kDa monomeric molecular weight, which we overexpressed in E. coli and purified to homogeneity. The protein forms ring-shaped complexes with a diameter of 125 Å as determined by electron microscopy. Using sedimentation equilibrium analysis we demonstrate that it assembles into hexamers over a wide concentration range both in presence and absence of ATP. As suggested by homology to other members of the AAA+ family, the complex binds and hydrolyzes ATP. Michaelis-Menten analysis revealed a k_{cat} of 118 min\(^{-1}\) and a K_{M} of 1.4 mM at 78 °C. This hyperthermophilic archaeal ATPase is stable to 86 °C and the ATPase activity is maximal at this temperature. The protein is most homologous to the AAA-domain of FtsH from bacteria, while the N-terminal domain shows predicted structural homology to members of the CDC48 family of AAA proteins. Possible roles of this new AAA+ protein are discussed.

Keywords: Archaea, Methanogens, Archaeoglobus fulgidus, AAA+, ATPase, ATA
1. Introduction

The AAA+ ATPases are found in all three kingdoms of life and comprise a family of oligomeric proteins featuring motor-domains (the AAA modules) that drive conformational changes by binding and hydrolyzing ATP (Erdmann et al., 1991; Iyer et al., 2004; Zwickl et al., 1999). The AAA module (AAA stands for ATPases associated with various cellular activities) is made of a classical P-loop NTPase domain followed by a smaller α-helical domain that communicates conformational changes to other regions in the protein complex in response to nucleotide state (Erdmann et al., 1991; Maupin-Furlow et al., 2001; Maurizi and Li, 2001). These nucleotide-hydrolysis-driven conformational changes ultimately lead to a variety of “mechanical” activities depending on the context in which the AAA module is used. Proteins containing AAA-domains are, for example, involved in membrane fusion, protein dissociation, protein unfolding and unwinding of nucleic acids (Lupas and Martin, 2002; Ogura and Wilkinson, 2001). In most of these systems they work in complex with other binding partners. A significant subset of AAA+ proteins is involved in the process of protein quality control (Wickner et al., 1999). These AAA+ ATPases participate in the disaggregation or unfolding of protein substrates and many of them also interact with protease partners to mediate the proteolytic destruction of their substrates (Langer et al., 2001; Maupin-Furlow et al., 2004; Maurizi and Xia, 2004).

Archaeoglobus fulgidus and the methanogenic archaea each possess more than a dozen AAA+ proteins, which are involved in such essential processes as DNA replication (members of the Mcm family, CDC6-like replication initiation factors) and ATP-dependent protein degradation (Lon-protease and proteasome regulatory subunit PAN) (Benaroudj and Goldberg, 2000; Besche et al., 2004; Davey et al., 2002; Groll et al., 2003). The roles of some of the classified AAA+ proteins from archaea are still poorly understood, like for example the role of the CDC48-like AAA+ proteins that display surprisingly high homology to the eukaryal members important for dislocation of proteins from the ER (Wang et al., 2004).

In order to further investigate the characteristics of a group of putative AAA+ proteins, which form a separate “cluster of orthologous groups”, COG 1223, we have characterized the protein from *A. fulgidus*. So far no member of COG 1223 has been investigated, but they have been predicted to be ATPases of the AAA+ superfamily (Maupin-Furlow et al., 2000). Amongst the 25 fully sequenced archaeal genomes, all methanogens and the strictly anaerobe sulphate-reducer *A. fulgidus* contain this AAA+ gene which we will herein refer to as “archaeal triple-A ATPase” (ata). Through a combination of structural, biophysical and kinetic experiments we show that ATA is a hexameric protein with peak ATPase activity at 86 °C.
2. Materials and Methods

2.1. Materials

CIAP, T4 DNA Ligase, the restriction enzymes NdeI and BamHI were from MBI Fermentas. PfuTurbo DNA polymerase was purchased from Stratagene. The plasmid pET-11a and E.coli BL21 (DE3) Rosetta™ cells were from Novagen. The reagents for the malachite green assay were from Sigma (malachite green), Fluka (sodium citrate and hexammonium molybdate) and Calbiochem (Tween-20). ATP and ATPγS were purchased from Roche Molecular Biochemicals. Prepacked columns, chromatography media and the high molecular weight calibration kit were obtained from Amershamb Biosciences.

2.2. Construction of the ATA-overexpressing plasmid

The gene AF1285 was amplified from genomic DNA (kindly provided to us by Karl Stetter, Archaeenzenrum Universitaet Regensburg) using PCR. The following primers were used:

Sense primer 5'-GTGCACATATGGCAAAGAGGGAAACGGC-3' and
antisense primer: 5'-CACCAGGATCCCTAAACGAACATTTGGGG-3'. After cleavage with NdeI and BamHI, the PCR product was ligated into the NdeI- and BamHI-digested pET-11a plasmid. This ligation resulted in a vector with the ata gene under control of the T7-promoter (pET-11a/ata). The sequence of the cloned gene was verified by DNA sequencing.

2.3. Expression and purification

ATA was purified from an IPTG-induced culture of E. coli BL21 (DE3) Rosetta™ cells harboring the pET-11a/ata plasmid. Cells were resuspended in buffer A (0.05 M HEPES pH 7.5 (RT); 0.03 M NaCl; 10% (v/v) glycerol; 2 mM DTT) and disrupted in an EmulsiFlex-C5 High Pressure Homogeniser at 105 kPa. The cleared lysate was heated to 60 °C for 30 minutes and then centrifuged again. The lysate was then passed over a Sepharose FastFlow Q anion exchange column equilibrated in buffer A. ATA eluted at 0.3 M NaCl in a linear gradient of 0.03 – 0.7 M NaCl. The pooled fractions were precipitated with 50% saturated ammonium sulfate solution. The pellet was resuspended in 30 ml buffer B (0.05 M HEPES pH 7.3 (RT); 0.03 M NaCl; 10% (v/v) glycerol; 2 mM DTT), dialyzed against the same buffer and applied on a Source 30 Q anion exchange column. ATA eluted at 0.27 M NaCl in a linear gradient of 0.03-1 M NaCl. The pooled fractions were dialyzed against buffer C (0.01 M HEPES pH 7.8 (RT); 0.03 M NaCl; 10% (v/v) glycerol; 2 mM DTT) and applied on a
MonoQ anion exchange column. The protein was eluted from the column at 0.34 M NaCl in a linear gradient of 0.03-0.7 M NaCl. The pooled fractions were dialyzed against buffer R (0.05 M HEPES pH: 7.5 (RT); 0.3 M NaCl; 0.02 M MgCl$_2$; 10% (v/v) glycerol) and run over a Superose 12 size exclusion column for further purification. The protein was then flash-frozen and stored at –80 °C.

2.4. ATPase activity assay

The ATPase activity of ATA was measured by the malachite green assay (Lanzetta et al., 1979). Calibration was carried out by using an aqueous solution of KH$_2$PO$_4$. Quenching of color development was obtained by addition of 34% citrate. For the activity/temperature profile, the concentration of ATA was 30 µM (monomeric) and the ATP concentration was 5 mM. The Michaelis-Menten analysis was carried out at 78 °C in buffer R. The concentration of ATA was 12 µM (monomeric) and the ATP concentrations varied from 200 µM to 20 mM. The obtained data were analyzed using the following equation:

\[ V_o = \frac{V_{\text{max}} \cdot [\text{ATP}]}{K_M + [\text{ATP}]} \]  

Where \( V_o \) is the initial velocity and \( V_{\text{max}} \) is the velocity at substrate saturation.

2.5 Circular dichroism spectra

Far-UV CD spectra were recorded on a JASCO J715 CD spectropolarimeter. Protein concentration was 6 µM ATA hexamer. The spectra were measured at 25 °C in 0.01 M HEPES, pH 7.5 (RT), 0.3 M NaCl and 10% glycerol in a 0.1 cm quartz cuvette. The recorded spectra were averages of 20 measurements and were corrected for buffer background. The temperature transition was recorded from 25 °C to 95 °C at 1 °C/min in buffer R in a 0.1 cm quartz cuvette. The protein concentration was 5 µM hexamer. The temperature dependent unfolding curve was fitted as described in Pace et al. (1998). As the process of unfolding was not reversible for the ATA complex, because it precipitated after heating to 95 °C, the enthalpy values are not meaningful and only the inflexion point is reported.
2.6. MALDI-MS analysis and Edman degradation

Mass spectra were recorded using a Voyager Elite MALDI-TOF mass spectrometer (Perspective Biosystems). Edman degradation reactions were carried out and analyzed on an Applied Biosystems Procise Sequencer.

2.7. Analytical size exclusion

Analytical size exclusion experiments were performed on an analytical (24 ml) Superose 6 column in buffer R in presence or absence of ATP. The flow rate was 0.2 ml/min. For calibration the Gel Filtration HMW Calibration Kit containing blue dextran 2’000 (~2’000’000 Da), thyroglobulin (669’000 Da) and catalase (232’000 Da) was used. 300 µg of each were loaded in a 100 µl volume. 700 µg of ATA were applied in a 100 µl volume. The results were analyzed by plotting $K_{av}$ versus log (Mr).

$$K_{av} = \frac{(V_e - V_o)}{(V_t - V_o)}$$

where $V_e$ is the elution volume, $V_o$ is the void volume and $V_t$ is the column volume. For the gel filtration runs in presence of nucleotide, an ATP concentration of 1 mM was used in the applied sample and the running buffer.

2.8. Analytical ultracentrifugation

Prior to ultracentrifugation experiments, proteins were dialyzed against two changes of buffer R. Densities were measured by pycnometry at 4 °C with a Model 48 densitometer (Anton Paar). The measured density of buffer R was 1.04055 g/ml. The partial specific volume of the protein was calculated from amino acid composition. Sedimentation velocity and sedimentation equilibrium experiments were performed using an XL-A analytical ultracentrifuge from Beckman. Analytical cells had 12 mm, 4°, Epon carbon-filled double sector centerpieces and plane, far-UV quartz windows. Loading volumes for velocity runs were 0.34 ml of protein solution (26 µM monomer) and 0.35 ml of the appropriate buffer R. Loading volumes for equilibration runs were 0.13 ml protein solution and 0.15 ml appropriate buffer. The protein concentration for the equilibration runs ranged from 0.02 – 0.5 mg/ml. All experiments were carried out at 4 °C. The XL-A was run at 3’000 rpm until the temperature was 4 °C and the vacuum established. When the rotor accelerated to the desired speed, scans were taken every three minutes in continuous mode for velocity runs and every 180 min in step mode for equilibration runs. The step size was always 0.001 cm. Equilibrium was reached when the subtraction of absorbance spectra, with six hours time difference, resulted in a randomly distributed deviation below 0.02 AU over the whole exponential part. Data were analyzed using the Beckman XL-A/XL-I Data Analysis software, UltraScan and sedfit (Demeler and Saber, 1998; Schuck, 2004).
The molecular weight was determined by fitting the absorbance data of the equilibration runs using an exponential equation:

\[
C(r) = C(r_0) \cdot e^{-\frac{(1-eomega^2)}{2RF}M(r^2 - r_0^2)}
\]

(3)

\(C(r)\): concentration at radius \(r\)
\(C(r_0)\): concentration at reference radius \(r_0\)
\(M\): monomer molecular weight
\(R\): universal gas constant
\(T\): temperature
\(omega\): angular velocity

2.9. Sequence alignments

The multiple sequence alignment of ATA, PAN and FtsH was carried out with ClustalW (Chenna et al., 2003) using the BLOSUM score matrix (Henikoff and Henikoff, 1992). The sequence alignment with multiple ATA sequences was carried out with T-COFFEE (Notredame et al., 2000).

2.10. Electron microscopy

A drop of the ATA protein solution at a concentration of 126 µg/ml containing 100 µM ATPγS was placed on a glow-discharged, carbon-coated grid 400 mesh for 30 seconds. The sample drop was blotted away and the grid was negatively stained with one drop of 1% aqueous uranyl acetate blotted immediately away after application. The samples were viewed in a Philips EM420 transmission electron microscope. Nominal magnification was 31’500. After a reference-free alignment procedure image classification was performed using the Imagic-5 software (van Heel et al., 1996).

2.11. Phylogenetic analysis

For the generation of the phylogenetic tree, archaeal AAA+ proteins from seven representative organisms (Archaeoglobus fulgidus, Methanocaldococcus jannaschii, Pyrococcus abyssi, Halobacterium salinarum, Thermoplasma acidophilum, Aeropyrum pernix, Sulfolobus solfataricus) and nine different clusters of orthologous groups (COG0464, COG0470, COG0714, COG1067, COG1222, COG1223, COG1224, COG1241, COG1474) (Tatusov et al., 2000; Tatusov et al., 2001) were used. Pairwise alignment was carried out using ClustalW applying a Gap open penalty of 10, a Gap
extension penalty of 0.1 and the BLOSUM weight matrix. The phylogenetic tree was drawn using Phylodendron (D.G. Gilbert, version 0.8d).

3. Results

3.1. Sequence analysis and similarities with other AAA+ proteins

The ata genes of different organisms encode proteins of 334 amino acids (Methanopyrus kandleri) to 372 amino acids (Methanosarcina acetivorans) in length. The ata gene of A. fulgidus is composed of 1059 base pairs and its locus in the genome is from base pair 1’144’771 to 1’145’829. It codes for a protein (AF1285) of 352 amino acids in length giving a calculated monomeric molecular weight of 39.9 kDa. The genomic context of the ata gene does not reveal any adjacent functionally related genes nor does it show co-occurrence with other genes that might provide clues to ATA’s functional role.

Protein sequence comparison was carried out in three different contexts: with all fully sequenced archaeal genomes (26), with all fully sequenced bacterial genomes (317), and with both archaeal and bacterial genomes simultaneously using the “basic local alignment search tool” (BLAST) (Altschul et al., 1997). As previously shown by exhaustive phylogenetic analysis of the AAA protein family (Frickey and Lupas, 2004), BLAST against the bacterial genomes revealed the FtsH ATPase domain as the most homologous protein sequence, while comparison with the archaeal genomes revealed AAA domains of CDC48-like ATPases as the closest relatives. The next closest relative is the proteasome-activating nucleotidase (PAN), the proteasome binding partner in archaea (Zwickl et al., 1999). In all cases, the sequence alignment excluded an N-terminal region of about 102 amino acids probably representing the nonconserved N-domain of A. fulgidus ATA. Pairwise alignment of ATA with the FtsH protein sequence from E. coli showed 36% sequence identity and 53% homology. Alignment of ATA with the proteasome-activating nucleotidase (PAN) from A. fulgidus revealed 33% sequence identity and 51% homology. Figure 1 A shows the multiple sequence alignment of the AAA modules of ATA, FtsH and PAN.
Figure 1: Sequence alignments. A) Alignment of the ATA protein from *A. fulgidus* with the protein sequences of FtsH from *E. coli* and PAN from *A. fulgidus*. (B) Multiple alignment of the ATA protein sequences from nine different archaea. The arrows above the sequence of *A. fulgidus* ATA denote the putative domain structure and the ATP-binding signature motifs. The N-domain is shown in blue, the AAA module is shown in green (P-loop domain) and light blue (α-helical domain). In both alignments, conserved residues are highlighted in red, equivalent residues are highlighted in yellow. The alignments were colored by ESPript (Gouet et al., 1999).
Multiple sequence alignment with ATA protein sequences from nine archaeal species using T-COFFEE showed that aside from an N-terminal region of about 100 amino acids, the different ATA proteins show very high sequence homology (Figure 1B). The conservation is highest around the ATP-binding signature motifs, the Walker A and Walker B domains.

3.2. Expression and purification of recombinant ATA protein

ATA was purified from *E. coli* BL21(DE3) Rosetta™ harboring the pET-11a/ata plasmid. Rosetta cells were used because the ata gene contains multiple rare codons with regard to *E. coli* codon usage. Expression tests showed highest expression levels in Rosetta cells grown at 37 °C. As a first crude purification step the supernatant after cracking of the cells was heated to 60 °C for 30 minutes. The cleared lysate was passed over a Sepharose FastFlow Q anion exchange column (ATA eluted at 0.3 M NaCl). The pooled fractions were precipitated with 50% saturated ammonium sulfate solution. The pellet was resuspended and subjected to two additional anion exchange chromatography steps at slightly different pH values using a Source 30 Q anion exchange column and a MonoQ anion exchange column. The protein solution was further separated by size exclusion chromatography using a Superose 12 column. The protein yield after the purification was 11.0 mg/L cell culture. The protein could be flash-frozen and stored at –80 °C without any loss of ATPase activity. The identity was verified by Edman-sequencing and MALDI-Mass spectrometry.

3.3. ATA protein forms ring-shaped complexes

To investigate the homogeneity and the assembly state of the ATA protein the sample was examined by negative stain electron microscopy. The recorded images show uniformly shaped and sized particles that are evenly distributed across the fields. The individual particles are clearly ring-shaped, displaying mostly top views with a discernable hole in the middle where the stain accumulates. The diameter of the rings is in the range of 120-130 Å. Images were recorded both in the absence and presence of ATPγS, as the assembly state of AAA+ proteins is often influenced by binding of nucleotide. Under both conditions, the same number of particles was observed and they displayed the same overall shape. Micrographs of evenly stained fields of ATA complexes in presence of ATPγS were digitized and 2000 individual particles were manually picked and grouped into classes based on similarity using Imagic-5 (Image Science). A large number of these groups show a six-lobed ring structure. Figure 2 shows the negatively stained electron micrograph and the inset is a representative view of the 150 class-averages of the hexameric ring obtained from 18 individual particles.
3.4. Analysis of the assembly state by size exclusion chromatography

A first indication of higher oligomerization of ATA into complexes originally came from the preparative size exclusion chromatography step used in the purification procedure. Therefore, a calibration on an analytical Superose 6 column was carried out using the Gel Filtration HMW Calibration Kit containing blue dextran 2000 (~2,000,000 Da), thyroglobulin (669,000 Da) and catalase (232,000 Da). The ATA protein sample was applied to the same Superose 6 column in reference to the standards and eluted after 12.9 ml (Figure 3 A). The determined molecular weight of the ATA complex resulting from the calibration is 494 kDa (see “Materials and methods”) (Figure 3 B). Surprisingly, this corresponds to an assembly state of 12 subunits per complex even though members of the AAA-protein family are known to form hexamers. The effect of ATP on the oligomerization state was analyzed by applying ATA to the Superose 6 column in the presence of 1 mM ATP. Under these conditions the ATA complex also eluted at 12.9 ml (data not shown).
3.5. **Analytical ultracentrifugation measurements**

To further analyze the oligomeric assembly state analytical ultracentrifugation measurements were carried out. First, to test the homogeneity and stability of the ATA complex and to obtain the sedimentation coefficient for the particle, sedimentation velocity runs were performed at 40'000 rpm in a Beckman XL-A. A sedimentation coefficient of $8.28 \times 10^{-13}$ s was determined from these measurements after correction to standard conditions ($S_{20,w}$). This represents the main sedimentation coefficient peak at $s^* = 4.7 \cdot 10^{-13}$ s with 91% of the total protein concentration (the sedimentation coefficient distribution peak ranges from 4–5 S). A smaller peak at $s^* = 7.4$ S ($S_{20,w} = 13.0 \cdot 10^{-13}$ s) appeared on all sedimentation velocity experiments accounting for 9% of the total concentration presumably representing aggregates.
Sedimentation equilibrium ultracentrifugation of the ATA complex at 9’000 rpm and 4 °C. (A) The different exponential absorbance curves belong to different protein concentrations between 0.02 and 0.5 mg/ml (see Table 1). The global fits carried out with Ultrascan are shown in red, the measured data points are shown as black squares (B) Residuals are plotted as black squares. Deviations of the measured curves to the calculated show an even random distribution under 0.02 AU indicating a good fit.

Table 1
Apparent molecular weights from sedimentation equilibrium measurements

<table>
<thead>
<tr>
<th>ATA concentration</th>
<th>Apparent molecular weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5 mg/ml</td>
<td>243.8 kDa</td>
</tr>
<tr>
<td>0.3 mg/ml</td>
<td>231.2 kDa</td>
</tr>
<tr>
<td>0.2 mg/ml</td>
<td>225.7 kDa</td>
</tr>
<tr>
<td>0.1 mg/ml</td>
<td>220.1 kDa</td>
</tr>
<tr>
<td>0.05 mg/ml</td>
<td>229.1 kDa</td>
</tr>
<tr>
<td>0.02 mg/ml</td>
<td>235.4 kDa</td>
</tr>
</tbody>
</table>

Sedimentation equilibria were then measured over a wide concentration range of ATA to detect possible lower oligomers (Table 1). The method chosen for the runs was meniscus depletion for $S_{20,w} = 8.28 \cdot 10^{-13}$ s. Therefore, the speed was set to 9’000 rpm and the temperature was set to 4 °C. Equilibrium was reached for all cells after 139 hours. A global fit for all spectra obtained from this sedimentation equilibrium run was done with UltraScan (Figure 4). The single species model for the global fit
produced the best result with a molecular weight of 236.8 kDa which corresponds to the ATA hexamer.

3.6. Characterization of the ATPase activity of the ATA complex

ATA contains a nucleotide binding domain and is very homologous to a certain class of chaperones that associate with protease partners and display ATPase activity. Therefore, the ATPase activity was measured using the malachite green assay (Lanzetta et al., 1979). This ATPase assay is based on formation of a complex between inorganic phosphate with ammonium molybdate and malachite green and can be carried out over a wide range of temperatures. First the temperature-dependence of the ATPase activity was examined. The activity of the ATA complex was measured from 25 to 90 °C at eight different temperatures. The data shown in Figure 5 have been corrected for spontaneous ATP hydrolysis. The ATA complex shows highest ATPase activity at 86 °C and at 90 °C it drops sharply suggesting loss of structure at that temperature. To confirm this, temperature denaturation curves for ATA were recorded to measure the transition from the folded to the unfolded state using the circular dichroism signal at 222 nm. ATA was heated from 25 °C to 95 °C at 1 °C/min. The inflexion point of the transition shows the state where half of the protein is native and the other half unfolded. This occurs at 87.3 ±0.05 °C (Figure 5 A, second y-axis). At the end point (95 °C) most of the protein is irreversibly aggregated rendering the transition irreversible. At 86 °C most of the protein is still folded.

Full Michaelis-Menten analysis was carried out at 78 °C with ATP concentration varying between 200 µM to 20 mM. The concentration of ATA was kept constant at 2 µM hexamer (Figure 5 B).

Figure 5: ATPase activity
(A) Temperature dependence of the ATPase activity of the ATA complex (open circles). The protein concentration was 5 µM (hexameric complex), the ATP concentration was 5 mM. Highest ATPase activity was measured at 86 °C. On the second Y-axis, the temperature transition profile of ATA measured by circular dichroism at 222 nm is shown (filled squares; fit shown as solid black line). The inflexion point of this irreversible transition is 87.3 ±0.05 °C (for details of the fit see “Materials and
Methods”). (B) Michaelis-Menten kinetic of the ATA complex at 78 °C. ATP concentration varied from 0.1 to 20 mM. The concentration of ATA was 2 µM (hexameric).

A temperature slightly below the optimal growth temperature of *A. fulgidus* was chosen to ensure that the ATA complex remained stable over the course of the measurement. From this analysis the catalytic constant for ATP hydrolysis at 78 °C was determined to be 118 ± 4 min⁻¹ (per hexamer) and the apparent dissociation constant *Kₐ* was determined as 1.4 ± 0.2 mM.

4. Discussion

The strict anaerobe *A. fulgidus* and methanogenic archaea contain a putative AAA+ ATPase with a single AAA module and a calculated molecular weight of about 40 kDa. Searching all fully sequenced archaeal genomes, eight orthologs of ATA could be identified that form a cluster of orthologous groups (COG1223). No bacterial members of this COG could be identified and, therefore, the protein was named “*Archaeal Triple-A ATPase*” (ATA). Phylogenetic analysis of archaeal AAA+ proteins from nine different COGs using seven different archaeal species places ATA on the same branch together with COG1222 (proteasome-activating nucleotidase) and COG0464 (CDC48-related proteins, VAT). The ATA COG1223, however, branches off early and does not cluster directly with either one of the other COGs (Figure 6).

This is in agreement with a more exhaustive phylogenetic analysis on the whole AAA protein family which describes this group as monophyletic occurring either within the metalloprotease clade or close to the root of the tree (Frickey and Lupas, 2004). ATA’s closest homologs in archaea and bacteria are thus part of the protein quality control system (FtsH, CDC48-like, PAN) and are more specifically involved in protein removal and degradation.

As a sequence homologue of the AAA+ family members ATA is a predicted ATPase that displays the conserved nucleotide binding site with the Walker A and Walker B signature motifs. The arrows in Figure 1B show the conserved features and the putative domain structure of the ATA proteins.
Almost all members of the AAA+ protein family assemble into ring-like higher order structures. The overall topology of ATA, as visualized by negative stain electron microscopy, was clearly shown to be ring-shaped irrespective of the presence or absence of ATPγS. As is characteristic for AAA complexes, the particles displayed predominantly as top-views resulting in ring-like projections with an outer diameter of about 125 Å and dark, stain-accumulating centers. The ring-diameter observed in the electron microscopic images agrees well with how VAT from T. acidophilum (Pamnani et al., 1997), p97 from Xenopus laevis (Peters et al., 1992) or CDC48 from yeast present in similar experiments (Frohlich et al., 1995).

Most of the ring-shaped complexes formed by AAA+ proteins are hexameric and often the assembly state is dependent on the binding of nucleotide (Lupas and...
Martin, 2002). We investigated the assembly state of ATA using both size exclusion chromatography and sedimentation equilibrium analysis by analytical ultracentrifugation. The first technique is based on calibration with standard proteins to get a relative estimate of the molecular weight and resulted in a molecular weight corresponding to a dodecameric assembly state. Sedimentation equilibrium analysis is a primary method and, as such, does not rely on calibration with molecular weight standards. Analysis with sedimentation equilibrium measurements demonstrated clearly that ATA exists as a hexamer under a wide range of concentrations and that the result obtained with gel filtration must have been influenced by factors other than molecular weight. The hydrodynamic radius of ring-shaped complexes is certainly larger than that of a globular assembly of the same molecular weight. Furthermore, unspecific interaction with the column matrix could also contribute to a shift in the elution profile toward higher apparent mass. To analyze the possible existence of lower oligomerization states of ATA, such as dimers or trimers, equilibria were measured at different concentrations of ATA, covering the widest possible range of concentration for these measurements resulting in a 25-fold difference between the lowest and the highest concentration. The boundaries for the measurable concentrations are given by the sensitivity of the optical system of the analytical ultracentrifuge and non-ideality at protein concentrations over 0.5 mg/ml. Under these conditions no shift of the apparent molecular weight towards lower masses with decreasing concentration could be detected. This means that the hexameric species is predominant and lower molecular species are below detectable quantities. As is the case for many AAA+ proteins with only one AAA-domain, ATA is already fully assembled in absence of the nucleotide and addition of ATP does not change its oligomeric state.

*Archaeoglobus fulgidus* is a hyperthermophilic archaea and grows between 60 and 95 °C with optimal growth temperature at 83 °C. Thermal unfolding measurements showed that ATA was stable up to 86 °C with an inflexion point in the temperature transition of 87.3 °C. Not surprising for a hyperthermophilic organism, no ATPase activity could be detected at room temperature. The ATPase activity rises with higher temperature to an optimum close to the optimal growth temperature of this organism. Michaelis-Menten analysis at 78 °C showed a $k_{cat}$ (118 min$^{-1}$ per hexamer) and $K_m$ (1.4 mM) value similar to other AAA+ proteins (Zwickl et al., 1999).

The N-terminal domains of AAA+ proteins have been shown to play an important role in the substrate and functional specificity of the various AAA+ ring complexes (Dougan et al., 2003; Gerega et al., 2005; Guo et al., 2002). The N-terminal domains function both as substrate binding regions and as interaction platforms for so-called adaptor proteins (Dougan et al., 2002). Alignment with the closest homologues that themselves do not belong to COG1223 shows sequence similarity starting at around residue 100 of ATA. In spite of the low level of homology in the N terminal region of the nine members of COG1223 (Figure 1B), it was possible to identify and align their N-terminal domains using the multiple sequence alignment algorithm implemented in T-COFFEE (Notredame et al., 2000). The resulting alignment file was subjected to
the FUGUE sequence-structure comparison program to identify potential structural homologues and alignments (Shi et al., 2001). In agreement with the finding by Frickey and Lupas (2004), the N-terminal domain of ATA is predicted to adopt a \( \beta \)-clam fold similar to the C-terminal subdomain of the VAT N-domain (Coles et al., 1999). This suggests that ATA might be involved in similar interactions as the CDC48-like AAA members. One of the prominent roles of eukaryotic CDC48 is in the ER-associated degradation of proteins by the proteasome (Elsasser and Finley, 2005; Lord et al., 2002). Here the CDC48 homologue is responsible for dislocation of ubiquitinated substrate proteins from the ER membrane (Ye et al., 2001). As archaea do not contain ubiquitin sequences in their genome, the mode of interaction between archaeal CDC48 homologues and their substrates and other binding partners is still very unclear. Furthermore, nothing is known about possible adaptor proteins for archaeal CDC48-like proteins. However, one common feature with their eukaryotic counterparts could be that CDC48 homologues in archaea also act near membranes and on membrane-associated substrates. Similarly, ATA could be involved in extraction of proteins from the membrane. The other close sequence-homologue of ATA in bacteria, the membrane-bound FtsH, has a very similar function in the removal of membrane-proteins. Its AAA-domain “pulls” proteins out of the membrane for subsequent degradation by its metalloprotease domain (Kihara et al., 1999).

Identification of interaction partners for this AAA complex could provide important clues about its role inside the cell. We are now in the process of characterizing a mesophilic member of this COG to search for binding partners using archaeal whole lysate.

Acknowledgement

Special thanks go to Daniel Boehringer for his help with the class averaging of electron micrographs. Electron microscopy was performed at the Electron Microscopy Center of the ETH Zürich (EMEZ). We are grateful to Heinz Gross and the scientific and technical staff of EMEZ whose efforts have made these experiments possible. This work was supported by the Swiss National Science Foundation.
References


IV Ongoing work with AfuATA

1. Crystallization trials with AfuATA

Crystallization experiments were carried out with ATA and His-ATA using the sitting drop method. The protein was dissolved in 10 mM HEPES/NaOH, pH: 7.5, 10 mM NaCl, 10% glycerol (v/v) at a concentration of 7.8 mg/ml. Needle clusters were obtained after four days at +20 °C (Figure 40) from two conditions: from 1.8 M lithium sulfate and 100 mM sodium cacodylate pH 6.5 at +20 °C (conditions from Clear Strategy Screen™ by Molecular Dimensions Ltd) and from 2 M ammonium sulfate, 5% 2-propanol (Fig. 40). The drops contained two volumes of protein and one volume of reservoir with a total drop volume of 300 nl equilibrated against 450 µl stock reservoir (scaled up later to 6 µl drops equilibrated against 750 µl reservoir).

Figure 40: Needle clusters from 100 mM sodium cacodylate at 20 °C. Protein concentration was 10 mM HEPES, pH: 7.5, 10 mM NaCl, 10% glycerol at a concentration of 7.8 mg/ml. Needle cluster from the screen set up with the robot using 0.3 µl drops.

2. Test for possible interaction with the proteasomal core particle

Because of the high sequence homology between PAN and ATA we postulated that ATA could be an alternative binding partner to PAN for the proteasome. To test a possible interaction several experiments were set up. In a first trial, analytical size exclusion chromatography was done in presence of ATP, but no stable interaction could be detected using this method (data not shown). Other trials were performed, with different nucleotidate concentrations and incubation temperatures and also with ATP instead of ATPγS, all without success.
3. Identification of binding partners of the AfuATA complex

Identification of binding partners for the ATA complex could provide important information about the function and in vivo activity of this AAA+ ATPase. Proteins interacting with ATA could either be substrates or adaptor proteins modulating ATA’s activity. Two types of experiments were carried out to identify interacting proteins both based on an approach that uses AfuATA as bait and A. fulgidus whole cell lysate as prey.

In the first experiment purified His-ATA was incubated with whole cell lysate of A. fulgidus cell pellets in presence of ATPγS at temperatures above 55 °C. After incubation the mixture was loaded onto a Ni-chelating column. After a washing step with imidazole-containing buffer, His-ATA was eluted from the column at a higher concentration of imidazole (for details see methods). Interaction partners binding to AfuATA in a stable manner should co-elute with the His-ATA complex. The collected fractions were concentrated and visualized via SDS-PAGE (Fig. 41). Bands of interest were blotted and sequenced. The flow through (FT) and the elution fractions 1-4 were applied on the gel, as well as a sample of the recombinant ATA and the concentrated fractions.

![Figure 41: Collected fractions binding to the Ni-chelating assay visualized by 15% reducing SDS-PAGE.](image)

Figure 41 shows that by this method no bands were visible that did not originate from the purified ATA sample itself. Other trials were performed, with different nucleotide concentrations and incubation temperatures and also with ATP instead of ATPγS. But also under those conditions, binding partners could not be found.

In a second experiment an anti-His antibody (AB) was used to co-immunoprecipitate His-ATA and potential interaction partners. Again His-ATA was used as bait and
was incubated with *Archaeoglobus fulgidus* whole cell lysate. Different incubation conditions were tested like various temperatures, incubation times and other nucleotides or nucleotide analogs. The principle of the used kit is based on the covalent coupling of the antibody to the provided amine reactive gel. The whole procedure is performed in a spin cup column. The binding with the protein of interest follows the antibody coupling. In the experiment there are two possibilities to proceed. The bait and prey can be incubated and then bound to the antibody or bait, prey and antibody can be incubated as a whole mixture. Washing and elution is done with provided buffers.

![15% reducing SDS-PAGE with eluted fractions from Co-IP](image)

Figure 42: 15% reducing SDS-PAGE with collected fractions from Co-IP.

The samples from washing were precipitated and solubilised in a smaller volume to increase the concentration and loaded on a 15% reducing SDS-PAGE gel (Fig. 42). The fractions from elution were put directly, without further concentration, on the gel.

The band of interest gave *Archaeoglobus fulgidus* CDC48 as result from in-gel digest. This protein is highly expressed in archaea and despite thorough washing there is still the possibility that this result is due to the abundance of CDC48. The remaining bands were too low concentrated for further analysis or were not good enough separated.

For the reason of limited methods due to the high temperatures, which may be important to support a stable interaction between binding partners, a mesophilic representative waits to be tested. Therefore MacATA and His-MacATA were cloned as well as *M. acetivorans* cells were cultivated to get cell material for further lysate testing and to hopefully find a possible interaction partner or substrate for ATA.
Prospects

The work presented in this thesis concerns three classes of AAA chaperones, the archaeal PAN complex and the actinobacterial ARC ring, likely both involved in mediating degradation of protein substrates by the proteasome, as well as an archaeal AAA ATPase with high structural homology to FtsH and p97.

I have cloned and expressed the AAA and proteasomal subunit genes from four organisms, the archaeal *Archaeoglobus fulgidus*, *Pyrococcus furiosus* and *Methanosarcina acetivorans*, as well as the gene for an actinobacterial AAA ATPase from *Mycobacterium tuberculosis*. One problem in purifying archaeal proteins out of *E. coli* lysate is caused by low expression yields. The fact that we are working with huge complexes may result in accumulation of desired protein in inclusion bodies and disturbance of normal cellular function.

In case of the *M. acetivorans* proteins, the main problem with expression is that the *Methanosarcina* genome has a lower GC content than *E. coli* and this codon usage can cause troubles while expressing recombinant proteins, even when using *E. coli* strains containing a plasmid encoding for rare codons. The co-expression of archaeal proteasomes is insufficient due to the fact that with advancing purification steps the complex gets disassembled resulting in separated α- and β-rings. The advantage of recombinant hyperthermophilic proteins is the possibility of including a heating step which has a large purification effect before further chromatographic purification procedures.

As in this thesis only one PAN (MA 4268) was cloned, it is still worthwhile cloning the second PAN gene (MA 4123) and expression of both, to test if there is a stable interaction with the proteasome. There is also a predicted alternative start in PAN (MA 4268), which could also be tested. Another experiment would be to test, if there are heterohexamers consisting of more than only one PAN homolog existing in *M. acetivorans*, which could be tested by co-expression of these genes in *E. coli* and proved via SDS-PAGE, Mass Spectroscopy and Edman sequencing. In addition the proteins can be tagged and pull down assays or Co-IPs can be done from *Methanosarcina acetivorans* whole cell lysate as this methanogenic organism could be cultivated successfully in our lab in a good enough yield for these testings.

From analytical gel filtration it was shown that the AAA+ proteins form higher oligomers. Since gel filtration is not an accurate method for MW determination the functional assembly state could not be determined. Medalia et. al. reported that the *M. mazei* PAN assembles as a dodecamer. This is the first time that an assembly state of 12 subunits per complex was shown as seen in EM.

In order to form a functional chaperone proteasome complex the proteasome core particle has to be assembled first. This process starts with the formation of αβ-half-proteasomes followed by auto catalytic cleavage of the propeptide when two half-proteasomes stack together forming active proteasome core particle.
Compared to archaeal unprocessed proteasome β-subunits, the propeptide of the β-subunits in the actinomycetes are longer and this indicates why it takes longer to process them (Striebel, personal communication). Also the suggested assembly pathway differs from the archaegal (see Introduction) and supports this speculation. Stable complex formation of the core particles with their respective AAA partners was in all cases unsuccessful under the conditions tested so far, even for those complexes that did show proof of transient interaction in the form of stimulated degradation of the model substrate protein GFP-SsrA. The GFP degradation was detectable for the PAN-proteasome complex. One of the greatest challenges was to identify substrate proteins for the archaenal and actinobacterial chaperone-proteasomes. So far no substrate proteins of these large degradation machines have been identified in these organisms. It is very well possible that presence of a true substrate protein might enhance the interaction between the ATPase and the proteasome core particle. Very recently a publication showed interaction between PAN and the proteasome from a mesophile source (Medalia et. al. 2006). In another publication the complex formation was shown using electron microscopy (Smith et. al. 2005). In this set-up the authors used Thermoplasma acidophilum proteasome and Methanocaldococcus jannaschii PAN. The fact that they used this hybrid complex reflects the difficulty in handling these complexes for EM. I have shown that only PAN was able to recognize the SsrA-tagged substrates. Similar results were also shown in the literature (Medalia et al., 2006). It is furthermore known that AAA+ proteins often interact with adaptor proteins that regulate their specificity and activity. Such binding partners could also play a role in stabilizing interactions between the core particle and the ATPase or between the ATPase ring and substrate proteins.

A potential alternative interaction partner for the proteasome in methanogenic archaea and in the sulphate reducer A. fulgidus, termed ATA, has been investigated. Interaction with the core particle could not be established by either gel filtration or electron microscopy. Phylogenetic analysis of archaeal AAA+ proteins from nine different COGs using seven different archaeal species places ATA on the same branch together not only with the proteasome-interacting PAN proteins but also with CDC48-related proteins.

One of the prominent roles of eukaryotic CDC48 is the involvement in the ER-associated degradation of proteins by the proteasome (Elsasser et al., 2005; Lord et al., 2002). Here the CDC48 homologue is responsible for dislocation of ubiquitinated substrate proteins from the ER membrane (Ye et al., 2001) without proof of any direct interaction between the proteasome core particle and CDC48. In an analogous way ATA could mediate degradation of certain protein substrates without itself forming a stable complex with the proteasome core.

One thing to do in future experiments would be to fish for substrates from whole cell lysate, as was done for AfuATA. In the presented set-ups with His-ATA as bait and A. fulgidus whole cell lysate as prey, no result for a binding partner could be obtained. One difficulty of pull down assays or Co-IP is that the interaction has to be strong...
enough to get efficient retrieval of proteins. Additionally the substrate turnover might be to fast for this method. Another method to show interaction would be to use crosslinking agents for protein-protein interactions study. Finding a target protein would lead to interesting new experimental possibilities and could lead to a better functional understanding of these proteins.

The pull down assay with AfuATA applied on the *Archaeoglobus fulgidus* cell lysate revealed the binding of sigma32 factor and GrpE. The Co-IP revealed one interesting band which was identified as CDC48 by sequencing. It has to be considered that CDC48 is a highly expressed house keeping protein in archaea; it could therefore well be that CDC48 co-elution is due to its high concentration within the whole cell lysate. It would be an interesting assignment to test if there is an interaction between ATA and CDC48 in vitro. The function of ATA in the cell is still unknown. ATA is an outstanding AAA+ protein because of its special N-domain. From structure modeling, the N-domain aligned perfectly well in the hexamer structure of *M. musculus* p97. CDC48 and p97 have an N-terminal double-psi barrel subdomain and were shown to be crucial for cell cycle control or membrane fusion. It is very doubtful that ATA mainly act as a chaperone as shown by Djuranovic et al. 2006, because it is structurally related to CDC48, which mainly functions in cell cycle control. Chaperone disaggregation activity was also tested for AfuATA in this thesis, but as these results gave no significant conclusions, they were not presented here. As the production of aggregates is hard to perform in vitro in a homogeneous and reproducible manner, tests to measure chaperone disaggregating activity are difficult to do in a reliable form. For future studies accurate and reproducible chaperone assays should be developed and the experiments could be repeated with the mesophilic ATA from *Methanosarcina*. ATA is only found in methanogenic archaea and the hyperthermophilic *Archaeoglobus fulgidus*. Interestingly *Archaeoglobus* contains at least five genes for methanogenesis but one important enzyme for methane production is missing (methyl Co-M reductase) and therefore only reverse methanogenesis occurs.

Until now, the distinct function of ARC is still unclear and possible functions for MtuARC maybe a role to protect or unfold oxidized proteins, which might be naturally occurring substrates. It was shown by Darwin et. al. 2003, that there is a virulence indication for the proteasome. The presence of the proteasome helps the *Mycobacterium* to develop resistance to macrophage attack and serves as a defense mechanism against oxidative or nitrosative stress. Macrophages produce nitric oxides and other reactive nitrogen intermediates to control infection. ARC is the suggested chaperone component of the bacterial proteasome but until today no results are presented about ARC stimulated proteasomal degradation of substrates in vitro. The future part of this project is to test, if there is an interaction between ARC and the proteasome.
Curriculum Vitae

Name: Heike Summer

Date and place of birth: October 4th 1978 in Feldkirch, Austria

Nationality: Austrian

Education:

2002-2006 Ph.D. studies at the Institute of Molecular Biology and Biophysics, ETH Zuerich

1998-2002 Diploma study in microbiology with elective subjects virology and molecular biology at the Leopold-Franzens-University Innsbruck (A)

2000-2001 Erasmus semester at the Technical University Darmstadt (D)

1997-1998 Study biology at the Karl-Franzens-University Graz (A)

1993-1997 Bundesoberstufenrealgymnasium, Matura natural science branch, Feldkirch (A)

1998-1993 Secondary school, Institute St. Josef in Feldkirch (A)

1985-1998 Elementary school in Feldkirch (A)
References


