ABSTRACT

CHANG, LARA SAMOFAL. Biochemical and biophysical characterization of compartmentalizing proteases from the hyperthermophilic microorganism *Pyrococcus furiosus*. (Professor Robert M. Kelly Committee Chair)

Proteases catalyze the cleavage of peptide bonds in peptides, polypeptides, and proteins using a hydrolysis reaction. From a biological standpoint, these enzymes are critical for cellular survival, particularly in removal of denatured proteins during stress events or of proteins that have completed their functions. Various proteases play distinct roles in the degradation of proteins, including proteinases that break down proteins and peptidases that break down the resulting oligopeptide products to single residues. The hyperthermophilic versions of proteases are useful for several reasons: they are easier to study because of their relative structural simplicity and, compared to their mesophilic counterparts, they are more stable in harsh conditions such as high heat.

The focus of this study was on the biochemical and biophysical characteristics of two multi-subunit compartmentalizing proteases from the hyperthermophilic archaeon *Pyrococcus furiosus* (T_{opt}=100ºC). The first protease was an oligopeptidase, PfpI (*Pyrococcus furiosus* protease I), and the second was a proteinase, called the proteasome. Both proteases are ubiquitous in all domains of life. However, they are theorized to have distinctly different roles within *P. furiosus*. The proteasome may be one of the primary proteinases, with access to its active sites tightly controlled by ATPase regulators that appear

to be dependent on cellular environment. In contrast, the role of PfpI may be degradation of the smaller peptides that result from proteasome and other proteinase action.

PfpI is a homo-multimer of 18.8-kDa subunits that assemble into hexameric rings. These rings then stack to form dodocamers and higher forms, with three active sites buried in hindered positions within each ring. Trimer, hexamer, and dodecamer forms were purified separately, with the dodecamer at least three-fold more specifically active than the smaller forms. It was also found that PfpI was only able to cleave oligopeptides up to 17 residues, preferring aromatic residues at the P_1 position. As the substrate length was increased, the cleavage by PfpI became less specific and confined to the C- and N-termini. The precise role of PfpI in *P. furiosus* still remains to be determined, with a particular need for studies of recombinantly expressed versions.

The 20S proteasome, along with a theorized ATP-dependent regulator PAN (proteasome-activating nucleotidase), was investigated from several angles. Both enzymes, including native and recombinant forms, were tested for biochemical and biophysical characteristics as isolated structures and in combination. In particular, the PAN ATPase activity was tested primarily to observe its effects on different forms of the proteasome. Furthermore, both were subjected to targeted cDNA microarray experiments during heat shock of native *P. furiosus*. The *P. furiosus* proteasome was the first archaeal form investigated that contains two forms of the beta subunit instead of one. Subsequently, one of the primary focuses of the study was to elucidate the roles of the two (48% identical) beta subunits.

Distinct differences in activity, stability, and level of ATPase-based stimulation were observed for the various proteasome forms. These differences were based on the presence or

absence of one of the three subunits and the assembly temperature. The beta-2 subunit appeared to be the catalytic center for proteinase activity, while the beta-1 subunit played a stabilizing role. PAN was able to stimulate the native form of the proteasome during degradation of polypeptides but inhibited the native heat-shocked form in the same reactions. It was concluded that PAN, which is highly up-regulated during heat shock, may stimulate the native proteasome form, while the heat-shocked proteasome (containing higher levels of beta-1) may associate with a different set of regulating proteins.

Biochemical and Biophysical Characterization of Compartmentalizing Proteases from the Hyperthermophilic Microorganism *Pyrococcus furiosus*

By

LARA SAMOFAL CHANG

A dissertation submitted to the Graduate Faculty of North Carolina State University in partial fulfillment of the requirements for the Degree of Doctor of Philosophy.

Chemical Engineering

Raleigh, NC

May, 2003

APPROVED BY

Dr. Todd R. Klaenhammer	Dr. David F. Ollis
Dr. Steven W. Peretti	Dr. Robert M. Kelly Chair of Advisory Committee

DEDICATION

The work summarized in this dissertation is dedicated to my parents, Alexander and Phyllis Samofal, who have always been there to lend their unending support and love.

BIOGRAPHY

Lara Samofal Chang was born Lara Ann Samofal on February 1, 1974 in the suburbs of Fairfax, Virginia. She remained in Northern Virginia until graduating from the magnet high school Thomas Jefferson High School for Science and Technology in 1992, where her interest in math and science were first cultivated. In spring 1996, she received a Bachelor of Science Degree in Engineering Science (with Distinction), with a specialization in Chemical Engineering, from University of Virginia. She took a year off to work for her undergraduate mentor, Dr. Erik Fernandez, in the Chemical Engineering Department at University of Virginia from 1996-1997. This was a continuation of work she had begun in her second summer at the University. In the fall of 1997, she then joined the Chemical Engineering Department at North Carolina State University (Raleigh, North Carolina), where she received her Master of Science in Chemical Engineering (Biotechnology minor) in December of 1999 with Dr. Robert M. Kelly.

During her time in Raleigh, she was married in November 1999 to Dr. Stephen T. Chang. She has also enjoyed volunteering with the Society for the Prevention of Cruelty to Animals (SPCA) Animal Shelter and the Helping Horse Therapeutic Riding Program, an incredible non-profit group that provides horseback riding lessons for mentally and physically challenged children and adults.

ACKNOWLEDGEMENTS

Many thanks to my advisor, Dr. Kelly, for his help and advice throughout this work. He has proven that his interests go beyond just the research; he cares for each and every one of his students both as researchers and as individuals. Financial support was provided by the Department of Education GAANN Biotechnology Fellowship and the Department of Energy (Energy Biosciences Program).

I would also like to thank all of my colleagues that have worked in the lab with me and given me many unforgettable lunch debates, the subjects of which spanned far beyond research. Thank you to my coffee friends for helping me through the last months. Many thanks to Dr. Amy Grunden for teaching me how to clone and express genes, skills that I used throughout my research.

Finally, I would like to thank my friends and family: mom, dad, Alex, Carol, Stephen, Teri and Liz, and many other wonderful friends. You have all been so supportive throughout my graduate career that I could not have finished without you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1:

Hyperthermophilic microorganisms

I. DEFINITION AND ECOLOGICAL SOURCES OF HYPERTHERMOPHILES

Extremophiles

Extremophiles are organisms, usually microorganisms, that can survive and reproduce at conditions that are considered extreme from the human vantage point (Madigan & Marrs, 1997). In most cases, extremophiles not only tolerate their extreme conditions, but also require them in order to reproduce. The most commonly known extremophiles include thermophiles or hyperthermophiles (heat-loving microbes), psychrophiles (cold-loving microbes), acidophiles (acidic-loving microbes), alkaliphiles (alkali-loving microbes), barophiles (high pressure-loving microbes), and halophiles (salt-loving microbes) (Adams *et al*., 1995; Madigan & Marrs, 1997). One of the most interesting characteristics of many of these extremophiles is that they cannot be placed into one classification. Instead, they often thrive in an environment with more than one extreme condition (Adams *et al.*, 1995). Another interesting observation is that they have some of the same physiological characteristics as many mesophilic microorganisms, or microorganisms that grow at more conventional conditions (temperature range of 15-45°C, neutral pH, atmospheric pressure) (Adams *et al.*, 1995; Stetter, 1996a). For example, there are many similarities between central metabolic pathways, such as glycolysis, for extremophiles and mesophiles. However, there are also many distinct differences in the physiology of extremophiles. These differences normally arise from the adaptation mechanisms that occur in extremophiles. It is these mechanisms that allow them to use a particular extreme environment to their advantage (Adams *et al.*, 1995).

Thermophiles and hyperthermophiles

The thermophiles and hyperthermophiles are the most studied groups of extremophiles. Thermophiles are microorganisms that prefer a temperature range of 45-80°C, whereas hyperthermophiles prefer temperatures above 80°C. At this time, the known upper limit for the growth temperatures of hyperthermophiles is 113°C (Adams, 1999). The one microorganism that is able to grow at this known upper temperature of life is *Pyrolobus fumarii*, found in the walls of deep-sea hydrothermal vents (Madigan & Marrs, 1997). It is a chemolithoautotrophic archaeon, or a microorganism that uses inorganic redox reactions as its energy source (chemolithotrophic) and needs carbon dioxide as its sole carbon source to build up organic cell material (autotrophic) (Stetter, 1999). This organism is so adapted to high temperatures that it cannot even grow at or below 85°C (Huber & Stetter, 1998)

Although the existence of thermophiles has been known for some time, hyperthermophiles are a relatively new discovery (Stetter, 1996a). Thomas D. Brock isolated one of the first discovered thermophiles in 1964 (Gross, 1998). This thermophilic bacterium, named *Thermus aquaticus*, was initially found in the hot springs of Yellowstone National Park. Since then, researchers began to find hyperthermophiles in a wide range of environments, both natural and artificial (Stetter, 1996a). These include locales near volcanic emissions, where deep magma chambers heat up soils and surface waters, smoldering coal refuse piles, and hot outflows from geothermal and atomic power plants (Stetter, 1996a). Most of the hyperthermophiles that have been isolated, however, originated from marine geothermal environments, including shallow- and deep-sea hydrothermal vents (Adams, 1999). Furthermore, approximately three-quarters of the hyperthermophiles known today come from deep-sea environments (Gross, 1998).

Since many hyperthermophiles are found in deep-sea environments, it also follows that almost all are strict anaerobes (Adams, 1999). Although they include species of methanogens, iron-oxidizers and sulfate reducers, most are obligate heterotrophs that depend on the reduction of elemental sulfur to hydrogen sulfide for maximum growth (Adams, 1999). This type of adaptation by hyperthermophiles from hydrothermal vents is due to the environment in and around these hot "smoker" chimneys. These "black smokers" were initially discovered in 1977 and have since been found in many tectonically active areas of the ocean floor (Prieur, 1997). They have been observed at depths of 800-3500 meters and in both the Atlantic and Pacific Oceans. These black smokers form when seawater penetrates through cracks in the sea floor and becomes heated upon reaching the magma chamber deep below. As the water heats, it leaches minerals from the magmatic basalts and, simultaneously, circulates back out into the cold ocean water. The resulting hydrothermal fluids are acidic, reduced, and enriched with chemicals such as heavy metals, methane and hydrogen sulfide. The smokers are formed from this event through the mixing of the hot hydrothermal fluids and the cold seawater, which causes mineral precipitations and, finally, the smokers. Table 1.1 lists a variety of hyperthermophiles that have been isolated from several different environments, including marine hydrothermal vents. In addition, some basic information, including temperature, pH, and oxygen preferences, is given.

Table 1.1: Basic features of hyperthermophiles (type species). Reconstructed from Stetter (1999).

Species	Min. growth temp. (°C)	Opt. growth temp. (°C)	Max. growth temp. (°C)	pH	Aerobic (ae) or anaerobic (an)	Biotope (marine (M) or terrestrial (T))	DNA G+C (mol %)	Morphology
Thermotoga maritima	55	80	90	5.5-9	an	M	46	Rods with sheath
Aquifex pyrophilus	67	85	95	5.4-7.5	ae	M	40	Rods
Sulfolobus acidocaldarius	60	75	85	1-5	ae	T	37	Lobed cocci
Metallosphaera sedula	50	75	80	1-4.5	ae	T	45	Cocci
Acidianus infernus	60	88	95	1.5-5	ae/an	T	31	Lobed cocci
Stygiolobus azoricus	57	80	89	1-5.5	an	T	38	Lobed cocci
Thermoproteus tenax	70	88	97	2.5-6	an	T	56	Regular rods
Pyrobaculum islandicum	74	100	103	5-7	an	T	46	Regular rods
Pyrobaculum aerophilum	75	100	104	5.8-9	ae/an	M	52	Regular rods
Thermofilum pendens	70	88	95	4-6.5	an	T	57	Slender regular rods
Desulfurococcus mobilis	70	85	95	4.5-7	an	M	51	Cocci
Thermosphaera aggregans	67	85	90	5-7	an	T	46	Cocci in aggregates
Sulfophobococcus zilligii	70	85	95	6.5-8.5	an	T	54	Cocci
Staphylothermus marinus	65	92	98	4.5-8.5	an	M	35	Cocci in aggregates

Table 1.1 ***(continued)***

Species	Min. growth temp. (°C)	Opt. growth temp. (°C)	Max. growth temp. (°C)	pH	Aerobic (ae) or anaerobic (an)	Biotope (marine (M) or terrestrial (T))	DNA G+C (mol %)	Morphology
Thermodiscus maritimus	75	88	98	5-7	an	M	49	Disks
Aeropyrum pernix	70	90	100	5-9	ae	M	67	Irregular cocci
Stetteria hydrogenophila	70	95	102	4.5-7	an	M	65	Irregular disks
Igneococcus islandicus	65	90	100	3.9-6.3	an	M	41	Irregular cocci
Pyrodictium occultum	82	105	110	5-7	an	M	62	Disks with cannulae
Hyperthermus butylicus	80	101	108	7	an	M	56	Lobed cocci
Pyrolobus fumarii	90	106	113	4.0-6.5	ae/an	M	53	Lobed cocci
Thermococcus celer	75	87	93	4-7	an	M	57	Cocci
Pyrococcus furiosus	70	100	105	5-9	an	M	38	Cocci
Archaeoglobus fulgidus	60	83	95	5.5-7.5	an	M	46	Irregular cocci
Ferroglobus placidus	65	85	95	6-8.5	an	M	43	Irregular cocci
Methanothermus sociabilis	65	88	97	5.5-7.5	an	T	33	Rods in clusters
Methanopyrus kandleri	84	98	110	5.5-7	an	M	60	Rods in chains
Methanococcus jannaschii	50	85	86	5.5-6.5	an	M	31	Irregular cocci
Methanococcus igneus	45	88	91	5-7.5	an	M	31	Irregular cocci

II. CLASSIFICATION AND CHARACTERISTICS OF HYPERTHERMOPHILES

Domains of life

Based on 16S rRNA sequence classification, one recent phylogenetic tree of life includes three domains: archaea (previously called archaebacteria), bacteria, and eukarya. The archaea can be further broken down into two major kingdoms: the *Crenarchaeota*, including the orders *Sulfalobales* and *Thermoproteales*, and the *Euryarchaeota*, including the hyperthermophiles, extreme halophiles, and methanogens. In addition, there has been recent evidence of a third kingdom, tentatively named *Korarchaeota* (Huber & Stetter, 1998). Figure 1.1 gives an illustration of this phylogenetic tree.

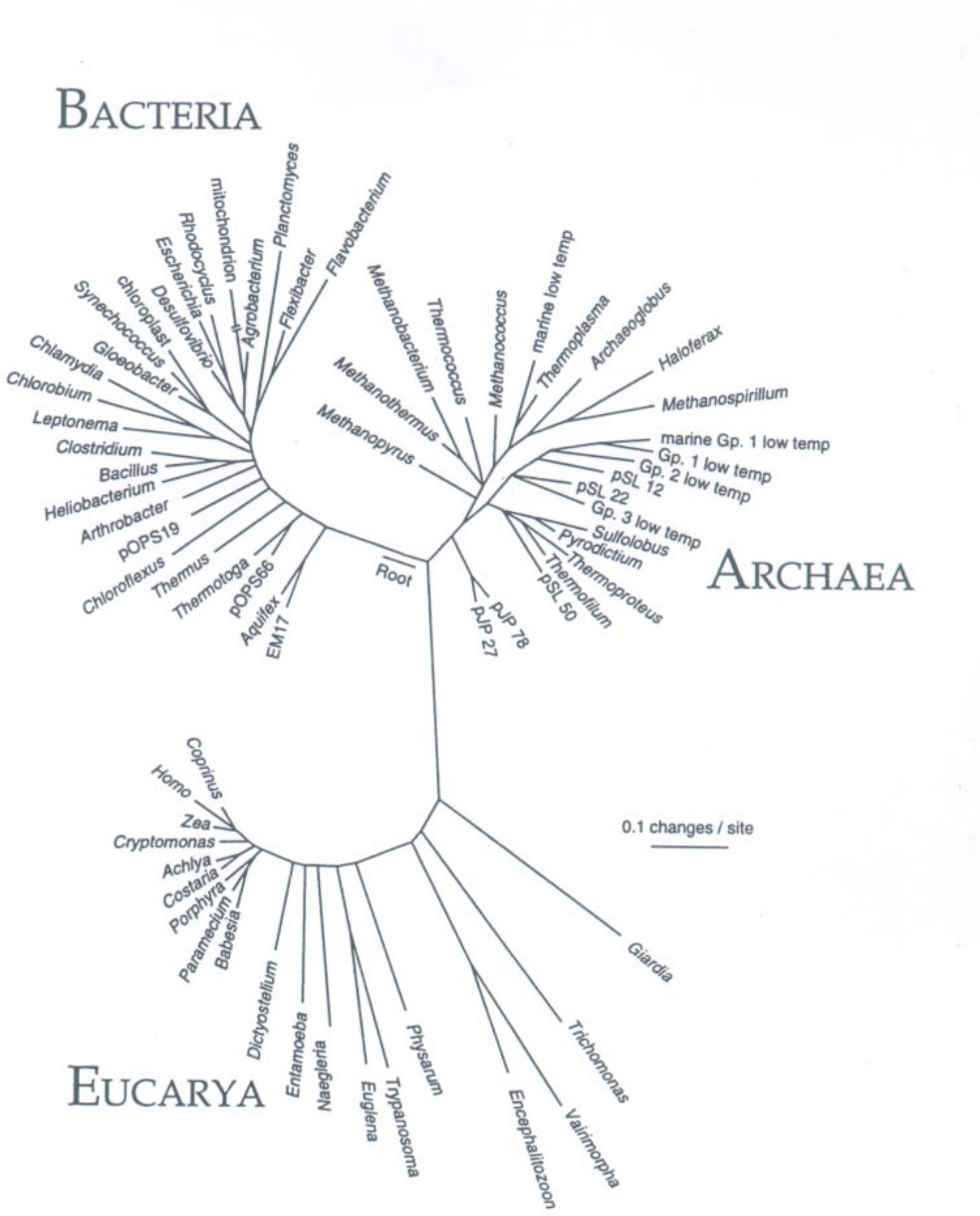


Figure 1.1: 16S rRNA-based phylogenetic tree showing the placement of the three domains of life. Reproduced from Hicks (1998).

Hyperthermophiles in the tree of life

Of the over 70 species of hyperthermophiles known today, a majority belong to the archaea (Adams, 1999; Prieur, 1997; Stetter, 1999). Within the bacteria, the only two sets of microorganisms that are hyperthermophilic include members of the Thermatogales and the genus *Aquifex* (Adams & Kelly, 1994; Stetter, 1999). No microbial eukarya yet discovered can tolerate long-term exposure to temperatures higher than about 60°C (Madigan & Marrs, 1997). From Figure 1.1, it is clear that the hyperthermophiles form a cluster around the root, suggesting a slow clock of evolution (Stetter, 1996b; Stetter, 1999). As a rule, these most primitive organisms with the deepest and shortest lineages exhibit the highest growth temperatures (Stetter, 1999).

The distinct characteristics that set the archaea apart from the other domains of life are threefold: 1) the presence of ether-linked, rather than ester-linked, membrane lipids, 2) the absence of murien in the cell wall, and 3) the expression of unusual enzymes for different metabolic pathways (Hicks, 1998; Stetter, 1999). In addition, there are many characteristics of archaea that are either similar to bacteria, but not eukarya, or vice versa (Adams *et al.*, 1995). For example, archaea possess histones phylogenetically related to eukaryotic core histones (Stetter, 1999). In addition, archaeal promoters, RNA polymerases, and transcription factors are similar to those in eukaryotic cells (Hicks, 1998). However, like bacteria, the archaea lack organelles and contain genomes that are double stranded circular DNA molecules of 1.7-2.9 Mb. Bacteria and eukarya have distinct codon preferences, while the archaea use certain codons preferred by bacteria and other codons preferred by eukarya (Adams *et al.*, 1995). Also, translation in archaea seems to be a combination of eukaryotic

and bacterial mechanisms. One main difference is that some ribosome binding sites are downstream of the start codon (Hicks, 1998).

III. APPLICATIONS OF HYPERTHERMOPHILES

The greatest application of hyperthermophiles is the use of their proteins, which are usually thermostable. One of the first important uses for thermostable proteins was as catalysts in the polymerase chain reaction (PCR) (Cowan, 1992). Although the original application makes use of thermostable *Taq* polymerase from *T. aquaticus*, which is not technically a hyperthermophile, many hyperthermophilic DNA polymerases are now being used for PCR reactions (Adams & Kelly, 1998). In addition, industrial applications are requiring the use of a more thermostable glucose isomerase that works at extreme temperatures (95°C) instead of the presently used 60°C. The more thermostable version of glucose isomerase is desired to help cut out a concentration step in making 55% fructose syrup (Zeikus *et al*., 1998). Another application could be the use of thermostable enzymatic "breakers" for hydrolyzing guar gum during hydraulic fracturing of oil and gas wells (Adams & Kelly, 1998).

By far, the largest contribution that could emerge from the study of hyperthermophilic proteins is knowledge of the mechanisms for protein stabilization and function at any temperature (Adams & Kelly, 1994). In the quest to find these mechanisms, some of the most recent research has been the direct comparison of related thermophilic and mesophilic enzymes (Beadle *et al*., 1999; Haney *et al*., 1999). It is recognized that increased

thermostability by a protein is often achieved by minor changes, rather than by major changes in primary, secondary, or tertiary structure. In direct comparisons between hyperthermophilic and mesophilic proteins, there were observed changes in number of ionic pairs per residue and subtle interactions that lead to global changes in structure, such as a decreased surface area to volume ratio for the more thermostable protein (Adams *et al.*, 1995). However, from all of these very focused studies, a general rule that allows one to create a thermostable enzyme from a mesophilic starting enzyme does not yet exist.

IV. *PYROCOCCUS FURIOSUS*: THE HYPERTHERMOPHILE OF THIS STUDY

The hyperthermophilic archaeon *Pyrococcus furiosus* was first isolated from geothermally-heated sediments near Vulcano Island, Italy (Fiala & Stetter, 1986). The optimum growth temperature of this archeaon is 100°C, while its minimum and maximum growth temperatures are 70°C and 105°C, respectively (Stetter, 1999). Members of the genus *Pyrococcus* normally ferment peptides and saccharides to carbon dioxide, hydrogen, and fatty acids. Significant amounts of alanine are produced in the absence of sulfur and, if a polysulfide source is available in the growth medium, large amounts of hydrogen sulfide will be produced. *P. furiosus* appears to grow the best on alpha-linked polysaccharides, including starch and glycogen, and disaccharides such as maltose and cellobiose. Although it was previously reported that *P. furiosus* could not grow on glucose or other beta-linked sugars, it has more recently been found to grow on several beta-linked glucans (Driskill *et al*., 1999).

Furthermore, glucose had only a slight effect on the growth of *P. furiosus* in batch culture, but a significantly positive effect in fed-batch conditions.

Another important characteristic of *P. furiosus* is that it possesses an altered version of the Embden-Meyerhof-Parnas (EMP) pathway compared to the one found in bacteria. For example, it involves ADP-dependent kinases, glucokinases, and phosphofructokinase (Kengen *et al*., 1996). Even though ADP and ATP have similar thermodynamic properties in regard to the free energy released upon hydrolysis, the reason for the presence of AMP-forming enzymes is not understood. In addition, the three oxidation steps between glucose and acetate are catalyzed by ferredoxin-reducing agents, rather than dehydrogenase-type NAD-reducing enzymes (Adams & Kelly, 1994). The reason for the use of ferredoxin may be because of its extreme thermostability compared to NAD (Adams & Kelly, 1994; Leuschner & Antranikian, 1995). In-depth studies on the physiological and bioenergetic characteristics of *P. furiosus* have been carried out and are further discussed by Kelly et al. (1994) and Kengen et al. (1996).

V. REFERENCES

Adams, M. W. W. (1999). The biochemical diversity of life near and above 100 degrees C in marine environments. *J Appl Microbiol* **85**, 108S-117S.

Adams, M. W. W. & Kelly, R. M. (1994). Thermostability and thermoactivity of enzymes from hyperthermophilic Archaea. *Bioorg Med Chem* **2**(7), 659-667.

Adams, M. W. W. & Kelly, R. M. (1998). Finding and using hyperthermophilic enzymes. *Trends Biotechnol* **16**(8), 329-332.

Adams, M. W. W., Perler, F. B. & Kelly, R. M. (1995). Extremozymes: expanding the limits of biocatalysis. *Biotechnology* **13**(7), 662-668.

Beadle, B. M., Baase, W. A., Wilson, D. B., Gilkes, N. R. & Shoichet, B. K. (1999). Comparing the Thermodynamic Stabilities of a Related Thermophilic and Mesophilic Enzyme. *Biochemistry* **38**(8), 2570-2576.

Cowan, D. A. (1992). Biotechnology of the Archaea. *Trends Biotechnol* **10**(9), 315-323.

Driskill, L. E., Kusy, K., Bauer, M. W. & Kelly, R. M. (1999). Relationship between glycosyl hydrolase inventory and growth physiology of the hyperthermophile *Pyrococcus furiousus* on carbohydrate-based media. *Appl Environ Microbiol* **65**(3), 893-897.

Fiala, G. & Stetter, K. O. (1986). *Pyrococcus furiosus* Sp-Nov Represents a Novel Genus of Marine Heterotrophic Archaebacteria Growing Optimally at 100 Degrees C. *Arch Microbiol* **145**(1), 56-61.

Gross, M. (1998). *Life on the Edge: Amazing Creatures Thriving in Extreme Environments*, Plenum Press, New York, NY.

Haney, P. J., Badger, J. H., Buldak, G. L., Reich, C. I., Woese, C. R. & Olsen, G. J. (1999). Thermal adaptation analyzed by comparison of protein sequences from mesophilic and extremely thermophilic *Methanococcus* species. *Proc Natl Acad Sci U S A* **96**(7), 3578-3583.

Hicks, P. M. (1998). Genetic, Biochemical, and Biophysical Characteristics of Intracellular Proteases from the Hyperthermophilic Archaeon *Pyrococcus furiosus* and the Hyperthermophilic Bacterium *Thermotoga maritima*. PhD dissertation, North Carolina State University.

Huber, H. & Stetter, K. O. (1998). Hyperthermophiles and their possible potential in biotechnology. *J Biotechnol* **64**(1), 39-52.

Kelly, R. M. & Adams, M. W. W. (1994). Metabolism in Hyperthermophilic Microorganisms. *Anton Leeuw Int J G* **66**(1-3), 247-270.

Kengen, S. W. M., Stams, A. J. M. & de Vos, W. M. (1996). Sugar metabolism of hyperthermophiles. *FEMS Microbiol Rev* **18**(2-3), 119-137.

Leuschner, C. & Antranikian, G. (1995). Heat-stable enzymes from extremely thermophilic and hyperthermophilic microorganisms. *World J Microbiol Biotechnol* **11**(1), 95-114.

Madigan, M. T. & Marrs, B. L. (1997). Extremophiles. *Sci Am* (April), 2-7.

Prieur, D. (1997). Microbiology of deep-sea hydrothermal vents. *Trends Biotechnol* **15**(7), 242-244.

Stetter, K. O. (1996a). Hyperthermophiles in the history of life. *Evolution of Hydrothermal Ecosystems on Earth (and Mars?)* **202**, 1-18.

Stetter, K. O. (1996b). Hyperthermophilic procaryotes. *FEMS Microbiol Rev* **18**(2-3), 149-158.

Stetter, K. O. (1999). Extremophiles and their adaptation to hot environments. *FEBS Lett* **452**(1-2), 22-25.

Zeikus, J. G., Vieille, C. & Savchenko, A. (1998). Thermozymes: biotechnology and structure-function relationships. *Extremophiles* **2**(3), 179-183.

CHAPTER 2:

Proteases: definitions and applications

I. DEFINITION AND CLASSIFICATION OF PROTEASES

The proteolytic reaction

Proteases are a group of proteins included in the subclass hydrolases, within the main class enzymes (Polgár, 1990). Using a hydrolysis reaction, they cleave specific peptide bonds that occur within the primary structure of polypeptides (Bauer *et al*., 1996; Moran *et al*., 1994). Many proteases catalyze the same general chemical reaction, a hydrolysis reaction similar to that shown in Figure 2.1 (Moran *et al.*, 1994). Peptide bond hydrolysis is an addition-elimination reaction, with the protease acting as a nucleophile or reacting with water to create one (Bauer *et al.*, 1996; Polgár, 1990). Generally, the nucleophile forms a tetrahedral intermediate with the carbonyl carbon of the peptide bond. An amine leaving group is consequently formed and diffuses from the active site, simultaneously replaced by a water molecule. Depending on the particular protease, an acyl-enzyme adduct can be formed, as in Figure 2.1. Finally, a second tetrahedral intermediate is formed, leading to the production of a carboxylate product, a proton, and a regenerated free enzyme. However, the specific mechanism varies among the proteases and this variation is what is generally used to classify them.

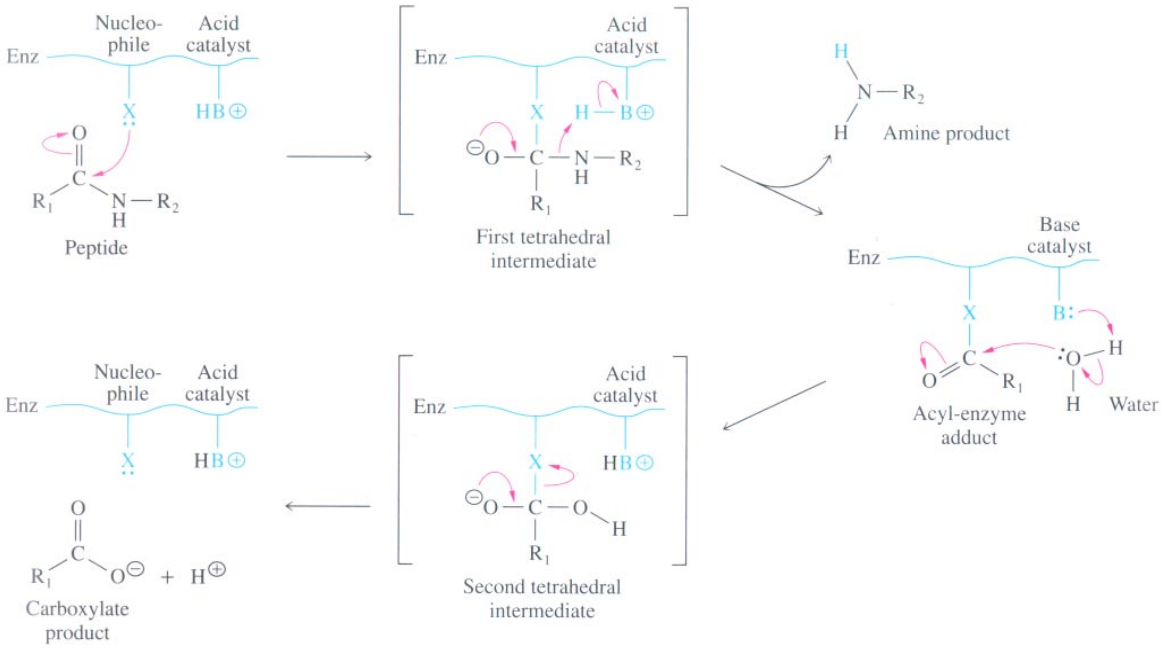


Figure 2.1: General mechanism for the enzymatic hydrolysis of a peptide substrate. Reproduced from Moran et al. (p 7.12) (1994).

Classes of proteases

Each type of protease has a specific preference for one or more peptide bonds, depending on the neighboring amino acid residues (Barrett, 1994; Polgár, 1990). This characteristic allows the proteases to be further broken down into two subclasses: peptidases (exopeptidases) and proteinases. Exopeptidases preferentially cleave at the carboxyl- or amino-terminus of a peptide chain and can be sorted into several subclasses depending on the position of the cleavage site in the polypeptide substrate and at which terminus the enzyme acts. These subclasses can be broken down into 6 different groups: those that cleave at the carboxyl- or amino- terminus and, further, those that cleave one, two, or three residues from the chosen target terminus (Barrett, 1994; Polgár, 1990). Table 2.1 lists the different groups of exopeptidases (Rao *et al*., 1998).

Table 2.1: Classification of exopeptidases by type of reaction catalyzed. The open circles represent amino acid residues and filled circles represent residues cleaved off the end of the target protein by the corresponding exopeptidase. The stars represent the blocked termini that provide substrates for some of the omega peptidases. Recreated from Barrett (1994) and Rao et al. (1998).

Protease	Mode of action	E.C. Number
Exopeptidases		
Aminopeptidases	●-↓-○-○-○-○---	3.4.11
Dipeptidyl peptidase	●-●-↓-○-○-○---	3.4.14
Tripeptidyl peptidase	●-●-●-↓-○-○---	3.4.14
Carboxypeptidase	---○-○-○-○-○-↓-●	3.4.16-3.4.18
Serine type protease		3.4.16
Metalloprotease		3.4.17
Cysteine type protease		3.4.18
Peptidyl dipeptidase	---○-○-○-○-↓-●-●	3.4.15
Dipeptidases	●-↓-●	3.4.13
Omega peptidases	*-●-↓-○-○--- ---○-○-○-↓-●-*	3.4.19 3.4.19
Endopeptidases	---○-○-○-↓-○-○-○---	3.4.21-3.4.34
Serine protease		3.4.21
Cysteine protease		3.4.22
Aspartic protease		3.4.23
Metalloprotease		3.4.24
Endopeptidases of unknown catalytic mechanism		3.4.99

Proteinases, although often treated as synonymous with proteases, only refer to those proteases that act at the inside of a polypeptide chain, termed endopeptidases. The endopeptidases generally are considered to cleave their specific peptide substrates through interaction with not only the amino acid residues immediately adjacent to the cleavage site, but also with neighboring residues. Figure 2.2 shows a schematic representation of an endopeptidase active site bound to a polypeptide substrate. Note that the substrate is almost always viewed with the amino-terminus on the left (Hicks, 1998). Cleavage of the substrate occurs at the bond between the two residues labeled P_1 and P_1' (Hicks, 1998; Keil, 1992).

The amino acid residues in all of the binding positions (P_3-P_3') can have an effect on the rate of hydrolysis based on two main effects: 1) steric hindrances and 2) the strength and stability of the interaction between the enzyme and substrate (Hicks, 1998).

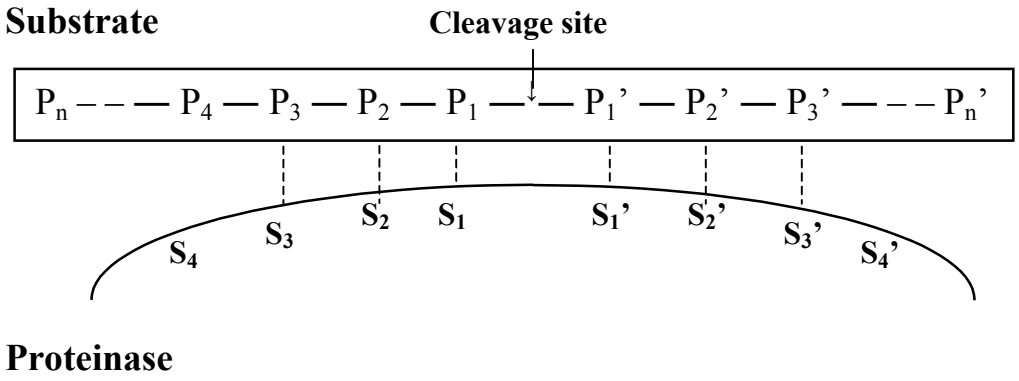


Figure 2.2: Schematic representation of the proteinase-substrate complex with six binding sites. Cleavage occurs between amino acid residues P_1 and P_1' (Hicks, 1998; Keil, 1992).

According to Enzyme Nomenclature, the endopeptidases are designated, based on active site geometry and enzymatic mechanism, into five groups (Barrett, 1994; Keil, 1992; Neurath, 1989; Polgár, 1990): serine, cysteine, aspartic, metallo-, and unknown proteinases. These groups are then further categorized into separate clans, based on the amino acid sequences of the active sites and the tertiary structures of the proteinases. Table 2.2 lists the distinguishing characteristics of the five proteinase groups and examples of the most commonly known proteinases within each group.

The clan grouping separates the serine proteinases into at least 6 different clans (Barrett & Rawlings, 1993; Rawlings & Barrett, 1994b). In addition, there are over 20 families that are formed based on evolutionary origin (Rawlings & Barrett, 1994b). Several

of these families are subsets within the clans, however, many cannot be grouped into the existing set of clans. The cysteine proteinases and aspartic proteinases are broken into at least 20 and 3 families, respectively (Barrett & Rawlings, 1993; Rawlings & Barrett, 1994a; Rawlings & Barrett, 1995b). Like the serine proteinases, many of the families are isolated instead of being grouped within the existing clans. The metalloproteinases are separated into at least five distinct clans based on the metal binding motif present at the active site (Hase & Finkelstein, 1993; Rao *et al.*, 1998; Rawlings & Barrett, 1995a). Among and beyond these 5 clans there are at least 30 families of metallopeptidases (Rawlings & Barrett, 1995a).

Table 2.2: Definitions of the protease groups and examples of known proteases within each group.

Protease	Definition	Examples
serine proteinases	possess catalytically competent Ser and His residues in their active centers; mechanism of action involves formation of a covalent acyl-enzyme intermediate (see Figure 2.1) *(this group is also found in the exopeptidase, oligopeptidase, and omega peptidase groups (Rao *et al.*, 1998))	-chymotrypsin family (chymotrypsin, trypsin, thrombin, plasmin) from clan SA -subtilisin family from clan SB -Clp family (ClpP) from clan SG (Rawlings & Barrett, 1994)
cysteine (thiol) proteinases	possess Cys and His residues in the active site; mechanism of action involves formation of a covalent acyl-enzyme intermediate (see Figure 2.1)	-papain and calpain families from clan CA -several viral endopeptidase families from clans CB, CC, or unclassified -caspase family (Steinberg, 1998; Stennicke & Salvesen, 1998)
aspartic proteinases	use two acidic (negatively-charged) residues in the catalytic process; mechanism of action involves direct hydrolysis by water	-A1 - pepsin family (pepsin A, chymosin) -A2 - retropepsin -A3 - pararetropepsin endopeptidases (Rawlings & Barrett, 1995; Rao & Tanksale, 1998)
metallo-proteinases	use a metal ion and glutamic acid residue in the catalytic process; mechanism of action involves direct hydrolysis by water – the tetrahedral transition state is formed by a pentacoordinated metal ion (usually zinc) (Bauer & Halio 1996; Hase & Finkelstein 1993)	-thermolysin family from clan MA (thermolysin, bacillolysin) -interstitial collagenase family from clan MB -carboxypeptidase A family (carboxypeptidase A subfamily, carboxypeptidase H subfamily) -pitrilysin family from clan of 'other metallopeptidases with known metal ligands' (Rawlings & Barrett, 1995)
proteinases of unknown catalytic tp.	proteinases that do not fit into the first four classes	-endopeptidase IV -aminopeptidase iap (Rawlings & Barrett, 1995)

There are many new proteases that have been discovered that do not fit into any one class or family. One type of non-classifiable protease is a proteasome that was found to contain a novel threonine active site. Two examples include the archaeal form (Maupin-Furlow & Ferry, 1995) and yeast form (Arendt & Hochstrasser, 1997; Groll *et al*., 1997) of this proteasome. This type of protease may be one of the first in an entirely new class. Another non-classifiable protease was recently isolated from *Archachatina ventricosa*. This endopeptidase was shown to be completely inactivated by the chelating agents EDTA and 1,10-phenanthroline, which are metalloprotease inhibitors. But, unlike the other known families of metalloproteinases, this protease preferentially cleaves areas where there is a threonine residue present at the P_1' position (see Figure 2.2) (Niamke *et al*., 1999).

A different set of proteases that do not fit into any one class or family includes those that fit into more than one of the four defined classes. In particular, the carboxypeptidases are zinc-containing metalloproteases, with a similar proteolytic mechanism to the family of metalloproteases in the endopeptidases. However, they act as exopeptidases, cleaving at the ends of polypeptide chains (Neurath, 1989).

From the few examples of non- or multi-classifiable proteases, it is clear that the present classification system is not foolproof. In general, this system has evolved from identification by molecular size, charge, or substrate specificity (Neurath, 1989), to classification by catalytic mechanism and active-site structure, and, further, to more specific classification by active site amino acid residues or sequences and/or tertiary structure. Most recently, the major criteria for classifying proteases have become: 1) the reaction catalyzed, 2) the chemical nature of the catalytic site, and 3) the evolutionary relationship/structural relationship (Barrett, 1994). However, in some cases, the endopeptidases are grouped by

source (microbial metalloproteases, lysozomal cys-proteinases) or biological function (kinases, collagenases) (Moran *et al.*, 1994). This type of classification, though not generally used, completely disregards protease specificity.

Protease inhibitors

In addition to classification by catalytic mechanism and active-site structure, proteases are further grouped by sensitivity to inhibitors (Bauer *et al.*, 1996; Bode & Huber, 1993; Salvesen & Nagase, 1989). Consequently, an unknown protease can be placed in one of the subclasses based on its activity, or lack of activity, in the presence of certain inhibitors (Dunn, 1989). Inhibitors are generally known to be any compound, including a competing substrate, that can decrease the measured rate of hydrolysis of a substrate (Salvesen & Nagase, 1989). In addition, this definition should be further modified to exclude enzyme inactivators including metal chelators and denaturants such as guanidinium chloride and urea. Therefore, true inhibitors are those that interact with the protease active site to form a complex (Salvesen & Nagase, 1989). Barrett (1994) offers a useful discussion on many specific inhibitors, as well as a laboratory procedure for using inhibitors to find the catalytic type of an unknown protease. Some of the most well known class-specific inhibitors include (Barrett, 1994; Bode & Huber, 1993; Hicks, 1998; Salvesen & Nagase, 1989):

1) **serine proteinase inhibitors:**
   - organophosphates (e.g., diisopropylphosphofluoridate (DFP))
   - coumarins (e.g., 3,4-dichloroisocoumarin)
   - serine protease inhibitors (serpins)

2) **cysteine proteinase inhibitors:**
   - peptide diazomethanes
   - peptide epoxides (e.g., L-trans-epoxysuccinyl-leucylamido(4-guanidino)butane (E-64))
   - cystatins

3) **aspartic proteinase inhibitors:**
   - pepstatin

4) **metallo-proteinase inhibitors:**
   - natural protein inhibitor TIMP (tissue inhibitor of metallo-proteinases)
   - specific synthetic inhibitors (e.g., phosphoramidon)
   - generally, inactivating metal chelators (not true inhibitors) are used in experiments (e.g., EDTA, EGTA, 1,10-phenanthroline)

Many of the listed inhibitors, although all possessing the ability to inhibit proteases, interact and inhibit in very different ways. For example, just within the group of different serine protease inhibitors, many of them are "conformationally rigid simulations of optimal serine proteinase substrates" (Wright, 1996). In contrast, members of the group of serine protease inhibitors (serpins), which are separate from the rest, are larger and tend to form tight, stable complexes with serine proteases. In addition, their function is restricted to single encounters with the target proteinase. Once they are recovered from an inhibition reaction, they are often found in a cleaved form. As a result, they are nicknamed the 'suicide substrate inhibitors' (Wright, 1996). Finally, unlike the smaller, rigid inhibitors, serpins do not present an optimal binding site for proteinases in their free state. Instead, they must change conformation in order to bind and react with proteinases (Wright, 1996).

In addition to the many class-specific protease inhibitors, there are several inhibitors that cross the imaginary lines between the different families of proteases. Some, such as *Streptomyces* Subtilisin Inhibitor (SSI) and the sulphonyl fluoride phenylmethylsulphonyl fluoride (PMSF), exhibit dual activity between two different families. SSI is active against both subtilisins and a metalloendopeptidase. PMSF is active against serine proteinases, by covalently modifying the active site serine (Hicks, 1998), and cysteine proteinases (Bode & Huber, 1993; Salvesen & Nagase, 1989). Other examples of dual-activity inhibitors are peptide chloromethyl ketones, which act in a substrate-like manner by reacting with the active-site histidine of serine proteases (Hicks, 1998; Salvesen & Nagase, 1989; Tsilikounas *et al*., 1996), but can also inhibit cysteine proteases (Salvesen & Nagase, 1989). Other inhibitors are even less discriminating in that they are almost completely non-specific. Examples of this type of inhibitor include the α-macroglobulins, which bind to and inhibit most proteinases (Salvesen & Nagase, 1989).

Simple versus complex proteases

Another way to differentiate between the proteases is to divide them into two main groups, based solely on overall structure. The first group is termed the 'simple' proteases and includes all of the single-domain proteases, such as protease La (Lon) and FtsH (HflB). The second group is named the 'complex' proteases and includes the multi-catalytic and multi-domain proteases such as the proteasome, Clp family, HslVU, and others.

II. APPLICATIONS OF PROTEASES

Industrial applications

For decades, proteases have been used in a wide range of industrial applications for both hydrolytic and synthetic reactions (Hicks, 1998). Specifically, proteases are used extensively in the detergent (von der Osten *et al*., 1993), leather, and food processing industries (Rao *et al.*, 1998). Many industrial proteins are obtained from microbial sources, and microbial proteases, in particular, account for nearly 40% of the total worldwide enzyme sales (Rao *et al.*, 1998). As a rule, these microbial sources must be safe, i.e., they must be non-toxic, non-pathogenic, and should not generally produce antibiotics.

Two main advantages to using microorganisms to produce enzyme products are: 1) the ability of the microorganisms to grow quickly and produce a large, steady supply of enzyme and 2) the fact that microbial enzymes are often more stable than their corresponding counterparts from plant and animal sources (Headon & Walsh, 1994). In particular, many industrial proteases used in detergents, brewing, baking, leather bating, and meat tenderization come from sources such as *Bacillus amyloliquefaciens*, *B. subtilis*, *Aspergillus oryzae*, and *Streptomyces* spp. These types of proteases are generally produced during a fermentation process and secreted into extracellular media as they are produced (Headon & Walsh, 1994). Figure 2.3 shows the basic sequence for the downstream purification of bulk amounts of industrial enzyme from a microbial source (Headon & Walsh, 1994).

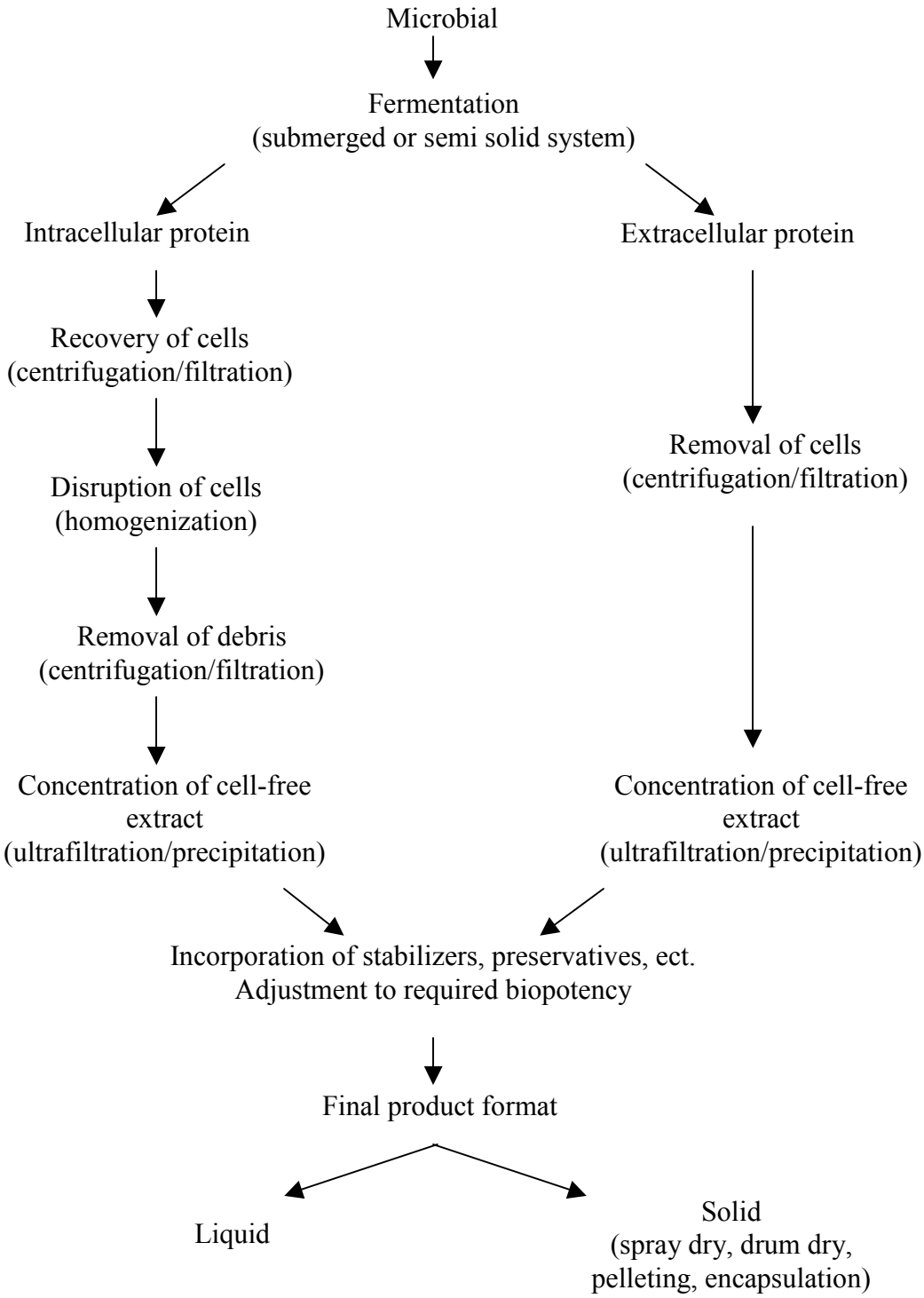


Figure 2.3: Generalized downstream purification scheme as often applied to the production of bulk industrial enzymes. Reconstructed from Headon et al. (1994).

Proteolytic reactions in industry

Bacterial proteases are used extensively in the detergent industry, which accounts for nearly 25% of the worldwide sale of enzymes (Rao *et al.*, 1998). Starting in 1913, proteases from crude pancreatic extracts were added to laundry detergents to achieve better performance in the removal of proteinaceous stains (Rao *et al.*, 1998). In the late fifties to early sixties, bacterial proteases were first used in a commercial detergent (Rao *et al.*, 1998; von der Osten *et al.*, 1993). Today, the most popular proteases for use in detergents are all subtilisin-like serine proteases from *Bacillus amyloliquefaciens*, *B. lichenformis*, highly alkalophilic bacilli such as *B. lentus* (Rao *et al.*, 1998; von der Osten *et al.*, 1993), and protein-engineered variants of these natural proteases (Nordisk, 1994). For example, Novo Nordisk produces a number of detergent proteases from each group: Alcalase® and Savinase® (natural bacterial serine proteases from *Bacillus* strains), Durazyme® and Everlase™ (protein-engineered variants of Savinase®), and Esperase® (bacterial protease effective under strongly alkaline conditions) (Nordisk, 1994). Esperase® and Savinase®, in particular, have one very advantageous characteristic: their isoelectric points (pI) are very high, approximately 11, which allows them to withstand high pH ranges. In general, the key parameter for the best performance of a protease in a detergent is the pI. The protease is considered most suitable for a detergent if its pI coincides with the pH of the detergent solution (Rao *et al.*, 1998).

Even though there is a long history of detergents with protease additives, there are still some major improvements to be made in the detergent industry. Many times, detergents with proteases do not have a long shelf life because the proteases inactivate over time. This is usually caused by oxidative inactivation by bleaching agents or autoproteolytic degradation in powder and liquid detergents, respectively. To help overcome these obstacles,

thermostable proteases have become the focus of the detergent industry (von der Osten *et al.*, 1993). Researchers are looking toward two avenues: using naturally thermostable proteases or altering the presently used proteases to make them thermostable. However, the latter choice has proven difficult because stabilization improvements among structurally related proteases lack universality (von der Osten *et al.*, 1993).

Although microbial sources are often advantageous to use in the production of industrial enzymes, there are some proteases of industrial significance that are obtained from plant sources. The most well known protease, derived from the plant source *Carica papaya*, is the cysteine protease papain, or vegetable pepsin (Headon & Walsh, 1994; Rao *et al.*, 1998). It is active between pH 5 and 9 and stable up to 80 or 90°C in the presence of substrates (Rao *et al.*, 1998). In industry, it is used in meat tenderization (Headon & Walsh, 1994). Some other popular plant-derived proteases are ficin, a debriding agent from *Ficus glabrata*, and bromelain, a protease from pineapple fruit and stem that is used for chill proofing of beers (Headon & Walsh, 1994). Finally, keratinases, obtained from botanical groups of plants, have the ability to degrade hair and, as such, are primarily used for the prevention of wastewater system clogging (Rao *et al.*, 1998).

Synthesis reactions in industry

Together with the use of proteolytic hydrolysis reactions, proteolytic synthesis reactions are used in industry for the production of peptide hormones, neuropeptides, and sweeteners (aspartame), and they are used in protein modifications for the partial synthesis of insulin, oligosaccharides, and antibiotics. Synthesis reactions by proteases can be created by reversal of the equilibrium of the hydrolysis reactions normally catalyzed by the proteases

(Hicks, 1998; Kasche, 1989). This alteration of protease activity can be created by changes in pH, temperature, ionic strength, and solvent composition (Kasche, 1989). Furthermore, it has only been within the past decade that the profound impact of solvent choice on enzyme selectivity has been realized. By simply changing the solvent for a particular enzyme, one can alter the selectivity of the enzyme as efficiently as protein engineering can (Wescott & Klibanov, 1994). The specific type of alterable selectivity can include substrate, enantio-, prochiral, regio-, and chemoselectivity of enzyme-catalyzed reactions (Wescott & Klibanov, 1994).

Beyond the ability of solvent choice to alter enzyme selectivity, there are several additional advantages to an enzyme solution with little or no water (Wescott & Klibanov, 1994). One advantage is that lipophilic substrates are soluble in organic solvents while being sparingly soluble in water. A second important advantage is that enzymes are often more thermostable when suspended in anhydrous solvents. This allows these enzymes to be used at temperatures as high as 100°C. In addition, since enzymes are usually insoluble in solvents, they are easily separated from the product reaction mixture by filtration. And, finally, solvent choice can often affect the formation of reaction products. Since the thermodynamic equilibrium is dependent on the solvent, one can choose the solvent that yields the desired product.

Biological importance of proteases

Proteases are found in all living organisms and are a vital part of cellular function (Hase & Finkelstein, 1993; Hicks, 1998). They participate, either directly or indirectly, in many diverse functions: cell growth, differentiation, and death, cell nutrition, cell

housekeeping and repair, cell migration and invasion, and fertilization and implantation (Hugli, 1996). In fact, proteolytic cleavage of peptide bonds is one of the most frequent and important enzymatic modifications of proteins (Neurath, 1989). Proteases are both intra- and extracellular (Bauer *et al.*, 1996). Most intracellular proteases are aspartic (viral proteases such as HIV, tissue proteases such as rennin), cysteine (cathepsin B and L), and threonine (proteasome proteases). Extracellular proteases are primarily serine (cell derived proteases such as elastases and blood proteases) and metallo- (matrix metalloproteases (MMP's), collagenases, gelatinases) (Hugli, 1996).

Initially, interest in proteases stemmed from the need for further knowledge about protein digestion in both humans and animals (Neurath, 1989). From this interest, primarily by early biochemists, detailed information about pancreatic and gastric proteases was sought and, in 1948, crystal structures of these proteases with their inhibitors were obtained. Those studies on pancreatic protease structure and function became the basis for our present knowledge about general protease function. Furthermore, it became the prototype for the elucidation of detailed structural and functional information on the many complex proteases that we have encountered since then (Neurath, 1993).

Protease inhibitors have been studied extensively for use in the health industry. One very important set of inhibitors includes those natural and synthetic inhibitors that regulate the activity of thrombin. Thrombin, a serine protease in the chymotrypsin family, is a key participant in the cascade of events that occurs during coagulation through conversion of soluble fibrinogen to insoluble fibrin (Bachand *et al*., 1999; Lombardi *et al*., 1999). Although the role of thrombin in regulation of platelet functions and coagulation has been well established, its many other cellular regulation functions are not well understood. A

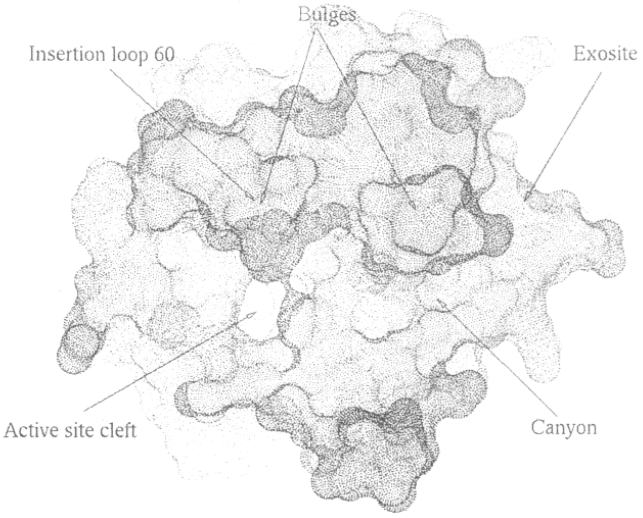
detailed explanation of the structure of thrombin is given Lombardi et al. (1999). An illustration of human α-thrombin, with the major interaction sites indicated, is shown in Figure 2.4a. Thrombin inhibitors are particularly helpful in the treatment of post-stroke and post-cardiac surgery patients, candidates for thrombosis (Hugli, 1996). Recently, some of these other functions, such as the formation of vascular lesions after vascular injury, have been closely studied for the purpose of controlling them through direct inhibition of thrombin and, in particular, the thrombin receptor (Herbert *et al*., 1997). These particular protease inhibitors have been studied for several years and are still being modified so that, in the future, treatment and/or prevention of cardiovascular diseases, such as heart attack and stroke, will be carried out in a much more efficient manner (Frost, 1999).

The most potent inhibitors of thrombin that have been isolated to date are from blood-sucking animals. Among these is the most potent anticoagulant agent known, hirudin (Hugli, 1996), which was isolated from the salivary glands of the leech *Hirudo medicinalis* (Lombardi *et al.*, 1999; Stone & Hofsteenge, 1986). It works by binding to thrombin, thereby preventing the cleavage of fibrinogen (the natural substrate of thrombin) and inhibiting the activation of platelets by thrombin. It has a wide variety of advantages over other anticoagulants, including the ability to inhibit thrombin in either the free or bound state, its long half life (50 min), and its lack of interference with the biosynthesis of clotting factors or other blood enzymes in the coagulation cascade (Lombardi *et al.*, 1999). An illustration of the interaction between hirudin and human α-thrombin is shown in Figure 2.4b. The most notable characteristic of the binding mode is that, unlike other serine protease inhibitors, hirudin makes contact with thrombin over an extended area, including regions far from the active site (Lombardi *et al.*, 1999). Other very potent inhibitors include triabin (Noeske-

Jungblut *et al*., 1995) (isolated from the saliva of the blood-sucking triatomine bug *Triatoma pallidipenis*), rhodniin (Friedrich *et al*., 1993) (isolated from the assassin bug *Rhodnius prolixus*, binds thrombin through a series of unique multiple interactions), and ornithodorin (van de Locht *et al*., 1996) (isolated from a blood sucking soft tick *Ornithodoros moubata*) (Lombardi *et al.*, 1999).

One of the most recent approaches to developing improved inhibitors for therapeutic applications is the use of natural thrombin inhibitors as template structures for the design of new enzymes with increases potency and selectivity (Hugli, 1996; Lombardi *et al.*, 1999). In particular, the hirudin structure has been used extensively in the design of synthetic inhibitors. Some of the first synthetic multi-site inhibitors included hirulogs and hirutonins. The main differences between the two families of inhibitors are in the nature of the linker and in the length of the fibrinogen recognition exosite (FRE) binding element (see Figure2.4b for location of FRE on thrombin). Recently, researchers have begun to develop a new class of synthetic thrombin inhibitors, the hirunorms. This class of inhibitors is being developed to mimic the hirudin mechanism of action: they interact with thrombin at both the active site, in a nonsubstrate mode, and the FRE (Lombardi *et al.*, 1999). Although the number of potent and selective synthetic inhibitors is growing, discovery of new mechanisms for thrombin inhibition is ongoing and, as a result, these newly designed inhibitors are continually being used to create novel therapeutic applications.

a)



b)

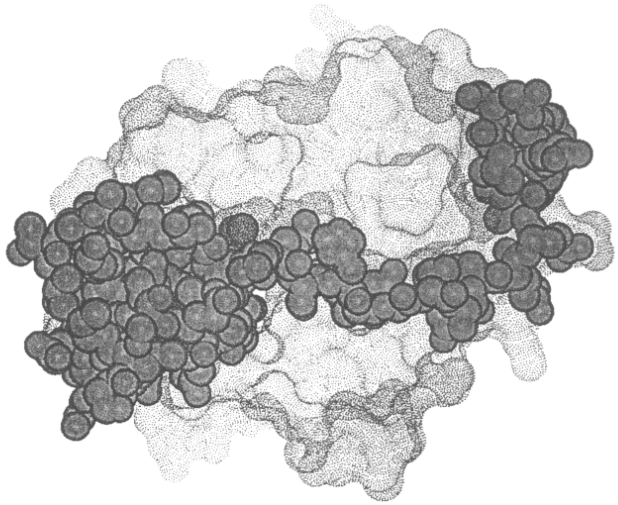


Figure 2.4: a) The surface structure of human α-thrombin with the major interaction sites labeled. The active site cleft contains a catalytic triad (Ser^{195}, His^{57}, Asp^{102}) and substrate binding regions. The insertion loop defines the upper rim of the active site and shields Ser^{195} and His^{57} from solvent. The exosite region, including the fibrinogen recognition exosite (FRE), is involved in the binding of substrates and inhibitors to thrombin. **b)** The interaction mode of hirudin with human α-thrombin. Hirudin binds not only to the active site cleft, but to extended regions far from the active site. Reproduced from Lombardi et al. (1999).

Many recent efforts have focused on the inhibition of proteases that participate in the reproduction of viruses. Researchers originally observed that retroviruses encode a proteolytic activity of their own as early as 1977. However, those in the health industry began to understand, only recently (early 1980's), that proteases play a large part in the reproductive cycle, and ultimately the survival, of nearly every class of virus (Pearl, 1990). Many commonly known examples of important viruses are HIV (human immunodeficiency virus), influenza, herpes, hepatitis and the common cold (human rhinoviruses) (Frost, 1999; Hugli, 1996). In order to impede the life cycle of such viruses, research is focused on pinpointing the precise proteases that participate in the reproduction, assembly, and survival of the viruses. Once specific proteases are found, efforts are directed toward the engineering of corresponding inhibitors with maximum potency and selectivity, in order to wipe out the virus both quickly and without side effects.

The cure for the common cold (human rhinoviruses, HRVs) has been the focus of intense study for the past 40-50 years (Johnson, 1997). The largest obstacles of this research are the high number of microorganisms that are associated with the common cold (about 200 different viruses) and the rapid mutation rates of some of these viruses leading to the presence of drug-resistant strains (Johnson, 1997). As is the case with the drugs used to treat HIV, there are difficulties with drug delivery, expense, and unwanted side effects (Johnson, 1997). Recently, inhibitors for the suppression of proteases that participate in the production of picornaviruses (human rhinovirus, hepatitis A), named 2A and 3C protease inhibitors, have been designed (Wang, 1999). Picornavirus proteases, such as the 2A protease, are often essential for viral particle assembly, can shut off host-cell protein synthesis, and are required as cofactors in viral RNA synthesis (Hugli, 1996). Since this information was found, direct

inhibition of these viral proteases has become an extremely promising avenue for treatment of the common cold and many other viruses.

In addition to their participation in the spread of viral infections, proteases play a large part in the survival of such diseases as asthma, rheumatoid arthritis, cancer, and diseases where programmed cell death (apoptosis) occurs. For all of these diseases, the importance of learning about protease regulation has grown as we realize just how ubiquitous proteases are. For instance, one pharmaceutical company, Corvas International, Inc., is presently attempting to develop a drug that will inhibit urokinase-type plasminogen activator (uPA). This particular protease is involved in the metastasis of some cancers and in angiogenesis (the growth of new blood vessels to help tumors expand and invade normal tissue) (Frost, 1999). In addition, through the use of a variety of studies such as inhibition, links were found between several proteolytic cathepsins, including cathepsins B, D, and L, and cancer cell invasion (Koblinski & Sloane, 1997; Tedone *et al*., 1997). However, because researchers are still in the process of attempting to fully understand the role of these proteases in the progression of cancer, drugs for their inhibition have yet to be put on the market.

III. HYPERTHERMOPHILIC PROTEASES

Strategies for thermostability

One of the most useful properties of proteases from hyperthermophilic microorganisms is their thermostability. In fact, all of the enzymes and, in particular, proteases purified thus far from hyperthermophiles are intrinsically thermostable (Adams *et al*., 1995). This means that their thermostability in the cell is not solely caused by a separate 'thermoprotectant' within the cell. Instead, these thermostable proteases can retain their catalytic activity for long periods at high temperatures and in their pure form in dilute buffer solution (Adams *et al.*, 1995). Furthermore, it is well accepted that resistance to heat is coupled with resistance to many denaturing and chaotropic agents. This relation between different types of stability occurs because the characteristics that create the thermostability are closely related to those that create other types of stability. Specifically, the increase in the net free energy of stabilization that leads to enhanced thermostability is caused by an increase in the interactions that stabilize proteins against denaturation by chemical agents (Daniel *et al*., 1995).

There are three main types of heat-induced mechanisms by which proteases can lose activity: 1) conformational unfolding (denaturation), 2) enhanced autolysis (self-digestion), and 3) other irreversible covalent modifications (i.e. deamidation) (Daniel *et al.*, 1995). These three mechanisms are closely linked and often seem to work together. For example, the third type of mechanism has been shown to occur much more quickly in proteases that are in some type of unfolded or denatured state (mechanism no. 1) than those that are in their native conformation.

Although knowledge about the resistance mechanisms of hyperthermophilic proteases against the three types of heat damage is incomplete, there is an awareness of several basic characteristics that may aid in their enhanced thermostability. Specifically, "stability refers to the maintenance of a defined functional state under extreme conditions" (Jaenicke & Böhm, 1998). One important example of a heat stability mechanism is the difference in the free energy of stabilization values between mesophilic proteases and their hyperthermophilic homologs. The free energy is larger for the more thermostable proteases, however, this difference is often quite small, only in the tens of kJ/mol (Beadle *et al*., 1999; Daniel *et al.*, 1995; Jaenicke & Böhm, 1998). A single salt bridge is able to contribute 4-12 kJ/mol to a proteolytic structure. Furthermore, a small number of additional hydrophobic interactions or hydrogen bonds can also confer an extra degree of stabilization. Therefore, instead of dramatic or obvious structural changes, only slight changes in amino acid sequences or in a few molecular interactions are needed to increase the free energy (Daniel *et al.*, 1995). Examples of the major causes of thermostability by several proteins from hyperthermophilic microorganisms are listed in Table 2.3.

Table 2.3: Strategies of thermal stabilization of selected hyperthermophilic enzymes. Multiple strategies for a single enzyme are listed in order of importance. Reconstructed from Jaenicke et al. (1998).

Protein [Microbial source]	Major cause(s) of thermostability
Citrate synthase [*P. furiosus*]	Increased compactness, enhanced subunit interactions, increased number of intersubunit ion pairs, shortening of loops.
Ferredoxin [*T. maritima*]	Structurally: stabilization of α helices; replacement of conformationally strained residues by glycines; strong docking of the amino-terminal methionine; increase in the number of hydrogen bonds. Thermodynamically: flat ΔG versus temperature profile caused by low ΔC_p of unfolding.
Ferredoxin [*Synechococcus elongatus*]	Extension of the hydrophobic core, a unique hydrophobic patch on the surface β sheet, two unique ion-pair networks.
Glutamate dehydrogenase (GluDH) [*P. furiosus*]	A series of extended ion-pair networks on protein subunit surfaces and ion-pair networks buried at interdomain and intersubunit interfaces, enhanced packing within the inner core, amino acid replacements increasing the hydrophobicity and sidechain branching (Val →Ili).
Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) [*T. maritima*]	Large number of additional salt bridges.
Indole 3-glycerol phosphate synthase [*Sulfolobus solfataricus*]	Large number of additional salt bridges (partly cross-linking adjacent helices), increased helix capping, dipole stabilization, increased hydrophobic interactions, strengthening of chain termini and solvent-exposed loops.
3-isopropylmalate dehydrogenase (IPMDH) [*T. thermophilus*]	Increased number of ion pairs and hydrogen bonds, extended hydrophobic subunit interactions and improved packing of the hydrophobic core, shortened chain termini.
Lactate dehydrogenase (LDH) [*T. maritima*]	Increased number of ion pairs, decreased hydrophobic surface area, increased helicity, less cavity volume.
Phosphoglycerate kinase (PGK) [*T. maritima*]	Increased rigidity by additional pairs, stabilization of helix and loop regions, stabilization by fusion with triosephosphate isomerase (TIM).

Applications for hyperthermophilic proteases

Although it is well known that hyperthermophilic enzymes and proteases are not necessarily more specifically active than their mesophilic counterparts (Adams & Kelly, 1994), the number of applications that make use of their intrinsic thermostability is growing (Daniel *et al.*, 1995). In general, the use of higher temperatures in industrial processes is advantageous for several reasons. First, temperatures above 70°C are normally high enough to kill many pathogenic bacteria and reduce the number of bacteria that can cause contamination of food processes. Higher temperatures also reduce the viscosities, which can often reduce the costs involved in pumping, filtration, and centrifugation. Stable proteases are also likely to allow the use of organic solvents in processes. Finally, since the major application of proteases involves hydrolysis of proteins, rather than peptides, it is advantageous to have high-temperature conditions that will cause some degree of denaturation of the substrate proteins. The one disadvantage that is coupled with the use of high-temperature proteases is that elevated temperatures are essential for maintaining high proteolytic activity.

Presently, few thermophilic or hyperthermophilic proteases are used extensively in industry, largely because it is so difficult to produce them in high quantities. Furthermore, most of the thermostable proteases that are used for industrial purposes are those that are most stable in the range of 30-60°C (Daniel *et al.*, 1995). Two main examples of these include thermolysin, which is used in the synthesis of the sweetening dipeptide aspartame precursor, and a *Thermus* protease (sold as Pretaq), which is used to clean up DNA for amplification in the PCR reaction (Daniel *et al.*, 1995).

There are some proposed uses for hyperthermophilic proteases that may help to improve present processes (Daniel *et al.*, 1995). For example, the ideal meat-tenderizing enzyme would seem to be one that works only during the cooking period, instead of during any type of storage period. A protease that is primarily active during this high temperature period may be useful to avoid the risk of under- or over-tenderization. Another example is in the cleaning of ultrafiltration membranes, particularly in the processing of whey. The use of very high temperature proteases would allow lower viscosities and reduce the need for strong acid or alkali, thereby lengthening the useful life of the membranes. A third example is use of proteases in laundry detergents. Much higher temperatures are desirable for laundry detergents used in hospitals and similar industries. Therefore, proteases that are primarily stable at these much higher temperatures are desirable for these types of detergents.

To help take advantage of the stable properties of many hyperthermophilic proteases, there are now many attempts to do more complete biochemical and biophysical characterization studies on these proteases. There are, at this point, a long list of proteases from hyperthermophiles, including the tricorn protease and pyrolysin. The focus in this research is on two proteases in particular. Both are from the model hyperthermophilic archaeon, *Pyrococcus furiosus*, and include *Pyrococcus furiosus* Protease I (PfpI) and the proteasome.

Examples of hyperthermophilic proteases

Tricorn protease

The tricorn protease (TRI) was isolated from the archaeon *Thermoplasma acidophilum* in 1996 (Tamura *et al*., 1996a). It has also been found in *Sulfolobus acidocaldarius* (Lupas *et al*., 1997). The purified form of TRI from *T. acidophilum* migrates at 720-730 kDa in gel filtration chromatography and is composed of a single 120-kDa polypeptide when separated by SDS-polyacrylamide gel electrophoresis (SDS-PAGE). TRI is named for the unusual three-dimensional shape it creates: similar to a tricorn, which is a hat with the brim turned up on three sides. Specifically, one full TRI complex contains six identical 120-kDa subunits that assemble into a trimer of dimers. The set of three dimers then form a toroid that contains a channel along its threefold axis (Schneider & Hartl, 1996). There is evidence that, *in vivo*, the hexameric toroids further assemble into a higher-order assembly, namely a 55-nm icosahedral capsid. Each capsid contains 20 copies of the tricorn toroid, resulting in the creation of a 14.6-MDa complex (Walz *et al*., 1997; Yao & Cohen, 1999). It is believed that the capsid structure, which has large void volumes in the shell of its structure, may be an organization center for the positioning of interacting factors (Tamura *et al*., 1998).

While TRI is similar to the proteasome in the fact that it appears to be a self-compartmentalized protease, there are several distinct differences between the TRI complex and the proteasome. The three TRI dimers enclose a channel that traverses the hexamer. This channel has 2.6-nm openings that lead into a cavity that is 10 nm across and 4.3 nm high. Therefore, the width of the channel is much larger than the 5.3-nm size of the *T. acidophilum* proteasome chamber. In addition, the overall structure is composed of only one

type of subunit in TRI, while the proteasome is composed of two types of subunits (α and β). The activities of the TRI and proteasome are also different. TRI has trypsin-like and very high chymotrypsin-like cleavage specificities, while the archaeal 20S proteasome has primarily chymotrypsin-like activity (Yao & Cohen, 1999). Furthermore, mutational studies on TRI suggest that a serine residue is involved in the active center and inhibition studies suggest that a histidine is involved in active site formation (Tamura *et al.*, 1996a). However, a threonine residue appears to be the center of the proteasomal active site.

Like the proteasome, TRI cooperates with other polypeptides to achieve full proteolytic activity. The activating polypeptides for TRI include at least three different low molecular mass components: F1, F2, and F3. F1, a 33-kDa polypeptide, was found to have homology to bacterial proline iminopeptidases (PIPs) and some studies have shown that it also has PIP activity (Tamura *et al*., 1996b). In addition, it is capable of hydrolyzing a wide variety of substrate peptides. F1 has several homologs in bacteria and eukaryotes that are all members of the α/β hydrolase superfamily (Tamura *et al.*, 1998; Tamura *et al.*, 1996a; Tamura *et al.*, 1996b). There is preliminary evidence that F2, which is 89 kDa in size, may be an amino-peptidase (Tamura *et al.*, 1996a). F3 is both the same size as F2 and is closely related to F2 (56.3% sequence identity). Homologs to F2 and F3 have been found in *S. cerevisiae*. These yeast homologs, Aap1p and Ape2p, have been identified as metallo-aminopeptidases. Although all three of the interacting factors can cleave very short peptides (2-4 residues) efficiently, they are all needed for separate functions. F1 is necessary for the release of proline residue, F2 is needed for the release of basic amino residues, and F3 is needed for the release of acidic residues (Tamura *et al.*, 1998).

Unlike the 26S proteasome, but like the 20S proteasome, there is thus far no evidence for an ATP requirement by TRI. Compared to the 26S proteasome, the central cavity of TRI, which is relatively large, appears to be more easily accessed by substrates. Furthermore, less extensive unfolding of substrates may be required for cleavage by TRI. At this time, the proposed function of TRI is as a participant in an entire sequence of events. It is believed that TRI actually acts as a scavenger for oligopeptides generated by the proteasome or other ATP-dependent proteases. Once TRI breaks down the oligopeptides (e.g., 6- to 12-mers) into 2-4 residue pieces, the aminopeptidase-interacting factors (F1, F2, F3) hydrolyze them further into free amino acids. The proposed pathway for degradation of proteins is illustrated in Figure 2.5.

The quaternary structure of TRI is similar to that of the Gal6 or bleomycin hydrolase from *S. cerevisiae*. Both proteases are hexamers with a central channel. However, the sequence of the *tri* gene, which encodes for TRI, is not related to the gene encoding for Gal6 (Schneider & Hartl, 1996; Tamura *et al.*, 1996a). In fact, only one TRI homolog has been found in the databases to date: a 37.1% sequence identity in the genome of the archeaon *Sulfolobus solfataricus*. A high-molecular weight (700 kDa) protease with similar enzymatic properties to TRI has also been isolated from *P. furiosus*. However, the *P. furiosus* protease only has slight sequence similarity to TRI in the amino-terminal half (Tamura *et al.*, 1998). An illustration of TRI, in comparison to Gal6, the 20S proteasome, and ClpP is shown in Figure 2.6.

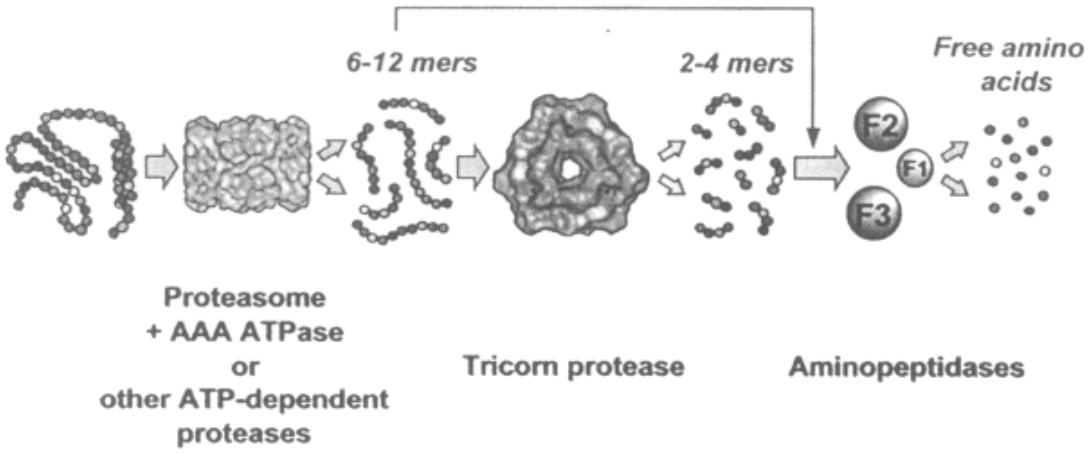


Figure 2.5: Proposed proteolytic pathway in *Thermoplasma acidophilum*. Unfolded target proteins are initially broken down into oligopeptides (6-12 mers) by the proteasome or other ATP-dependent proteases. Secondly, the tricorn protease breaks the oligopeptides into 2-4 residue fragments. Finally, the aminopeptidase-interacting factors (F1, F2, F3) break the 2-4 mers into free amino acids. Reproduced from Tamura et al. (1998).

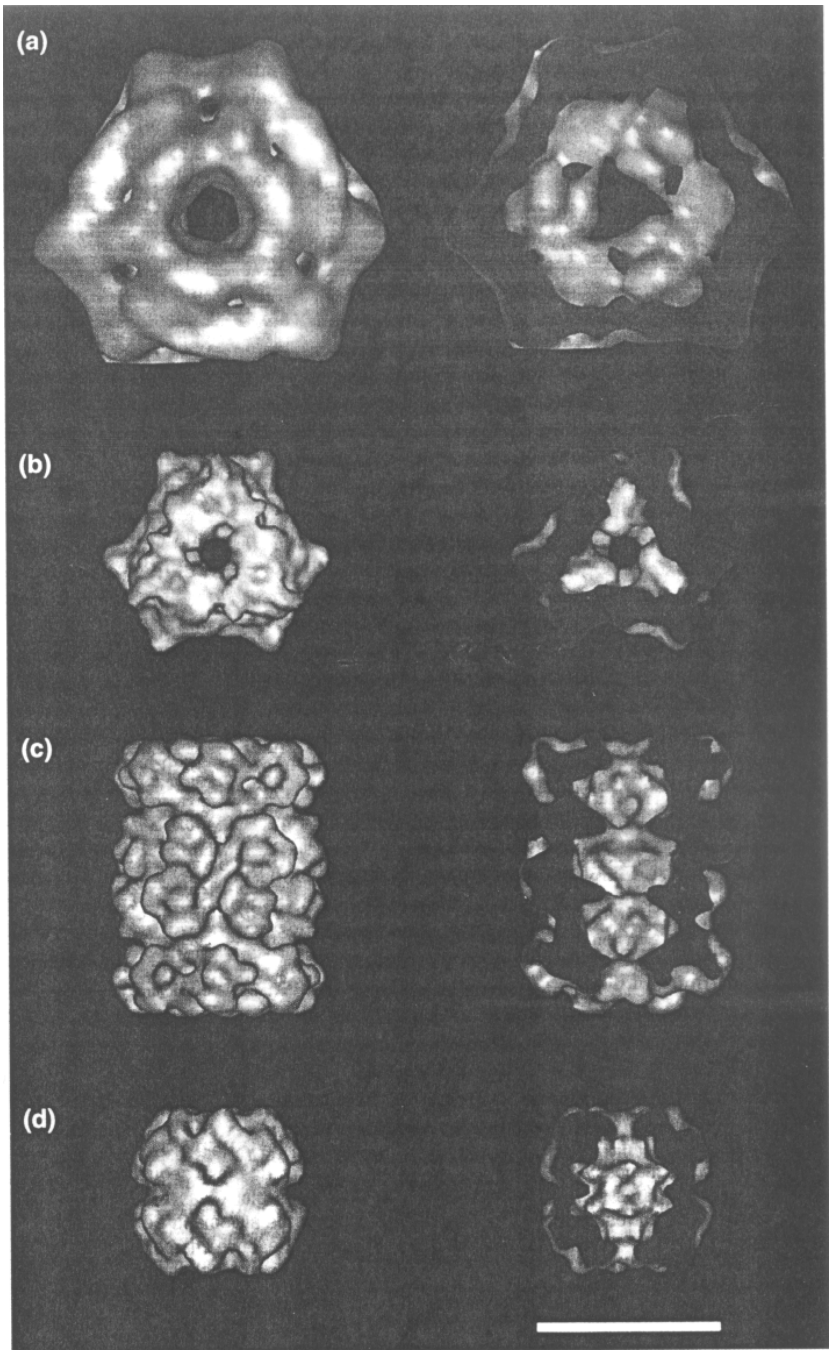


Figure 2.6: Space-filling views of several self-compartmentalizing proteases (left) and cross-sections showing the inner cavities (right). (a) tricorn, (b) Gal6 (bleomycin hydrolase), (c) 20S proteasome, and (d) ClpP. The tricorn protease is shown at a resolution of 2.5 nm and the other structures are shown at a resolution of 1.2 nm. The bar represents 10 nm. Reproduced from Lupas et al. (1997).

Pyrolysin

The protease pyrolysin was initially isolated from *P. furiosus* in 1990 and was given its name as a reference to the Greek translation for 'pyro' (fire) and the Latin translation for 'lysin' (to cleave) (Eggen *et al*., 1990). From growth studies of *P. furiosus*, it was found that pyrolysin is largely a cell envelope-associated protease. Specifically, 75% of pyrolysin activity was associated with the cell envelope in initial studies. It is active in a wide range of pH values, exhibiting 50% of its maximal activity between pH values of 6.5 and 10.5. In addition, it can retain activity up to a temperature of 115°C and has a half-life of more than 96 hours at 80°C and 4 hours at 100°C. From sequence and inhibitor studies, it was found to have the highest homology to the subgroup of the subtilisin-like serine proteases (Eggen *et al.*, 1990; Voorhorst *et al*., 1996). Subtilisin-like proteases are characterized by the Asp-His-Ser sequential order of residues in the active site and the high degree of homology in the amino acid sequence around the residues that form the catalytic triad (Voorhorst *et al.*, 1996).

From purification of pyrolysin, it has been found to occur in two different forms: high molecular weight (HMW) and low molecular weight (LMW) forms (Eggen *et al.*, 1990; Voorhorst *et al.*, 1996). The two forms of pyrolysin, which are 130 and 150 kDa in size, have identical amino-terminal sequences and are both glycosylated. Furthermore, a long incubation period of purified HMW pyrolysin at 95°C yields a decrease in the quantity of HMW and a simultaneous increase in the amount of LMW. With the appearance of increased LMW pyrolysin, there is also an appearance of a new proteolytic activity. This result suggests that the LMW form may be a processing product of the HMW form. It has been proposed that the HMW becomes the LMW pyrolysin through autoproteolytic removal of a carboxyl-terminal part of the HMW form. This type of activity has been proposed for

other serine proteases from archaeal microorganisms (i.e. *Thermococcus stetteri*) and is, in fact, a common feature of serine proteases (Voorhorst *et al.*, 1996).

The deduced amino acid sequence of pyrolysin shows the highest degree of homology with the tripeptidyl peptidases II (TPPs) from eukaryotic organisms. Specifically, there is 28-32% identity in the catalytic domains of pyrolysin and TPPs from humans, mice, *Drosophila melanogaster*, and the nematode *Caenorhabditis elegans*. The eukaryotic TPPs, unlike all other known subtilisin-like serine proteases, are intracellular proteases with exopeptidase activity. However, in substrate specificity studies on pyrolysin, it was found that this protease does not exhibit the exopeptidase activity of its closest homolog. Instead, it exhibits endopeptidase activity, as observed during the incubation of pyrolysin with several types of caseins and casein fragments. All of these peptides were degraded into distinct intermediates and, finally, were completely degraded by pyrolysin (Voorhorst *et al.*, 1996). Conformational characteristics of the catalytic domain and the substrate binding regions of pyrolysin have been modeled by Voorhorst et al. (1997).

Pyrococcus furiosus protease I (PfpI)

PfpI was initially observed and named S66 in 1990, following an unusual purification procedure. After heating *P. furiosus* cell extracts at 98°C for 24 hrs in the presence of 1% w/v SDS, Blumentals et al. (1990) found that one of the few remaining proteins had proteolytic activity. Once this proteolytically active sample was run on a denaturing gel, it exhibited a size of approximately 66 kDa. This led to its initial designation, S66. However, by 1996, it was understood that there were at least two functional forms of this protease and that, because of its insensitivity to SDS, the 66-kDa size observed on SDS-PAGE gels did

not accurately represent this protease (Halio *et al*., 1996; Hicks, 1998). It was subsequently re-named PfpI (*Pyrococcus furiosus* protease I).

PfpI is composed of 18.8-kDa subunits that, in isolated form, do not appear to be active (Halio *et al.*, 1996). However, during *in vitro* studies, it was found that these subunits can assemble into trimeric, hexameric, and higher forms of three (Halio *et al.*, 1996; Hicks, 1998). It has been observed in structures with sizes of up to 275 kDa, which is the approximate size of a 15-mer form of the subunit (Hicks & Kelly, 1998). Du et al. (2000) reported the crystal structure of the *P. horikoshii* form and concluded that it was a homo-hexameric ring of subunits. It is most likely the case that these hexamers then stack to form the dodocameric and higher structures.

A trimeric structure of PfpI is the smallest group of subunits that exhibits proteolytic activity. This activity was found to be chymotrypsin- and trypsin-like (Blumentals *et al.*, 1990; Halio *et al*., 1997). This characteristic was determined from the fact that PfpI is specific toward both basic and bulky hydrophobic P_1 amino acid residues in peptide substrates. In particular, it is most active toward the synthetic substrate N-succinyl-alanine-alanine-phenylalanine-7-amido-4-methylcoumarin (AAF-MCA) (Halio *et al.*, 1997). Furthermore, large proteins, including azocasein and gelatin, are degraded by PfpI. Specific cleavage sites for these types of degradation reactions have not yet been identified (Hicks, 1998).

The trimeric and hexameric forms of PfpI were found to have higher temperature optima (75-90°C) compared to the dodocamer form (75-80°C), which is the only higher form (above hexameric) of PfpI studied (Hicks & Kelly, 1998). However, toward the MCA-linked peptides AAF, AFK, and LLVY, the smaller forms exhibited lower specific activity over the

temperature range 75-100°C. Specifically, the dodecamer of PfpI was three times more active toward the chymotrypin and trypsin substrates AAF- and AFK-MCA, respectively. As temperatures were increased to 100°C, the specific activity of the larger form decreased more rapidly, reflecting a possible degradation into the smaller forms. The reverse activity of PfpI, namely the assembly of trimeric forms into larger forms, was observed during incubation of the protein at temperatures suboptimal for catalysis (Halio *et al.*, 1996). Finally, the half-life values for all forms of PfpI were extremely variable (ranging from <30 minutes to >24 hours for temperatures 75-98°C), depending on temperature and PfpI form.

The calculated isoelectric point (pI) from amino acid sequence data is 6.1 for the 18.8-kDa subunit of PfpI. Experimentally determined pI values for the trimer and hexamer are 6.1 and 3.7-3.9, respectively. However, if all asparagine and glutamine residues are changed to their corresponding acidic residues, the pI of the monomer is calculated to be 4.8. This information suggests that, at high temperatures, it is possible that PfpI is deamidated, especially the trimeric form since more surface area is left unprotected. This may help to explain differences in the measured and predicted pI values (Halio *et al.*, 1997).

Analysis of the *pfpI* gene, which encodes for PfpI, did not show any structural motifs (Halio *et al.*, 1996). Many homologous protein sequences have been found, however, and there appears to be sequence consensus near a serine, a histidine, and two aspartate residues. This information, along with results from preliminary inhibitor studies, suggests that PfpI could be a serine protease. Furthermore, PfpI appears to have homologs in all three domains of life, including *Archaeoglobus fulgidus*, *E. coli*, and *Arabidopsis thaliana*, and its sequence is well conserved. Unfortunately, like PfpI, all of the putative homologs have unknown functions in their respective cells. Table 2.4 lists the proteins that are most similar to PfpI

and specifies the degree of similarity each protein sequence has in comparison to the amino acid sequence of PfpI.

Table 2.4: Top 14 proteins or putative proteins with similarity to PfpI.
Computations were performed by BLAST 2.0 Program, National Center for Biotechnology Information Web Site: www.ncbi.nlm.nih.gov.

Organism (protein)	**% Identity (amino acid)**	**% Similarity (amino acid)**	**# Gaps (%)**	**# aa compared / length of protein**
Pyrococcus horikoshii (protease I)	150/166 (90%)	160/166 (96%)	0/166 (0%)	166/166
Pyrococcus abyssi (protease I)	142/166 (85%)	159/166 (96%)	0/166 (0%)	166/166
Aeropyrum pernix (protease I)	101/165 (61%)	129/165 (78%)	0/165 (0%)	165/180
Archaeoglobus fulgidus (protease I)	93/166 (56%)	124/166 (75%)	0/166 (0%)	166/168
Aquifex aeolicus (protease I)	69/165 (42%)	98/165 (59%)	0/167 (0%)	165/167
Bacillus subtilis (YfkM)	75/168 (45%)	105/168 (62%)	3/168 (2%)	168/172
Escherichia coli (hypothetical protein)	79/167 (47%)	103/167 (62%)	3/167 (2%)	167/186
Bacillus subtilis (hypothetical protein)	63/150 (42%)	89/150 (59%)	1/150 (0%)	150/154
Staphylococcus aureus (hypothetical protein)	71/168 (42%)	96/168 (57%)	3/168 (2%)	168/171
Streptomyces coelicolor (putative protease)	71/172 (41%)	99/172 (58%)	10/172 (6%)	172/180
Burkholderia cepacia (unknown)	61/183 (33%)	96/183 (52%)	19/183 (10%)	183/197
Methanococcus jannaschii (intracellular protease I)	51/152 (33%)	86/152 (57%)	3/152 (2%)	152/205
Arabidopsis thaliana (unknown)	50/174 (29%)	88/174 (51%)	18/174 (10%)	174/398
Borrelia burgdorferi (biosynthesis protein)	47/169 (28%)	85/169 (50%)	5/169 (3%)	169/184

Halio et al. (1996) previously expressed the gene for PfpI, *pfpI*, in *Escherichia coli* using the T7 promoter and infection with λ-phage CE6. Unfortunately, with this expression system, expression levels were extremely low. Specifically, they were able to obtain approximately 1 mg protein per 100 ml culture when a histidine tag was used to make a fusion protein. But even with the success of obtaining this small volume of protein, the recombinant PfpI was significantly less stable than its native form. Further attempts were made to obtain recombinant forms of this protease through the use of different expression systems and hosts, but have been unsuccessful to this point (Hicks, 1998). In general, vector constructs are often unstable and toxicity has consistently been a problem during attempts to produce recombinant forms of PfpI. The next two chapters concentrate on the most recent efforts to elucidate the role of PfpI in the cell. Furthermore, Chapters 5 and 6 concentrate on the proteasome, which may be an integral part of the activity by PfpI.

IV. REFERENCES

Adams, M. W. W. & Kelly, R. M. (1994). Thermostability and thermoactivity of enzymes from hyperthermophilic Archaea. *Bioorg Med Chem* **2**(7), 659-667.

Adams, M. W. W., Perler, F. B. & Kelly, R. M. (1995). Extremozymes: expanding the limits of biocatalysis. *Biotechnology* **13**(7), 662-668.

Arendt, C. S. & Hochstrasser, M. (1997). Identification of the yeast 20S proteasome catalytic centers and subunit interactions required for active-site formation. *Proc Natl Acad Sci U S A* **94**(14), 7156-7161.

Bachand, B., DiMaio, J. & Siddiqui, M. A. (1999). Synthesis and Structure-Activity Relationship of Potent Bicyclic Lactam Thrombin Inhibitors. *Bioorg Med Chem Lett* **9**(7), 913-918.

Barrett, A. J. (1994). Classification of Peptides. In *Proteolytic Enzymes: Serine and Cysteine Peptidases* First edit. (Barrett, A. J., ed.), Vol. 244, pp. 1-15. Academic Press, Inc., San Diego, CA.

Barrett, A. J. & Rawlings, N. D. (1993). The Many Evolutionary Lines of Peptidases. In *Innovations in Proteases and their Inhibitors* First edit. (Avilés, F. X., ed.), pp. 13-30. Walter de Gruyter & Co., Berlin.

Bauer, M. W., Halio, S. B. & Kelly, R. M. (1996). Proteases and Glycosyl Hydrolases from Hyperthermophilic Microorganisms. *Adv Protein Chem* **48**, 271-310.

Beadle, B. M., Baase, W. A., Wilson, D. B., Gilkes, N. R. & Shoichet, B. K. (1999). Comparing the Thermodynamic Stabilities of a Related Thermophilic and Mesophilic Enzyme. *Biochemistry* **38**(8), 2570-2576.

Blumentals, I. I., Robinson, A. S. & Kelly, R. M. (1990). Characterization of sodium dodecyl sulfate-resistant proteolytic activity in the hyperthermophilic archaebacterium *Pyrococcus furiosus*. *Appl Environ Microbiol* **56**(7), 1992-1998.

Bode, W. & Huber, R. (1993). Structural Basis of the Proteinase-Protein Inhibitor Interaction. In *Innovations in Proteases and their Inhibitors* First edit. (Avilés, F. X., ed.), pp. 81-122. Walter de Gruyter & Co., Berlin.

Daniel, R. M., Toogood, H. S. & Bergquist, P. L. (1995). Thermostable proteases. *Biotechnol Genet Eng* **13**, 51-100.

Du, X., Choi, I. G., Kim, R., Wang, W., Jancarik, J., Yokota, H. & Kim, S. H. (2000). Crystal structure of an intracellular protease from Pyrococcus horikoshii at 2-A resolution. *Proc Natl Acad Sci U S A* **97**(26), 14079-84.

Dunn, B. M. (1989). Determination of Protease Mechanism. In *Proteolytic enzymes: a practical approach* (Beynon, R. J. & Bond, J. S., eds.), pp. 57-81. IRL Press, Oxford.

Eggen, R., Geerling, A., Watts, J. & de Vos, W. M. (1990). Characterization of pyrolysin, a hyperthermoactive serine protease from the archaebacterium *Pyrococcus furiosus*. *FEMS Microbiol Lett* **71**(1-2), 17-20.

Friedrich, T., Kroger, B., Bialojan, S., Lemaire, H. G., Hoffken, H. W., Reuschenbach, P., Otte, M. & Dodt, J. (1993). A Kazal-Type Inhibitor with Thrombin Specificity from *Rhodnius Prolixus*. *J Biol Chem* **268**(22), 16216-16222.

Frost, P. (1999). Battling Disease by Blunting "Mother Nature's Swiss Army Knife". *Chemistry (Am Chem Soc)*(Spring), 15-19.

Groll, M., Ditzel, L., Löwe, J., Stock, D., Bochtler, M., Bartunik, H. D. & Huber, R. (1997). Structure of 20S proteasome from yeast at 2.4 Å resolution. *Nature* **386**(6624), 463-471.

Halio, S. B., Bauer, M. W., Mukund, S., Adams, M. W. W. & Kelly, R. M. (1997). Purification and Characterization of Two Functional Forms of Intracellular Protease PfpI from the Hyperthermophilic Archaeon *Pyrococcus furiosus*. *Appl Environ Microbiol* **63**(1), 289-295.

Halio, S. B., Blumentals, I. I., Short, S. A., Merrill, B. M. & Kelly, R. M. (1996). Sequence, Expression in *Escherichia coli*, and Analysis of the Gene Encoding a Novel Intracellular Protease (PfpI) from the Hyperthermophilic Archaeon *Pyrococcus furiosus*. *J Bacteriol* **178**(9), 2605-2612.

Hase, C. C. & Finkelstein, R. A. (1993). Bacterial Extracellular Zinc-Containing Metalloproteases. *Micrbiol Rev* **57**(4), 823-837.

Headon, D. R. & Walsh, G. (1994). The Industrial Production of Enzymes. *Biotechnol Adv* **12**(4), 635-646.

Herbert, J. M., Guy, A. F., Lamarche, I., Mares, A. M., Savi, P. & Dol, F. (1997). Intimal Hyperplasia Following Vascular Injury Is Not Inhibited by an Antisense Thrombin Receptor Oligodeoxynucleotide. *J Cell Physiol* **170**(2), 106-114.

Hicks, P. M. (1998). Genetic, Biochemical, and Biophysical Characteristics of Intracellular Proteases from the Hyperthermophilic Archaeon *Pyrococcus furiosus* and the Hyperthermophilic Bacterium *Thermotoga maritima*. PhD dissertation, North Carolina State University.

North Carolina State University. (1998). Subunit (18.8 kDa) of protease PfpI from the hyperthermophile *Pyrococcus furiosus* self-assembles into active trimeric, hexameric and larger forms in vitro. Hicks, P. M. & Kelly, R. M. July 23.

Hugli, T. E. (1996). Protease inhibitors: novel therapeutic application and development. *Trends Biotechnol* **14**(11), 409-412.

Jaenicke, R. & Böhm, G. (1998). The stability of proteins in extreme environments. *Curr Opin Struct Biol* **8**(6), 738-748.

Johnson, S. L. (1997). Problems and prospects of developing effective therapy for common colds. *Trends Microbiol* **5**(2), 58-63.

Kasche, V. (1989). Proteases in peptide synthesis. In *Proteolytic enzymes: a practical approach* (Beynon, R. J. & Bond, J. S., eds.), pp. 125-143. IRL Press, Oxford.

Keil, B. (1992). *Specificity of Proteolysis*. First edit, Springer-Verlag, Berlin.

Koblinski, J. E. & Sloane, B. F. (1997). Is Altered Localization of Cathepsin B Causally Related to Malignant Progression? In *Medical Aspects of Proteases and Protease Inhibitors* First edit. (Katunuma, N., Kido, H., Fritz, H. & Travis, J., eds.), Vol. 15, pp. 185-194. 15 vols. IOS Press, Burke, VA.

Lombardi, A., De Simone, G., Galdiero, S., Staiano, N., Nastri, F. & Pavone, V. (1999). From Natural to Synthetic Multisite Thrombin Inhibitors. *Biopolymers (Peptide Science)* **51**(1), 19-39.

Lupas, A., Flanagan, J. M., Tamura, T. & Baumeister, W. (1997). Self-compartmentalizing proteases. *Trends Biochem Sci* **22**(10), 399-404.

Maupin-Furlow, J. A. & Ferry, J. G. (1995). A Proteasome from the Methanogenic Archaeon *Methanosarcina thermophila*. *J Biol Chem* **270**(48), 28617-28622.

Moran, L. A., Scrimgeour, K. G., Horton, H. R., Ochs, R. S. & Rawn, J. D. (1994). *Biochemistry*. Second edit, Prentice Hall, Inc., Upper Saddle River.

Neurath, H. (1989). The diversity of proteolytic enzymes. In *Proteolytic enzymes: a practical approach* (Beynon, R. J. & Bond, J. S., eds.), pp. 1-13. IRL Press, Oxford.

Neurath, H. (1993). The Regulation of Protease Action: An Overview. In *Innovations in Proteases and their Inhibitors* First edit. (Avilés, F. X., ed.), pp. 3-12. Walter de Gruyter & Co., Berlin.

Niamke, S., Sine, J. P., Guionie, O. & Colas, B. (1999). A novel endopeptidase with a strict specificity for threonine residues at the P1' position. *Biochem Biophys Res Comm* **256**(2), 307-312.

Noeske-Jungblut, C., Haendler, B., Donner, P., Alagon, A., Possani, L. & Schleuning, W. D. (1995). Triabin, a Highly Potent Exosite Inhibitor of Thrombin. *J Biol Chem* **270**(48), 28629-28634.

Novo Nordisk. (1994). Enzyme Business. Nordisk, N. January.

Pearl, L. H. (1990). Introduction. In *Retroviral Proteases: Maturation and Morphogenesis* First edit. (Pearl, L. H., ed.), pp. 1-3. Macmillan Press Ltd, New York, NY.

Polgár, L. (1990). *Mechanisms of Protease Action*. Second edit, CRC Press, Inc., Boca Raton.

Rao, M. B., Tanksale, A. M., Ghatge, M. S. & Deshpande, V. V. (1998). Molecular and Biotechnological Aspects of Microbial Proteases. *Microbiol Mol Biol Rev* **62**(3), 597-635.

Rawlings, N. D. & Barrett, A. J. (1994a). Families of Cysteine Peptidases. In *Proteolytic Enzymes: Serine and Cysteine Peptidases* First edit. (Barrett, A. J., ed.), Vol. 244, pp. 461-486. Academic Press, Inc., San Diego, CA.

Rawlings, N. D. & Barrett, A. J. (1994b). Families of Serine Peptidases. In *Proteolytic Enzymes: Serine and Cysteine Peptidases* First edit. (Barrett, A. J., ed.), Vol. 244, pp. 19-61. Academic Press, Inc., San Diego, CA.

Rawlings, N. D. & Barrett, A. J. (1995a). Evolutionary Families of Metallopeptidases. In *Proteolytic Enzymes: Aspartic and Metallo Peptidases* First edit. (Barrett, A. J., ed.), Vol. 248, pp. 183-228. Academic Press, Inc., San Diego, CA.

Rawlings, N. D. & Barrett, A. J. (1995b). Families of Aspartic Peptidases, and Those of Unknown Catalytic Mechanism. In *Proteolytic Enzymes: Aspartic and Metallo Peptidases* First edit. (Barrett, A. J., ed.), Vol. 248, pp. 105-120. Academic Press, Inc., San Diego, CA.

Salvesen, G. & Nagase, H. (1989). Inhibition of proteolytic enzymes. In *Proteolytic enzymes: a practical approach* (Beynon, R. J. & Bond, J. S., eds.), pp. 83-104. IRL Press, Oxford.

Schneider, C. & Hartl, F. U. (1996). Hats Off to the Tricorn Protease. *Science* **274**(22 November), 1323-1324.

Stone, S. R. & Hofsteenge, J. (1986). Kinetics of the Inhibition of Thrombin by Hirudin. *Biochemistry* **25**(16), 4622-4628.

Tamura, N., Lottspeich, F., Baumeister, W. & Tamura, T. (1998). The Role of Tricorn Protease and Its Aminopeptidase-Interacting Factors in Cellular Protein Degradation. *Cell* **95**(5), 637-648.

Tamura, T., Tamura, N., Cejka, Z., Hegerl, R., Lottspeich, F. & Baumeister, W. (1996a). Tricorn Protease - The Core of a Modular Proteolytic System. *Science* **274**(22 November), 1385-1389.

Tamura, T., Tamura, N., Lottspeich, F. & Baumeister, W. (1996b). Tricorn protease (TRI) interacting factor 1 from *Thermoplasma acidophilum* is a proline iminopeptidase. *FEBS Lett* **398**(1), 101-105.

Tedone, T., Correale, M., Barbarossa, G., Casavola, V., Paradiso, A. & Reshkin, S. J. (1997). Release of the aspartyl protease cathepsin D is associated with and facilitates human breast cancer cell invasion. *FASEB J* **11**(10), 785-792.

Tsilikounas, E., Tao, T., Gutheil, W. G. & Bachovchin, W. W. (1996). N-15 and H-1 NMR Spectroscopy of the Catalytic Histidines in Chloromethyl Ketone-Inhibited Complexes of Serine Proteases. *Biochemistry* **35**(7), 2437-2444.

van de Locht, A., Stubbs, M. T., Bode, W., Friedrich, T., Bollschweiler, C., Höffken, W. & Huber, R. (1996). The ornithodorin-thrombin crystal structure, a key to the TAP enigma? *EMBO J* **15**(22), 6011-6017.

von der Osten, C., Branner, S., Hastrup, S., Hedegaard, L., Rasmussen, M. D., Bisgård-Frantzen, H., Carlsen, S. & Mikkelsen, J. M. (1993). Protein engineering of subtilisins to improve stability in detergent formulations. *J Biotechnol* **28**(1), 55-68.

Voorhorst, W. G. B., Eggen, R. I. L., Geerling, A. C. M., Platteeuw, C., Siezen, R. J. & de Vos, W. M. (1996). Isolation and Characterization of the Hyperthermostable Serine Protease, Pyrolysin, and Its Gene from the Hyperthermophilic Archaeon *Pyrococcus furiosus*. *J Biol Chem* **271**(34), 20426-20431.

Voorhorst, W. G. B., Warner, A., de Vos, W. M. & Siezen, R. J. (1997). Homology modelling of two subtilisin-like serine proteases from the hyperthermophilic archaea *Pyrococcus furiosus* and *Thermococcus stetteri*. *Protein Eng* **10**(8), 905-914.

Walz, J., Tamura, T., Tamura, N., Grimm, R., Baumeister, W. & Koster, A. J. (1997). Tricorn Protease Exists as an Icosahedral Supermolecule *In Vivo*. *Mol Cell* **1**(1), 59-65.

Wang, Q. M. (1999). Protease inhibitors as potential antiviral agents for the treatment of picornaviral infections. *Prog Drug Res* **52**, 197-219.

Wescott, C. R. & Klibanov, A. M. (1994). The solvent dependence of enzyme specificity. *Biochim Biophys Acta* **1206**(1), 1-9.

Wright, H. T. (1996). The structural puzzle of how serpin serine proteinase inhibitors work. *BioEssays* **18**(6), 453-464.

Yao, T. & Cohen, R. E. (1999). Giant proteases: Beyond the proteasome. *Curr Biol* **9**(15), R551-R553.

CHAPTER 3:

Protease I (PfpI) from *Pyrococcus furiosus*

Lara S. Chang, Paula M. Hicks, and Robert M. Kelly*

Department of Chemical Engineering
North Carolina State University
Raleigh, NC 27695-7905

Published in:

Methods of Enzymology (2001), Volume 330, pp. 403-413:

Robert M. Kelly and Michael W. W. Adams, Editors

*Address correspondence to: **Robert M. Kelly**
Department of Chemical Engineering
North Carolina State University
Raleigh, NC 27695-7905
Phone: (919) 515-6396
Fax: (919) 515-3465
email: rmkelly@eos.ncsu.edu

Introduction

Pyrococcus furiosus is a hyperthermophilic archaeon from the order Thermococcales that is capable of growth on a variety of proteinaceous and carbohydrate-containing substrates.[1] Analysis of gelatin-containing SDS-PAGE gels indicate that at least 11 endoproteinases are active in the cell extracts of this organism[2,3] and the following proteases have been characterized: protease I (PfpI),[2,4,5,6,7] pyrolysin,[8,9] proteasome,[10] prolyl oligopeptidase,[11,12] and proline dipeptidase.[13] Blumentals et al.[2] found that when *P. furiosus* cell extracts were heated at 98°C for 24 hours in the presence of 1% w/v SDS, one of the few remaining proteins had proteolytic activity. This proteolytically active sample, which was initially named S66, was approximately 66 kDa in size as determined by SDS-PAGE. However, at least two functional forms of this protease exist and the smallest functional form is a trimer of 19-kDa subunits. Furthermore, it is now recognized that, because of its insensitivity to SDS, the 66-kDa size observed on SDS-PAGE gels does not accurately represent the true molecular mass of this protease.[4,7] It was subsequently re-named PfpI (*Pyrococcus furiosus* protease I). Figure 1 illustrates the presence of at least two forms of PfpI observed after heat treatment for 24 hours in the presence of SDS. The smallest functional form of the protease (a homotrimer) is the predominant structure after extended incubation under denaturing conditions.

The gene encoding the 19-kDa subunit of PfpI has homologs in nearly every organism and cell examined to date, ranging from *Escherichia coli* to *Homo sapiens*; this ubiquity and evolutionary conservation indicates that it may play a fundamental physiological role.[7] Efforts to study this issue have been exacerbated by difficulties encountered in obtaining significant amounts of a particular assembly of PfpI in either a

native or recombinant form. Native PfpI undergoes autoproteolysis and/or disassembly during direct purification from *P. furiosus* biomass, and exists in multiple (single- to multi-subunit) forms *in vitro*.[5,6,7,14] The production of a recombinant form of PfpI is also problematic due to its toxicity toward mesophilic hosts.[7] Several methods that have been used to purify PfpI directly from *P. furiosus* cell extracts are described here, together with an assay to detect proteolytic activity, a procedure to determine its molecular mass, and approaches to minimize PfpI-catalyzed proteolysis of other *P. furious* proteins.

Methods

Biochemical assay

PfpI is specific toward basic and bulky hydrophobic P_1[15] amino acid residues in peptide substrates. In general, peptidase activity (proteolysis at the C- or N-terminus of a peptide chain) of PfpI is detected by the release of the fluorescent chemical 7-amido-4-methylcoumarin (MCA) from the carboxyl terminus of amino-terminally blocked substrates (Sigma Chemical Company, St. Louis, MO) using endpoint assays in microtiter plates. It is most active with the chymotrypsin-like substrate N-succinyl-alanine-alanine-phenylalanine 7-amido-4-methylcoumarin (AAF-MCA).[5] Assay procedures have been previously described by Halio et al.[5] Proteinase activity (proteolysis at the inside of a peptide chain) for PfpI can be followed using gelatin-containing zymograms (Novex, San Diego, CA) or by caseinolytic assay using BODIPY FL casein (Molecular Probes, Inc., Eugene, OR). It should be noted that Tris buffers are often replaced with sodium phosphate buffers at higher temperatures, and that the gelatin-containing gels can only be heated to 70°C before autohydrolysis occurs.

In addition, the zymograms should be run at 4°C to prevent proteolysis during electrophoresis, since PfpI is not denatured or inhibited by SDS.[7]

Purification of PfpI from P. furiosus

All purification steps are carried out at room temperature on a Pharmacia LKB FPLC system (Pharmacia, Uppsala, Sweden). Samples are either clarified with a 0.2 μm filter or centrifuged to remove cellular debris before being subjected to chromatographic separation. All concentration steps are done with stirred-cell concentrators (Amicon, Beverly, Mass.) using filters of 10-kDa molecular weight cut-off. Elution of PfpI is followed by assaying for hydrolysis of the substrate AAF-MCA, as described in the assay protocol above. Five-minute incubations at 85°C are used during all of the assays.[7]

DEAE chromatography

Cell extract (420 ml) is applied to a 600-ml DEAE CL-6B XK50 column (Pharmacia, Uppsala, Sweden), which is initially equilibrated with 4 L of 50 mM sodium phosphate buffer pH 8. After 2 L of pass-through, an 8 L linear gradient from 0-1 M NaCl in 50 mM sodium phosphate buffer pH 8 is used to elute the proteins collected in 24 ml fractions at a flow rate of 4 ml/min. The majority of PfpI activity elutes between 168-240 mM NaCl.[7]

Hydrophobic interaction chromatography (HIC)

The active fractions are concentrated to 150 ml, equilibrated with 61 g/L ammonium sulfate (using repeated dilution and concentration steps with a stirred-cell concentrator), and loaded at 2 ml/min onto a 70 cm x 5 cm XK50 column (Pharmacia, Uppsala, Sweden)

packed with 750 ml Phenyl-Sepharose 650 M (Toso Haas, Montgomeryville, PA). The column is equilibrated with 2 L of 122 g/L ammonium sulfate in 50 mM sodium phosphate buffer pH 7. After a 750 ml pass-through, proteins are eluted with a 250 ml decreasing gradient from 122-36 g/L ammonium sulfate in sodium phosphate buffer pH 7, followed by a 1 L decreasing gradient from 36-0 g/L ammonium sulfate in sodium phosphate buffer pH 7. The flow rate is 5 ml/min and 25 ml fractions are collected. PfpI activity elutes from 17-0 g/L ammonium sulfate in two peaks, which are separately combined into two pools of samples (275-ml and 200-ml volumes). By SDS-PAGE analysis, the first peak (275 ml) appears as primarily the hexameric form of PfpI, while the second peak (200 ml) corresponds to the trimeric form of PfpI.[7]

Hydroxylapatite (HAP) chromatography

The second (trimer) pool from the HIC run is concentrated from 200 to 28 ml and loaded onto an 80 ml hydroxylapatite (Calbiochem, La Jolla, CA) XK 16 column (Pharmacia, 1.6 x 40 cm) equilibrated with 400 ml of 50 mM sodium phosphate buffer pH 7. Proteins are eluted with a 300 ml linear gradient from 0-2 M NaCl in approximately 280 ml. This step is omitted for the hexameric HIC pool.[7]

Strong anion exchange (Mono Q) chromatography

The 2 M NaCl eluant from HAP is de-salted and concentrated to 6 ml and the first pool (hexamer) from the HIC run is concentrated to 7 ml. Both pools are further purified separately by multiple loadings of 0.5 ml onto a 1 ml Mono Q column (HR 5/5, Pharmacia) at a flow rate of 0.1 ml/min. The column is equilibrated with 10-20 ml of 100 mM sodium

phosphate buffer pH 8 prior to loading, and washed with 8 ml of the same buffer after loading. Proteins are eluted with sequential linear gradients: 0-120 mM NaCl in 100 mM sodium phosphate buffer pH 8 (8 ml), 120-130 mM NaCl (9 ml), 130-400 mM NaCl (20 ml), and 0.4-1.5 M NaCl (5 ml). A flowrate of 0.5-1.0 ml/min is used, depending on the back pressure, and 2 ml fractions are collected. Generally, PfpI elutes in the second gradient (120-130 mM NaCl).[7]

Gel filtration chromatography

Mono Q eluant from each of the HIC pools are combined separately and concentrated to 1-3 ml before loading (0.5 ml at a time) onto a Superdex 200 HiLoad 16/60 column (Pharmacia, Uppsala, Sweden) equilibrated with 50 mM sodium phosphate buffer pH 7 containing 150 mM NaCl. A flowrate of 0.2-0.25 ml/min is used. Peaks corresponding to molecular masses of 107-112 kDa for the hexamer form and 47-56 kDa for the trimer form are typical. Certain proteolyzed fractions may elute much later, as noted in the report by Halio et al.[16] Some activity may need to be sacrificed in order to avoid fractions that contain multiple forms of the PfpI enzyme. In past experiments, the void volume of the column was found to be 52.2 ml when measured immediately prior to the run that generated a dodecamer form of PfpI. Blue dextran (2000 kDa) is used to determine the void volume and the following proteins standards are used for the calibration curve: thyroglobulin (669 kDa), apoferritin (443 kDa), α-amylase (200 kDa), and bovine albumin serum (66 kDa).[7]

Affinity column purification

Affinity purification of PfpI is used primarily with cell extracts by allowing PfpI to degrade other proteins through storage or heating in small amounts of detergent, or it can be used after an initial ion exchange step (DEAE or Mono Q). Affinity purification is not efficient for obtaining large amounts of homogeneous PfpI, but it does allow for the larger forms of PfpI (dodecamer or 15-mer) to be isolated. These larger forms are not purified using the standard procedure because the HIC and HAP chromatography steps most likely disrupt hydrophobic interactions between the subunits of PfpI, leading to dissociation of the multimers. Therefore, affinity column purification is the preferred purification method to obtain moderate amounts of the larger forms of PfpI that are not degraded into the smaller forms (trimer or hexamer).[7]

Construction of affinity column for PfpI

The affinity column is constructed based on the affinity of PfpI for the peptide AAF-X.[5] Pre-packed NHS-activated HiTrap™ columns (1 ml) are obtained from Pharmacia (Uppsala, Sweden). AAF-MCA (Sigma Chemical Co.) or AAF (Synpep Corporation, Dublin, CA) is dissolved in 0.2 M $NaHCO_3$, 0.5 M NaCl, pH 8.5 at a concentration of 5 mg/ml. In the case of AAF-MCA, 15% DMSO is used as a co-solvent to increase its solubility. Ligand coupling is performed as directed by vendor's instructions, and excess active groups are deactivated with ethanolamine. The column is stored at 4°C in 50 mM sodium phosphate buffer pH 7, with 0.05% sodium azide. Storage in ethanol has been found to be detrimental to the column. Immediately prior to use, the affinity column is washed,

alternating between binding buffer (50 mM sodium phosphate buffer, pH 7, 150 mM NaCl) and elution buffer (50 mM sodium acetate pH 3, 0.5 M NaCl).[7]

Use of affinity chromatography for PfpI

The first affinity column for PfpI was based on AAF-MCA, which has a binding constant of 152 μM.[7] Elution was done at 4°C using a Luer-Lok syringe to deliver the elution buffer described in the previous section. It was thought that binding, but not hydrolysis, would occur at this temperature. However, the enzyme would not elute from the column and a steady amount of MCA was detectable in the wash. It was subsequently determined that PfpI does, in fact, have measurable activity at 4°C. Therefore, a second column was constructed using only the peptide AAF. One ml of cell extract, which had already been enriched by PfpI with long-term storage (the same can be accomplished with heating and small amounts of denaturants), was loaded at 0.1 ml/min to the column (at 25°C). Table 1 shows PfpI activity that eluted at the end of the pass-through (fractions 4-5), the beginning of the pH gradient (fractions 7-8), at pH 3 (fractions 10-11), and after several column volumes exposure to low pH (fractions 19-22). The resulting overall purification fold is lower than previously published for PfpI[5,7] as the protein samples were already enriched with PfpI by incubation of the cell-free extract. The specific activity is also lower than previously published, perhaps due to prolonged exposure of the protein to low pH, changes in buffer composition, and the age of the protein sample. As shown in Figure 2, the eluted protein is predominantly the larger form of the protein. Furthermore, it is of similar purity to the PfpI that had been purified from the same pool by both ion exchange and gel-

filtration chromatography. The zymogram in Figure 2 shows that, although later fractions are very dilute, they still have detectable gelatinase activity.[7]

This affinity purification approach is particularly useful for proteases with affinity for substrates toward which other proteases in the cell extract do not have activity. In the future, it would be desirable to use reversible aldehyde inhibitors to construct the column; if the K_i is sufficiently low, it may be necessary to use chaotropic agents such as guanidinium chloride for elution. In some cases, it may be possible to elute with a substrate for which the protease has a lower K_m. Separation from the substrate can be easily achieved using ultrafiltration or gel filtration chromatography.[7]

Estimation of molecular mass by cross-linking

Because PfpI occurs in multiple active forms, it is difficult to assess the amounts of these forms even after several purification procedures. One approach that can be used is to cross-link the various PfpI forms with dimethyl suberimidate (Sigma Chemical, St. Louis, MO).[17] Cross-linker (10 mg/ml) is dissolved in 200 mM triethanolamine pH 8.5, and mixed with varying amounts of concentrated protein to yield final total protein concentrations around 0.5 mg/ml. The mixtures are incubated at room temperature for 3 hours, heated at 90°C in 1% SDS and 1% β-mercaptoethanol for 1 hour, and analyzed by 7.5-10% SDS-PAGE. These conditions are such that the predominant amount of cross-linking occurs only between structures that are already associated through non-covalent interactions. Aldolase from rabbit muscle (Boehringer Mannheim, Indianapolis, IN), a homo-tetramer of 39.2 kDa subunits, is used as a control. This method has demonstrated that PfpI occurs in several

active forms (multiples of three up to a 15-mer),[7] in addition to the trimeric and hexameric structures previously reported.[5]

Use of irreversible inhibitors to prevent PfpI from destroying intracellular proteins

PfpI has been shown to be capable of degrading nearly all other intracellular proteins upon extended incubation of the cell-free extracts in either the presence[2] or absence[7] of SDS and other denaturants. It is a particular problem during attempts to purify other proteins from *P. furiosus*. Specifically, PfpI directly interferes with the purification of the PEPase[11,12] if the two proteins are not separated immediately after cell disruption.[7] Titration studies with the inhibitor AAF-CMK (chloromethyl ketone) showed that active forms of PfpI (trimer, hexamer, nonamer) contain approximately 2 reactive sites per monomer.[7] Purification studies indicate that PfpI can account for 0.4-0.5% of the total protein in *P. furiosus* cell extracts under certain conditions.[5,7] Therefore, for every 200 mg of cellular protein, there is approximately 1 mg or 0.053 μmoles of PfpI monomer. To inhibit PfpI activity during purification of *P. furiosus* proteins, it is useful to add a three-fold excess (0.16 μmoles per 0.053 μmoles PfpI) of inhibitor to account for other AAF-active proteases and aminopeptidases in the cell extracts (an amino-terminally blocked inhibitor would be more desirable, if commercially available). When the inhibition procedure was previously done, there was 62% inhibition of PfpI activity after incubating the cell extract with inhibitor AAF-CMK for 15 minutes at 70°C. Complete reaction with inhibitor occurs between 4-12 hours at 50-65°C. More inhibitor can be used to expedite the reaction. This same protocol can be used with an excess of reversible inhibitor in the case of storing autoproteolytic enzymes.[7]

Cloning and expressing the pfpI gene

The gene encoding PfpI (GenBank accession number U57642) can be cloned and expressed as a fusion protein with a histidine tag.[4] The *pfpI* gene is amplified by PCR using Vent DNA polymerase, primers complementary to each end of the gene, and *P. furiosus* genomic DNA. The upstream and downstream primers are designed with an NdeI restriction site at the N-terminal methionine of the *pfpI* gene and a BamHI restriction site immediately following the *pfpI* stop codon, respectively. The PCR product is gel purified and ligated to a pET-15b plasmid (Novagen). The resulting vectors are electrotransformed into *E. coli* BL21 (Novagen) and induced with CE6 λ phage, after growth at 30°C to an optical density at 600 nm of 0.6. The *E. coli* host cells are harvested by centrifugation and disrupted by sonication. The recombinant protein is separated from other soluble proteins by using the HisBind Resin and Buffer Kit (Novagen). After purification, the product is dialyzed and concentrated, and the histidine tag was removed by addition of thrombin protease.[4]

The histidine fusion protein expression product that was obtained by Halio et al.[4] was less stable and had different properties than the native PfpI. In addition, only small amounts of the recombinant protein were produced. The temperature optimum and half-life values for the recombinant protease were 95°C and 19 minutes (at 95°C), suggesting that the recombinant form was less stable than the native form. As a result, several other expression methods have been attempted with minimal success, including the T7 promoter in pET-22b+ (inducer CE6 phage), the P_{BAD} promoter in pBAD24 (inducer arabinose), the ϕ31 promoter in pTRK360 (inducer ϕ31 phage), and several other systems.[18] The search to find a better expression system for PfpI is continuing.

Biochemical and biophysical properties of PfpI

The PfpI monomer contains 166 amino acid residues, is approximately 19 kDa in size, but proteolytically inactive.[4] However, *in vitro* these subunits assemble into proteolytically active trimeric, hexameric, and higher trimer-based forms.[5,7] The largest structures seen by Western blots and zymogram gels is 275 kDa, the approximate size of a 15-mer.[7] The trimer form (the smallest active form) is the smallest assembly observed under non-denaturing conditions[7] and is the primary form present after prolonged heating in the presence of SDS.[2,4,5] The trimer (59±3 kDa) and hexamer (124±6 kDa) forms of PfpI are not fully denatured by SDS and, on denaturing gels, have apparent masses of 66 and 86 kDa, respectively.[5] The trimer form self-assembles into the hexamer upon incubation at catalytically sub-optimal temperatures (4°C).[5]

PfpI has both chymotrypsin- and trypsin-like activities and is affected by serine inhibitors.[2] To date, this protease is most active toward the synthetic substrate N-succinyl-alanine-alanine-phenylalanine 7-amido-4-methylcoumarin (AAF-MCA).[5] Large proteins, such as azocasein and gelatin, are also degraded by PfpI.[6,7] Using AAF-MCA, the temperature and pH optima of PfpI are 86°C and 6.3, respectively.[5] The experimentally determined pI values of the trimer and hexamer are 3.8 and 6.1, respectively, while the calculated pI of the 18.8-kDa subunit is 6.1. If all asparagine and glutamine residues are changed to their corresponding acidic residues, the pI of the monomer is calculated to be 4.8. At high temperatures it is possible that the protein is deamidated, especially is its trimeric form, since more thermally labile amino acid residues (i.e., glutamine and asparagine) are left unprotected. This could explain the difference between the experimentally determined and calculated pI values.[5]

The role of PfpI in *P. furiosus* has yet to be determined, although its ability to hydrolyze the majority of cell protein during extended incubation of cell extracts under denaturing conditions suggests it is a predominant protease. It is not known whether the multiple active forms seen *in vitro* also exist *in vivo*. Efforts are underway to explore the role of PfpI in *P. furiosus* as well as in other organisms that contain PfpI homologs.

Acknowledgments

This work was supported in part by grants from the Energy Biosciences Program of the Department of Energy and the National Science Foundation. LSC acknowledges the support of the Department of Education GAANN Fellowship.

References

[1] G. Fiala and K. O. Stetter, *Arch. Microbiol.*, **145**, 56 (1986).

[2] I. I. Blumentals, A. S. Robinson and R. M. Kelly, *Appl. Environ. Microbiol.*, **56**, 1992 (1990).

[3] H. Connaris, D. A. Cowan and R. J. Sharp, *J. Gen. Microbiol.*, **137**, 1193 (1991).

[4] S. B. Halio, I. I. Blumentals, S. A. Short, B. M. Merrill and R. M. Kelly, *J. Bacteriol.*, **178**, 2605 (1996).

[5] S. B. Halio, M. W. Bauer, S. Mukund, M. W. W. Adams and R. M. Kelly, *Appl. Environ. Microbiol.*, **63**, 289 (1997).

[6] P. M. Hicks and R. M. Kelly, In A. J. Barrett, N. D. Rawlings and J. F. Woessner (eds.), *PfpI Protease I (Pyrococcus furiosus)*. Academic Press Limited, London, 1579 (1998).

[7] P. M. Hicks, Ph.D. Thesis, North Carolina State University, Raleigh, NC, Ch. 5 and 8 (1998).

[8] R. Eggen, A. Geerling, J. Watts and W. M. de Vos, *FEMS Mirobiol. Lett.*, **71**, 17 (1990).

[9] W. G. B. Voorhorst, R. I. L. Eggen, A. C. M. Geerling, C. Platteeuw, R. J. Siezen and W. M. de Vos, *J. Biol. Chem.*, **271**, 20426 (1996).

[10] M. W. Bauer, S. H. Bauer and R. M. Kelly, *J. Bacteriol.*, **63**, 1160 (1997).

[11] K. A. Robinson, D. A. Bartley, F. T. Robb and H. J. Schreier, *Gene*, **152**, 103 (1995).

[12] V. J. Harwood, J. D. Denson, K. A. Robinson-Bidle and H. J. Schreier, *J. Bacteriol.*, **179**, 3613 (1997).

[13] M. Ghosh, A. M. Grunden, D. M. Dunn, R. Weiss and M. W. W. Adams, *J. Bacteriol.*,

180, 4781 (1998).

[14] P. M. Hicks, L. S. Chang, and R. M. Kelly, manuscript in preparation.

[15] B. Keil, *Specificity of Proteolysis*. Springer-Verlag, Berlin (1992).

[16] S. B. Halio, Ph.D. Thesis, North Carolina State University, Raleigh, NC (1995).

[17] G. E. Davies and G. R. Stark, *Proc. Natl. Acad. Sci. USA*, **66**, 651 (1970).

[18] P. M. Hicks and R. M. Kelly, unpublished data.

CHAPTER 4:

Pyrococcus furiosus protease I (PfpI) is a homomultimeric oligopeptidase with a distinct physiological role that differs from its homologs

Lara S. Chang[1], *Paula M. Hicks*[1], *Keith R. Shockley*[1], *Donald E. Ward and Robert M. Kelly*[1]*

[1]Department of Chemical Engineering
North Carolina State University
Raleigh, NC 27695-7905

Submitted to: **FEBS Letters** (June, 2003)

Running Title: The oligopeptidase PfpI from *Pyrococcus furiosus* is highly and constitutively expressed

*Address inquiries to: **Robert M. Kelly**
Department of Chemical Engineering
North Carolina State University
Raleigh, NC 27695-7905

Phone: (919) 515-6396
Fax: (919) 515-3465
Email: rmkelly@eos.ncsu.edu

Abbreviations: PfpI, *Pyrococcus furiosus* Protease I; Gal6, bleomycin hydrolase; LAP, leucine aminopeptidase; SPB, sodium phosphate buffer; AS, ammonium sulfate; HIC, hydrophobic interaction chromatography; HAP, hydroxylapatite; MCA, 7-amido-4-methylcoumarin; pNA, p-nitroaniline; GGF, N-succinyl-glycine-glycine-phenylalanine; TFA, triflouroacetic acid; RP-HPLC, reverse-phase high-performance liquid chromatography; ADH, adrenocorticotropic hormone; AAF, N-succinyl-alanine-alanine-phenylalanine; AFK, N-succinyl-alanine-phenylalanine-lysine; LLVY, N-succinyl-leucine-leucine-valine-tyrosine.

I. ABSTRACT

Pyrococcus furiosus protease I (PfpI), from the hyperthermophilic archaeon *P. furiosus*, is a homomultimer composed of identical hexameric rings based on a single 18.8-kDa subunit. *In vitro*, PfpI occurs in at least three functional forms, including trimer, hexamer, and dodecamer, which could be purified separately by column chromatography. The dodecameric assembly is at least three-fold more specifically active than the trimeric and hexameric forms, suggesting a catalytically important structural interaction between the hexameric rings. Biochemical characterization revealed that PfpI is an oligopeptidase, preferring aromatic amino acids, especially phenylalanine, in the P_1 position. While PfpI was able to cleave peptides up to 17 residues in length, little or no activity was found on longer peptides. Additionally, the cleavage pattern by PfpI became less specific and was confined to the N- and C-terminal ends of polypeptides longer than 9 residues. PfpI hydrolyzed the synthetic substrate GGF-pNA with a K_m of 0.85 mM at 85°C, compared to a K_m of 1.7 mM at 95°C; at both temperatures the ν_{max} was approximately 0.20 μM/sec-μg. These parameters are comparable to those determined for other ring-shaped multimeric peptidases, such as the leucine aminopeptidase (LAP) from *Aquifex aeolicus*. The *pfpI* gene was expressed at high levels when *P. furiosus* was grown with cellobiose or tryptone as the primary carbon and energy source. These results suggest that PfpI plays a role in the utilization of peptides as a carbon and energy source as well as in normal turnover of oligopeptides in *P. furiosus*.

II. INTRODUCTION

The hyperthermophilic archaeon *Pyrococcus furiosus* (15) is known to produce several proteases (5,10,13,17,21,43-46), including two intracellular, multimeric, ring-shaped proteases: a version of the archaeal proteasome (3) and PfpI (*Pyrococcus furiosus* protease I, PF1719) (19). The *P. furiosus* proteasome appears to be similar to other archaeal versions of this protease, consisting of stacked rings of α- and β- subunits (3). PfpI is based on a single 18.8-kDa subunit (19), which has putative homologs in many other cells and microorganisms representing the three domains of life, including *Methanococcus jannaschii* (7), *Bacillus subtilis* (2,27,50), *Escherichia coli* (4,22,30,33,34,51) and *Arabadopsis thaliana* (28,39), but not the hyperthermophilic bacterium *Thermotoga maritima* (32) nor *Saccharomyces cerevisiae* (38). PfpI bears some structural relationship to other multimeric proteases; it was recognized on Western blots by anti-mammalian proteasome antibodies (42) and anti-*E. coli* ClpP antibodies (18). The three-dimensional structure of the PfpI homolog in *Pyrococcus horikoshii* (PhpI), which is 90% identical to PfpI at the amino acid level (14), supports the previously reported biophysical information on PfpI (19,20,23): PhpI consists of hexameric rings, each with axes of symmetry such that it contains a dimer of trimers or a trimer of dimers (14). In addition, based on its crystal structure, PhpI may be classified together with the ATP-independent cysteine proteases bleomycin hydrolase (Gal6) (24), which is found only in eukarya and bacteria, and leucine aminopeptidase (LAP) (8), which is found in all three domains of life (14). Both of these broadly specific aminopeptidases are composed of hexameric rings, such that each monomer contains an active site of minimal accessibility that lines an inner channel through the middle of each ring. Although it appears that PhpI

contains only three active sites per hexamer, its three-dimensional structure is reminiscent of the barrel-like compartmentalization of active sites in Gal6 and LAP, as well as in the ATP-dependent 20S proteasome and ClpP. The high level of amino acid identity between PfpI and PhpI, along with the similarities between the determined structure of PhpI (14) and biophysical information on PfpI (19,20,23), lead to the conclusion that PfpI has a three-dimensional structure that is virtually identical to that reported for PhpI (14).

A primary issue to resolve for PfpI is how its molecular properties relate to physiological role. Here, we examine biochemical and biophysical characteristics of PfpI in its various catalytically active forms, determine the substrate specificity and preferred sites of peptide cleavage, and consider its physiological role in *P. furiosus* with respect to other organisms.

III. MATERIALS AND METHODS

Isolation of PfpI functional forms. *P. furiosus* (DSM 3638) biomass (170 g) was kindly provided by M. W. W. Adams, University of Georgia, Athens, GA. Cells, grown on maltose-based medium in the absence of sulfur (1), were re-suspended in approximately 400 mL of 50 mM sodium phosphate buffer (SPB) pH 7, disrupted in a French-pressure cell (18,000 psi), according to manufacturer's protocols (SLM Aminco, Urbana, IL), and centrifuged (10,000 x g, 4ºC) for 30 minutes, yielding a crude extract of 11.5 mg protein/ml. All purification steps for PfpI were carried out at room temperature with a Pharmacia LKB FPLC system (Pharmacia, Uppsala, Sweden). Samples were either clarified with a 0.2 μm filter or centrifuged to remove cellular debris before being subjected to chromatographic separation. All concentration steps were done with stirred-cell concentrators (Millipore, Bedford, MA), using filters of 10 kDa molecular weight cut-off (MWCO). Elution of PfpI was followed by assays for hydrolysis of AAF-MCA as described below.

Ion exchange (DEAE) chromatography. Cell extract (420 ml) was first separated on a 600-ml DEAE CL-6B XK50 column (Pharmacia, Uppsala, Sweden) that was previously equilibrated with 4 L of 50 mM SPB pH 8. After 2 L of pass-through, an 8-L linear gradient (4 ml/min) from 0-1 M NaCl in 50 mM SPB pH 8 was used to elute protein in 24-ml fractions. The majority of PfpI activity eluted between 168-240 mM NaCl.

Hydrophobic interaction chromatography (HIC). The DEAE fractions with PfpI activity were combined and concentrated to 150 ml, equilibrated with 61 g/L ammonium sulfate (AS), and loaded at 2 ml/min onto a 70 cm x 5 cm XK50 column (Pharmacia, Uppsala, Sweden), packed with 750 ml of Phenyl-Sepharose 650 M (hydrophobic interaction

chromatography, HIC) (Toso Haas, Montgomeryville, PA), previously equilibrated with 2 L of 122 g/L AS in 50 mM SPB pH 7. After a 750 ml pass-through, protein was eluted with a 250-ml gradient from 122-36 g/L AS in 50 mM SPB pH 7, followed by a 1-L gradient from 36-0 g/L AS in 50 mM SPB pH 7, at 5 ml/min in 25-ml fractions. PfpI activity eluted from 17-0 g/L AS, in 2 separate peaks and a shoulder. The first peak (200 ml pool) and the shoulder (75 ml pool) were the same protease; the second peak (200 ml pool) corresponded to a lower molecular weight when separated by SDS-PAGE. These two peaks were then separated into pools of 275 ml (pool 1) and 200 ml (pool 2).

Hydroxylapatite (HAP) chromatography. Pool 2 from HIC was concentrated from 200 to 28 ml, and loaded onto an 80-ml hydroxylapatite (HAP) (Calbiochem, La Jolla, CA) XK 16 column (Pharmacia, 1.6 x 40 cm) previously equilibrated with 400 ml of 50 mM SPB pH 7. Proteins were eluted in 10-ml fractions with a 300 ml linear gradient from 0-2 M NaCl in 100 mM SPB pH 7 at a flow rate of 1.5 ml/min. PfpI activity eluted during the end wash at 2 M NaCl in approximately 280 ml.

Ion exchange (Mono Q) chromatography. The HAP fractions (pool 2 from HIC) were de-salted and concentrated to 6 ml, and the pool 1 fractions from HIC were concentrated to 7 ml; both of these were separately loaded at 0.5 ml at a time onto a 1-ml Mono Q column (HR 5/5, Pharmacia) at a flow rate of 0.1 ml/min. The column was equilibrated with 10-20 ml of 100 mM SPB pH 8 (buffer A) prior to loading and with 8 ml after loading. Protein was eluted with 4 different linear gradients: 0-120 mM NaCl (in 100 mM SPB pH 8) over 8 ml, 120-130 mM NaCl over 9 ml, 130-400 mM NaCl over 20 ml, and 0.4-1.5 M NaCl over 5 ml. A flow rate of 0.5-1.0 ml/min was used, depending on the back-

pressure, and 2-ml fractions were collected. PfpI activity eluted most frequently in the second gradient (120-130 mM NaCl).

Gel filtration chromatography. Fractions from the Mono Q step, which included pool 1 from the HIC step, were combined separately from the HIC pool 2 / HAP / Mono Q fractions. Each pool was concentrated to 1-3 ml before loading at a flow rate of 0.2-0.25 ml/min (0.5 ml at a time) onto a Superdex 200 HiLoad 16/60 (Pharmacia, Uppsala, Sweden), which was previously equilibrated with 50 mM SPB at pH 7 containing 150 mM NaCl. Peaks corresponding to molecular masses of 107-112 kDa for hexamer and 47-56 kDa for trimer were typical, along with proteolytically-active fractions eluting much later as noted in previous work (20). Blue dextran (2000 kDa) was used to determine void volume, and the following protein standards were used for the calibration curve: thyroglobulin (669 kDa), apoferritin (443 kDa), β-amylase (200 kDa), and bovine serum albumin (BSA) (66 kDa).

Differential scanning calorimetry. Differential scanning calorimetry was performed on a nano-differential scanning calorimeter (Calorimetry Sciences Corp., Provo, UT) from 25-125ºC (both heating and cooling) at a scan rate of 1ºC/min. PfpI was concentrated to 0.41 mg/ml, and dialyzed against 50 mM SPB pH 7. Predicted curves and theoretical enthalpy calculations were determined according to manufacturer's protocol (9).

Peptidase activity assays. Proteolytic activity for screening during purification and substrate preference determination was detected (typically at 85ºC) by release of 7-amino-4-methylcoumarin (MCA) from the carboxyl terminus of N-terminally blocked peptides (Sigma-Aldrich, St. Louis, MO) in 50 mM SPB pH 7, as described previously (20).

Substrate concentrations at 0.25-0.5 mM were used with enzyme concentrations between 0.1-1.0 μg/ml. Temperature optima and substrate specificity comparisons were done in triplicate. Temperature optima were determined using a thermal cycler as reported previously (20), except that 0.25 mM substrate was used, and the specificity comparison was done in a total volume of 100 μL with 0.5 mM substrate in a 96 U-well microtiter plate. Total protein concentration was determined using a bicinchoninic acid protein assay reagent kit (Pierce, Rockford, IL) with BSA as the standard.

Kinetic data was determined spectrophotometrically (Perkin Elmer Lambda Bio 20 UV/Vis spectrometer with heated chamber) by detection of p-nitroaniline (pNA) liberated from the chromogenic substrate N-succinyl-glycine-glycine-phenylalanine p-nitroanilide (GGF-pNA) (Sigma-Aldrich, St. Louis, MO). Substrate was dissolved to 50 mM in 100% dimethyl sulfoxide (DMSO) and stored in the dark at -20^{o}C for no longer than 2 months. All assay mixtures contained 50 mM SPB pH 7, with 1.6% DMSO in a total volume of 820 μl. Buffer, substrate, and DMSO were preheated at the assay temperatures for 4 min in sealed quartz cuvettes in the spectrometer, at which time buffer (for reference cell) or preheated (1 min) enzyme was added. Change in absorbance over time was recorded in triplicate at 405 nm continuously for a maximum of 5 min for each substrate concentration. Calculated molar absorption coefficients (ε) for pNA were approximately 10,400 $M^{-1}cm^{-1}$ and 10,300 $M^{-1}cm^{-1}$ at 85^{o}C and 95^{o}C, respectively. Kinetic parameters were calculated using a nonlinear regression method based on average initial velocity values. One unit of activity is defined as the amount of enzyme that hydrolyses 1 μmol of pNA per second under assay conditions.

Polypeptide activity assays. To examine PfpI cleavage of polypeptides, substrates (Sigma-Aldrich, St. Louis, MO) were incubated with PfpI at an average enzyme/substrate ratio of 1/35 (by weight) in 50 mM potassium phosphate buffer pH 7.2 at 85ºC. All reactions were quenched with 1% triflouroacetic acid (TFA) and processed through cellulose 30 kDa MWCO filters (Millipore, Bedford, MA). Flow-through from filtration was separated by reverse-phase high-performance liquid chromatography (RP-HPLC) using a C_{18} Nucleosil column in a TFA / acetonitrile solvent system. The peptide molecular weights were then determined with an LCQ Mass Spectrometer System (Wake Forest University Medical Center, Wake Forest, NC). For substrates bradykinin, angiotensin, and adrenocorticotropic hormone (ADH) segment 1-17, the assay incubation time was 1 hour. For neurotensin, insulin chains A and B, and ADH segment 1-24 the incubation time was 4 hr. For all substrates except ADH, an assay concentration of 200 μM was used. For both ADH 1-17 and 1-24, a concentration of 50 μM was used.

Transcriptional analyses of *pfpI*. *P. furiosus* was grown at 90ºC in a 500-ml culture bottle with either 10 mM cellobiose + 1 g/L yeast extract, or 5 g/L (w/v) tryptone + 2 g/L (w/v) yeast extract. For the tryptone-based medium, elemental sulfur was added at 1% (w/v). Sea salts (40 g/L) and PIPES (3.1 g/L) were used in all media formulations. Also, 2.5 ml of resazarin (0.4 g/L) solution and 0.1% (v/v) trace element solution were added before adjusting the pH to 6.8 with NaOH. The trace element solution was composed of (grams per 100 ml): nitrilotriacetic acid, 1.50; $FeCl_2 \cdot 6\ H_2O$, 0.50; $Na_2WO_7 \cdot 2\ H_2O$, 0.30; $MnCl_2 \cdot 4\ H_2O$, 0.40; $NiCl_2 \cdot 6\ H_2O$, 0.20; $ZnSO_4 \cdot 7\ H_2O$, 0.10; $CoSO_4 \cdot 7\ H_2O$, 0.10; $CuSO_4 \cdot 5\ H_2O$ (10 mg/ml), 1 ml; $Na_2MoO_4 \cdot 5\ H_2O$ (10 mg/ml), 1 ml. Before inoculation, 1.1 ml of a 10% (w/v) Na_2S solution was added and the bottle was sealed and flushed with N_2 for 2 min. Cultures were

inoculated with a 0.5% inoculum from a 60-ml culture grown overnight. Growth was monitored by epifluorescence microscopy as described previously (11). *P. furiosus* was grown until early- to mid-exponential phase at 90ºC and then harvested by spinning at 7500 rpm and 4ºC for 22 min. For the heat shock condition, the cells were grown until early- to mid-exponential phase at 90ºC and then immediately subjected to heat shock at 105ºC for up to 1 hr in a silicone oil bath. Total RNA extraction was performed using the RNAqueous^{TM} kit (Ambion, Austin, TX).

Northern analysis was carried out as described previously (12) except that 20 µg of total RNA was loaded per lane. The *pfpI* probe was amplified from *P. furiosus* genomic DNA using Pfu Turbo DNA polymerase (Stratagene, La Jolla, CA), with primer sequences of 5’ - GTG ATA CAT ATG AAG ATA CTG TTC TTG AGT GCA A - 3’ for forward and 5' - AGA AAG GAT CCT TAC AAT GAT CAC TTA AGT AAT TTA ACA AAT - 3' for reverse. As seen in Figure 4.3, Northern blot analysis revealed that the *pfpI* gene was expressed at comparable levels on both cellobiose- and tryptone-based media.

IV. RESULTS AND DISCUSSION

Characterization of PfpI functional forms. *In vitro*, PfpI can exist in several functional forms that are readily apparent when the protease is incubated at physiological temperatures (20). Whether this phenomenon is significant *in vivo* is not known. One of the issues to be examined was whether the larger catalytically active form of PfpI (α_{12}) differed in its biochemical and biophysical properties from the trimeric and hexameric forms. However, the difficulty in achieving this stems from the consistent observation that the structure is dynamic, exhibiting a tendency to assemble into larger forms upon incubation at temperatures sub-optimal for catalysis (19). The precise rate at which the subunits assemble remains to be determined, however, it appears to be high enough that the ratio of trimer to hexamer to dodecamer (and higher forms) is not reliably constant over longer experimental periods. Therefore, in many experimental procedures, the assumed protein sample must remain an assumed mixture of the various subunit assemblies.

Table 4.1 summarizes the approach used to purify the various forms from *P. furiosus* cell extracts. Although similar to the approach described previously (20), strong activity towards AAF-MCA could be separated into two distinct pools of eluent from the HIC step, designated pools 1 and 2. After both additional ion exchange and gel filtration, pool 1 led to the purification of the trimeric (α_3, 52.5 kDa by gel filtration) and hexameric (α_6, 111 kDa by gel filtration) forms, with a ratio of specific activities (α_6/α_3) equal to 1.3. Pool 2 led to the purification of α_3 (46.4 kDa by gel filtration) and α_6 (107.4 kDa by gel filtration) forms, following HAP, ion exchange and gel filtration, with a ratio of specific activities (α_6/α_3) equal to 1.4. Pool 2 also yielded a larger form, which was most likely the dodecamer (α_{12}),

with an estimated molecular mass of 198.3 kDa compared to an expected mass of 225.6 kDa from amino acid sequence information. Note that for this gel filtration column, the larger form eluted before the 200-kDa standard (β-amylase). Protease activities reported in Table 4.1 are averages of all contributions detected following the gel filtration step. The specific activity towards the substrates tested for the α_{12} was significantly higher than for the α_3 and α_6 forms, e.g., for AAF-MCA, $\alpha_{12}/\alpha_6/\alpha_3$ was 2.6/1.4/1 (see Table 4.2), suggesting a catalytically important structural interaction between the α_6 rings. Du et al. (14) reported that the PhpI hexamer, which appears nearly identical to PfpI in structure, has three active sites located in hindered positions; the configuration of the active sites at the α_{12} level (and higher levels) is unknown. The relatively low levels of activity observed by the α_3 and α_6 forms may be a result of the sheltered active sites. It is possible, however, that the interaction between α_6 rings that creates the α_{12} form alters the configuration of the active sites enough to make them more accessible to certain substrates, thereby increasing their relative activities.

The melting temperature (T_m) for PfpI recovered in Pool 1 was determined to be 102ºC (Figure 4.1), at which temperature the catalytic activity significantly decreased. After heating this sample to 125ºC and holding it at that temperature for 20 min, 10% of initial activity could be recovered from the sample after it was cooled to room temperature and assayed at 85ºC. At temperatures well below the T_m, the three forms of PfpI had broad temperature optima (T_{opt}), such that specific activity toward AAF-MCA varied little over the range 75-90ºC (Figure 4.2). However, as the temperature was increased to 100ºC, close to the melting point, the specific activity of the three forms appeared to be converging, which

may reflect disassembly into the trimeric form, a characteristic that has been observed previously (5,19).

Kinetic parameters at (85ºC) and above (95ºC) the T_{opt} were also compared. PfpI degraded GGF-pNA with a K_m of 0.85 mM at 85ºC, compared to a K_m of 1.7 mM at 95ºC; at both temperatures the ν_{max} remained at approximately 0.20 μM/sec-μg. Based on the structural data of PhpI (14), the hexameric form contains 3 active sites. If the PfpI used in the kinetic assays is assumed to be a homogeneous mixture of hexamers, each containing 3 active sites, or dodecamers, each containing 6 active sites, the resulting k_{cat} values are nearly identical at approximately 6.1 sec^{-1} at 85ºC and 6.6 sec^{-1} at 95ºC. As such, the k_{cat}/K_m values were 7200 $M^{-1}sec^{-1}$ at 85ºC and 3900 $M^{-1}sec^{-1}$ at 95ºC. Beyond the report that it is a cysteine protease (14), PfpI has not yet been classified, but its kinetics are comparable to other multisubunit proteases and cysteine proteases. For example, the archaeal proteasome from *Thermoplasma acidophilum* exhibits peptidase activity with a K_m of 0.085 mM and k_{cat} of 0.03 sec^{-1} (29). Cathepsins K, S, L, and B (lysosomal cysteine peptidase) have a range of k_{cat}/K_m values on the order of 10^5-10^6 $M^{-1}sec^{-1}$ (6). Very similar to PfpI, LAP from *Aquifex aeolicus* was found to have a K_m of 1.8 mM, k_{cat} of 1.1 sec^{-1}, and k_{cat}/K_m 595 $M^{-1}sec^{-1}$ when tested on its preferred substrate, L-leucine-pNA (25).

As shown in Table 4.2, AAF-MCA was preferred over AFK-MCA and LLVY-MCA for all forms of PfpI. The dodecamer was approximately three times more specifically active on AAF-MCA and AFK-MCA and approximately two times more active on LLVY-MCA than either the trimer or hexamer. According to LC/MS analysis of various polypeptide degradation patterns (Table 4.3), PfpI was able to cleave only small polypeptide chains (≤ 17 amino acids in length) preferring, but not solely specific toward, aromatic amino acids in the

P_1 position. Specifically, it readily cleaved bradykinin after 1 hr of incubation, where phenylalanine represented P_1 at two positions in the substrate. With substrates longer than 9 residues, PfpI was only able to cleave at the ends of the polypeptides and did so less specifically by cleaving at positions where an aromatic residue represented P_1, P_2, or P_1'. For instance, weak cleavage at various positions in angiotensin and ADH 1-17 after 1 hr of incubation was observed. Substantial cleavage of neurotensin at various positions was noted, but only after an extended 85ºC incubation of 4 hr. Contrary to observations with bradykinin, PfpI was unable to cleave near the centrally located residues in a longer polypeptide, such as ADH 1-17, even when an aromatic residue was present. Furthermore, no cleavage was observed on ADH 1-24 or either chain of insulin, even after 24 hr of incubation. Inferring from the PhpI structure (14), this may relate to the location of the active sites in the PfpI hexamer, which appear to be hindered to the point that even small globular proteins would not have access to the catalytic sites. In addition, the broad specificity, particularly with longer polypeptides, may be due to the lack of the active site cleft that normally defines specificity (14). Both the inability to cleave longer polypeptides and the broad specificity support classification of PfpI with Gal6 and LAP, which are both peptidases with a wide range of specificity.

Physiological role of PfpI. Given the significant homology of *pfpI* to genes in a wide variety of cells and organisms, some conservation of physiological function might be expected. In this study, *pfpI* was expressed under normal growth conditions in both cellobiose-base and tryptone-based media. This result is consistent with the recent finding in the microarray-based studies of Schut et al. (40) which showed that *pfpI* was highly

expressed in a sulfur-independent manner for maltose-grown *P. furiosus*. The fact that the expression of PfpI is not affected by the primary carbon source suggests a role in protein turnover in the cell. However, it cannot be ruled out that PfpI may also play a significant role in the utilization of peptides as carbon and energy sources since it is expressed at such high levels (40).

Differences have been found among various homologs that suggest it may play different physiological roles that are dependent on the organism. For example, Table 4.4 summarizes the effects of heat shock on PfpI and two of its bacterial homologs. The PfpI results were found using a protease gene-based targeted DNA array and show that expression of PfpI is also affected when the cells are exposed to heat shock at 105ºC (Shockley et al., unpublished results), particularly during the early stages of the stress event. This is not unexpected, as the *pfpI* homolog in *B. subtilis* (*yfk*M) was found to be among the σ^B-dependent general stress genes (35,36). However, the *pfpI* homolog in *E. coli* (*yhbO*) was not induced for cells grown under any normal or abnormal circumstances examined (http://www.genome.wisc.edu/). This result is consistent with our finding that an *E. coli* mutant with the *yhbO* gene knocked out showed no noticeable differences compared to the wild-type for growth on several substrates or when exposed to heat shock (Miller and Kelly, unpublished results). More recently, a mammalian protein named DJ-1, which was found to be an activated *ras*-dependent oncogene product (31) and an infertility-related protein affected by sperm toxicants (26, 47-49), was found to have similar structure to PfpI (personal communication; Ariga, H. Hokkaido University, Sapporo, Japan). The eukaryotic DJ-1 also exhibited some weak proteolytic activity, most likely necessary during its actions of attachment and penetration of egg surfaces during fertilization (personal communication;

Ariga, H. Hokkaido University, Sapporo, Japan). Thus, while more information is needed to draw general conclusions about the physiological role of this protease, it appears that it varies from case to case. The *P. furiosus* form exhibited high levels of activity on small peptides, which do not appear to be regulated by APTase activities (unpublished data), and a clear inability to cleave polypeptides ≥24 residues long. Furthermore, it is present in high levels in *P. furiosus* cells under varying growth conditions. From these observations, the most likely physiological role for PfpI is in the degradation of small polypeptide fragments that have been already processed by proteinases such as the proteasome (3). It remains to be seen how PfpI and its homologs contribute to overall intracellular proteolytic function within their respective host organisms and cells.

V. ACKNOWLEDGEMENTS

LSC and KRS acknowledge support from Department of Education GAANN Fellowships. PMH acknowledges support from a National Science Foundation graduate fellowship. This work was supported by grants to RMK from the Department of Energy (Energy Biosciences Program) and National Science Foundation (Biotechnology Program).

VI. REFERENCES

1. **Adams, M. W. W.** 1995. Large-scale growth of hyperthermophiles, p. 47-49. *In* F. T. Robb, A. R. Place, K. R. Sowers, H. J. Schreier, S. DasSarma, and E. M. Fleischmann (ed.), Archaea: a laboratory manual. Cold Spring Harbor Laboratory Press, New York.
2. **Antelmann, H., J. Bernhardt, R. Schmid, H. Mach, U. Völker, and M. Hecker.** 1997. First steps from a two-dimensional protein index towards a response-regulation map for *Bacillus subtilis*. Electrophoresis. **18:**1451-1463.
3. **Bauer, M. W., S. H. Bauer, and R. M. Kelly.** 1997. Purification and characterization of a proteasome from the hyperthermophilic archaeon *Pyrococcus furiosus*. J. Bacteriol. **63:**1160-1164.
4. **Blattner, F. R., G. Plunkett, C. A. Bloch, N. T. Perna, V. Burland, M. Riley, J. Collado-Vides, J. D. Glasner, C. K. Rode, G. F. Mayhew, J. Gregor, N. W. Davis, H. A. Kirkpatrick, M. A. Goeden, D. J. Rose, B. Mau, and Y. Shao.** 1997. The complete genome sequence of *Escherichia coli* K-12. Science. **277:**1453-1474.
5. **Blumentals, I. I., A. S. Robinson, and R. M. Kelly.** 1990. Characterization of sodium dodecyl sulfate-resistant proteolytic activity in the hyperthermophilic archaebacterium *Pyrococcus furiosus*. Appl. Environ. Microbiol. **56:**1992-1998.
6. **Brömme, D.** 1998. Cathepsin K. *In* F. Woessner, N. Rawlings, and A. Barrett (ed.), Handbook of Proteolytic Enzymes. Academic Press Limited, London, pp. 624-628.
7. **Bult, C. J., O. White, G. J. Olsen, L. Zhou, R. D. Fleischmann, G. G. Sutton, J. A. Blake, L. M. FitzGerald, R. A. Clayton, J. D. Gocayne, A. R. Kerlavage, B. A. Dougherty, J. F. Tomb, M. D. Adams, C. I. Reich, R. Overbeek, E. F. Kirkness, K.**

G. Weinstock, J. M. Merrick, A. Glodek, J. L. Scott, N. S. M. Geoghagen, and J. C. Venter. 1996. Complete Genome Sequence of the Methanogenic Archaeon, *Methanococcus jannaschii*. Science **273:**1058-1073.

8. **Burley, S. K., P. R. David, A. Taylor, and W. N. Lipscomb.** 1990. Molecular structure of leucine aminopeptidase at 2.7-Å resolution. Proc Natl Acad Sci U S A. **87:**6878-6882.
9. **Calorimetry Sciences Corporation**. 1995. CSC5100 NDSC Users manual v 1.1. Provo, Utah.
10. **Cheng, T. C., V. Ramakrishnan, and S. I. Chan.** 1999. Purification and characterization of a cobalt-activated carboxypeptidase from the hyperthermophilic archaeon *Pyrococcus furiosus*. Protein Sci. **8:**2474-2486.
11. **Chhabra, S., K. N. Parker, D. Lam, W. Callen, M. A. Snead, E. J. Mathur, M. Short, and R. M. Kelly.** 2001. β-Mannanases from Thermotoga species. Method. Enzymol. **330:** 224-238.
12. **Chhabra, S., K. R. Shockley, D. E. Ward, and R. M. Kelly.** 2002. Regulation of endo-acting glycosyl hydrolases in the hyperthermophilic bacterium *Thermotoga maritima* grown on glucan- and mannan-based polysaccharides. Appl. Environ. Microbiol. (accepted pending revision).
13. **Connaris, H., D. A. Cowan, and R. J. Sharp.** 1991. Heterogeneity of proteinases from the hyperthermophilic archaeobacterium *Pyrococcus furiosus*. J. Gen. Microbiol. **137:**1193-1199.
14. **Du, X., I.-G. Choi, R. Kim, W. Wang, J. Jancarik, H. Yokota, and S.-H. Kim.** 2000. Crystal structure of an intracellular protease from *Pyrococcus horikoshii* at 2-Å resolution. Proc Natl Acad Sci U S A. **97:**14079-14084.

15. *Escherichia coli* genome project, University of Wisconsin, Madison, WI. http://www.genome.wisc.edu/.

16. **Fiala, G., and K. O. Stetter.** 1986. *Pyrococcus furiosus* sp. nov. represents a novel genus of marine heterotrophic archaebacteria growing optimally at 100°C. Arch. Microbiol. **145:**56-60.

17. **Ghosh, M., A. M. Grunden, D. M. Dunn, R. Weiss, and M. W. W. Adams.** 1998. Characterization of native and recombinant forms of an unusual cobalt-dependent proline dipeptidase (prolidase) from the hyperthermophilic archaeon *Pyrococcus furiosus*. J Bacteriol. **180:**4781-4789.

18. **Halio, S. B.** 1995. Ph. D. thesis, North Carolina State University, Raleigh, NC.

19. **Halio, S. B., I. I. Blumentals, S. A. Short, B. M. Merrill, and R. M. Kelly.** 1996. Sequence, expression in *Escherichia coli*, and analysis of the gene encoding a novel intracellular protease (PfpI) from the hyperthermophilic archaeon *Pyrococcus furiosus*. J. Bacteriol. **178:**2605-2612.

20. **Halio, S. B., M. W. Bauer, S. Mukund, M. W. W. Adams, and R. M. Kelly.** 1997. Purification and characterization of two functional forms of intracellular protease PfpI from the hyperthermophilic archaeon *Pyrococcus furiosus*. Appl. Environ. Microbiol. **63:**289-295.

21. **Harwood, V. J., J. D. Denson, K. A. Robinson-Bidle, and H. J. Schreier.** 1997. Overexpression and characterization of a prolyl endopeptidase from the hyperthermophilic archaeon *Pyrococcus furiosus*. J. Bacteriol. **179:**3613-3618.

22. **Hayashi, T., K. Makino, M. Ohnishi, K. Kurokawa, K. Ishii, K. Yokoyama, C.–G. Han, E. Ohtsubo, K. Nakayama, T. Murata, M. Tanaka, T. Tobe, T. Iida, H.**

Takami, T. Honda, C. Sasakawa, N. Ogasawara, T. Yasunaga, S. Kuhara, T. Shiba, M. Hattori, and H. Shinagawa. 2001. Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genomic comparison with a laboratory strain K-12. DNA Res. **8:**11-22.

23. **Hicks, P. M., and R. M. Kelly.** 1998. *Pyrococcus furiosus* protease I (PfpI). *In* F. Woessner, N. Rawlings, and A. Barrett (ed.), Handbook of Proteolytic Enzymes. Academic Press Limited, London, pp. 1579-1581.
24. **Joshua-Tor, L., H. E. Xu, S. A. Johnston, and D. C. Rees.** 1995. Crystal structure of a conserved protease that binds DNA: the bleomycin hydrolase, Gal6. Science. **269:**945-950.
25. **Khan, A. R., S. Nirasawa, S. Kaneko, T. Shimonishi, and K. Hayashi.** 2000. Characterization of a solvent resistant and thermostable aminopeptidase from the hyperthermophillic bacterium, *Aquifex aeolicus*. Enzyme Microb. Technol. **27:**83-88.
26. **Klinefelter, G. R., J. W. Laskey, J. Ferrell, J. D. Suarez, and N. L. Roberts.** 1997. Discriminant analysis indicates a single sperm protein (SP22) is predictive of fertility following exposure to epididymal toxicants. J. Androl. **18:**139-150.
27. **Kunst, F., N. Ogasawara, I. Moszer, A. M. Albertini, G. Alloni, V. Azevedo, M. G. Bertero, P. Bessieres, A. Bolotin, S. Borchert, R. Borriss, L. Boursier, A. Brans, M. Braun, S. C. Brignell, S. Bron, S. Brouillet, C. V. Bruschi, B. Caldwell, V. Capuano, N. M. Carter, S. K. Choi, J. J. Codani, I. F. Connerton, A. Danchin A, et al.** 1997. The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. Nature. **390:**249-256.

28. **Lin, X., S. Kaul, C. D. Town, M.–I. Benito, T. H. Creasy, B. Haas, D. Wu, C. M. Ronning, H. Koo, C. Y. Fujii, T. R. Utterback, M. E. Barnstead, C. L. Bowman, O. White, W. C. Nierman, and C. M. Fraser.** *Arabidopsis thaliana* chromosome III BAC F13E7 genomic sequence. Unpublished.

29. **Lupas, A., M. Kania, and W. Baumesiter.** 1998. Archaean proteasome. *In* F. Woessner, N. Rawlings, and A. Barrett (ed.), Handbook of Proteolytic Enzymes. Academic Press Limited, London, pp. 486-489.

30. **Makino, K., K. Yokoyama, Y. Kubota, C. H. Yutsudo, S. Kimura, K. Kurokawa, K. Ishii, M. Hattori, I. Tatsuno, H. Abe, T. Iida, K. Yamamoto, M. Ohnishi, T. Hayashi, T. Yasunaga, T. Honda, C. Sasakawa, and H. Shinagawa.** 1999. Complete nucleotide sequence of the prophage VT2-Sakai carrying the verotoxin 2 genes of the enterohemorrhagic *Escherichia coli* O157:H7 derived from the Sakai outbreak. Genes Genet. Syst. **74:** 227-239.

31. **Nagakubo, D., T. Taira, H. Kitaura, M. Ikeda, K. Tamai, S. M. M. Iguchi-Ariga, and H. Ariga.** 1997. DJ-1, a novel oncogene which transforms mouse NIH3T3 cells in cooperation with *ras*. Biochem. Biophys. Res. Commun. **231:**509-513.

32. **Nelson K. E., R. A. Clayton, S. R. Gill, M. L. Gwinn, R. J. Dodson, D. H. Haft, E. K. Hickey, J. D. Peterson, W. C. Nelson, K. A. Ketchum, L. McDonald, T. R. Utterback, J. A. Malek, K. D. Linher, M. M. Garrett, A. M. Stewart, M. D. Cotton, M. S. Pratt, C. A. Phillips, D. Richardson, J. Heidelberg, G. G. Sutton, R. D. Fleischmann, J. A. Eisen, and C. M. Fraser.** 1999. Evidence for lateral gene transfer between Archaea and bacteria from genome sequence of *Thermotoga maritima*. Nature. **399:**323-329.

33. **Ohnishi, M., T. Murata, K. Nakayama, S. Kuhara, M. Hattori, K. Kurokawa, T. Yasunaga, K. Yokoyama, K. Makino, H. Shinagawa, and T. Hayashi.** 2000. Comparative analysis of the whole set of rRNA operons between an enterohemorrhagic *Escherichia coli* O157:H7 Sakai strain and an *Escherichia coli* K-12 strain MG1655. 2000. Syst. Appl. Microbiol. **23:**315-324.

34. **Perna, N. T., G. Plunkett, V. Burland, B. Mau, J. D. Glasner, D. J. Rose, G. F. Mayhew, P. S. Evans, J. Gregor, H. A. Kirkpatrick, G. Posfai, J. Hackett, S. Klink, A. Boutin, Y. Shao, L. Miller, E. J. Grotbeck, N. W. Davis, A. Lim, E. Dimalanta, K. Potamousis, J. Apodaca, T. S. Anantharaman, J. Lin, G. Yen, D. C. Schwartz, R. A. Welch, and F. R. Blattner.** 2001. Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. Nature. **409:**529-533.

35. **Petersohn, A., M. Brigulla, S. Haas, J. D. Hoheisel, U. Völker, and M. Hecker.** 2001. Global analysis of the general stress response of *Bacillus subtilis*. J. Bacteriol. **183:**5617-5631.

36. **Price, C. W., P. Fawcett, H. Ceremonie, N. Su, C. K. Murphy, and P. Youngman.** 2001. Genome-wide analysis of the general stress response in *Bacillus subtilis*. Mol. Micro. **41:**757-774.

37. **Richmond, C. S., J. D. Glasner, R. Mau, H. Jin, and F. R. Blattner.** 1999. Genome-wide expression profiling in *Escherichia coli* K-12. Nucleic Acids Res. **27:**3821-3825.

38. *Saccharomyces* genome database, http://genome-www.stanford.edu/Saccharomyces/.

39. **Salanoubat, M., K. Lemcke, M. Rieger, W. Ansorge, M. Unseld, B. Fartmann, G. Valle, H. Blocker, M. Perez-Alonso, B. Obermaier, M. Delseny, M. Boutry, L. A. Grivell, R. Mache, P. Puigdomenech, V. De Simone, N. Choisne, F. Artiguenave, C.**

Robert, P. Brottier, P. Wincker, L. Cattolico, J. Weissenbach, W. Saurin, F. Quetier, et al. 2000. Sequence and analysis of chromosome 3 of the plant *Arabidopsis thaliana*. Nature. **408:**820-822.

40. **Schut, G. J., J. Zhou, and M. W. W. Adams.** 2001. DNA microarray analysis of the hyperthermophilic archaeon *Pyrococcus furiosus*: evidence for a new type of sulfur-reducing enzyme complex. J. Bacteriol. **183:**7027-36.
41. **Shockley, K. R., D. E. Ward, S. R. Chhabra, S. B. Conner, C. I. Montero, and R. M. Kelly.** 2003. Heat shock response by the hyperthermophilic archaeon *Pyrococcus furiosus*. Appl. Environ. Microbiol. **69:**2365-2371.
42. **Snowden, L. J., I. I. Blumentals, and R. M. Kelly.** 1992. Regulation of proteolytic activity in the hyperthermophile *Pyrococcus furiosus*. Appl. Environ. Microbiol. **58:**1134-1140.
43. **Tsunasawa, S., Y. Izu, M. Miyagi, and I. Kato.** 1997. Methionine aminopeptidase from the hyperthermophilic Archaeon *Pyrococcus furiosus*: molecular cloning and overexpression in *Escherichia coli* of the gene, and characteristics of the enzyme. J Biochem (Tokyo). **122:**843-850.
44. **Tsunasawa, S.** 1998. Purification and application of a novel N-terminal deblocking aminopeptidase (DAP) from *Pyrococcus furiosus*. J Protein Chem. **17:**521-522.
45. **Tsunasawa, S., S. Nakura, T. Tanigawa, and I. Kato.** 1998. Pyrrolidone carboxyl peptidase from the hyperthermophilic Archaeon *Pyrococcus furiosus*: cloning and overexpression in *Escherichia coli* of the gene, and its application to protein sequence analysis. J Biochem (Tokyo). **124:**778-783.

46. **Voorhorst, W. G. B., R. I. L. Eggen, A. C. M. Geerling, C. Platteeuw, R. J. Siezen, and W. M. de Vos.** 1996. Isolation and characterization of the hyperthermostable serine protease, pyrolysin, and its gene from the hyperthermophilic archaeon *Pyrococcus furiosus*. J. Biol. Chem. **271:**20426-20431.

47. **Wagenfeld, A., J. Gromoll, and T. G. Cooper.** 1998. Molecular cloning and expression of rat contraception associated protein 1 (CAP1), a protein putatively involved in fertilization. Biochem. Biophys. Res. Commun. **251:**545-549.

48. **Wagenfeld A., C. H. Yeung, K. K. Strupat, and T. G. Cooper.** 1998. Shedding of a rat epididymal sperm protein associated with infertility induced by ornidazole and alpha-chlorohydrin. Biol. Reprod. **58:**1257-1265.

49. **Welch, J. E., R. R. Barbee, N. L. Roberts, J. D. Suarez, and G. R. Klinefelter.** 1998. SP22: a novel fertility protein from a highly conserved gene family. J. Androl. **19:**385-393.

50. **Yamamoto, S. Uchiyama, and J. Sekiguchi.** 1996. Cloning and sequencing of a 40.6 kb segment in the 73 degrees-76 degrees region of the *Bacillus subtilis* chromosome containing genes for trehalose metabolism and acetoin utilization. Microbiology. **142:**3057-3065.

51. **Yokoyama, K., K. Makino, Y. Kubota, M. Watanabe, S. Kimura, C. H. Yutsudo, K. Kurokawa, K. Ishii, M. Hattori, H. Abe, T. Iida, K. Yamamoto, T. Hayashi, T. Yasunaga, T. Honda, C. Sasakawa, and H. Shinagawa.** 2000. Complete nucleotide sequence of the prophage VT1-Sakai carrying the Shiga toxin 1 genes of the enterohemorrhagic *Escherichia coli* O157:H7 strain derived from the Sakai outbreak. Gene.**258:**127-139.

Table 4.1: Purification of PfpI functional forms from *Pyrococcus furiosus* cell extracts

		Protein concentration (mg/ml)	Activity (U/ml)	Specific Activity (U/mg)	Total Volume (ml)	Total Activity (U) x 10-6	Purification (fold)	Recovery (%)
	Cell extract	11.5	1432	125	420	601	1	100
	DEAE	10.1	2554	253	150	383	2.0	63.7
POOL 1 from HIC	**HIC**	10.6	4282	404	7	30.0	3.2	5.0
	MonoQ	3.7	4054	1096	5	20.3	8.8	3.4
	S200	0.0831	981	11805	100	98.1	94.4	16.3
POOL 2 from HIC	**HIC**	1.7	2965	1744	28	83.0	14.0	13.8
	HAP	0.8	3070	3837	6	18.4	30.7	3.1
	MonoQ	0.3	2167	7223	5	10.8	57.8	1.8
	S200	0.0118	353	29915	51	18.0	239.3	3.0

Table 4.2: Biochemical properties of functional forms of PfpI[a]

Form	Specific Activity (U/mg)		
	AAF-MCA	***LLVY-MCA***	***AFK-MCA***
Trimer, α_3	32,640	680	4,090
Hexamer, α_6	35,660	530	3,940
Dodecamer, α_{12}	93,620	1,190	12,060

[a]Purified from pool 2 of the HIC column; all assays done at 85ºC.

Table 4.3: Cleavage sites by PfpI on polypeptides

Polypeptide	Length (amino acids)	Cleavage sites (indicated by arrow)
Bradykinin	9	R-P-P-G-F-$^{\downarrow}$-S-P-F-$^{\downarrow}$-R
Angiotensin	10	D-R-V-Y-I-H-P-$^{\downarrow}$-F-$^{\downarrow}$-H-$^{\downarrow}$-L
Neurotensin	13	pE-L-$^{\downarrow}$-Y-$^{\downarrow}$-E-N-K-P-R-R-P-$^{\downarrow}$-Y-$^{\downarrow}$-I-$^{\downarrow}$-L
Adrenocorticotropic hormone, human segment 1-17	17	S-Y-$^{\downarrow}$-S-M-E-H-F-R-W-G-K-P-V-G-$^{\downarrow}$-K-$^{\downarrow}$-K-R
Adrenocorticotropic hormone, human segment 1-24	24	S-Y-S-M-E-H-F-R-W-G-K-P-V-G-K-K-R-R-P-V-K-V-Y-P No cleavage
Insulin chain A (oxidized ammonium salt)	21	G-I-V-E-Q-C-C-A-S-V-C-S-L-Y-Q-L-E-N-Y-C-N No cleavage
Insulin chain B (oxidized free acid)	30	F-V-N-Q-H-L-C-G-S-H-L-V-E-A-L-Y-L-V-C-G-E-R-G-F-F-Y-T-P-K-A No cleavage

Table 4.4: Heat-shock differential gene expression for PfpI and bacterial homologs

ID[a] (gene name)	Function (identity)[b]	Fold-change per heat-shock time point[c]			
		0-B	5-B	7-B	10-B
PF1719 (*pfpI*)	PfpI	3.3[d]	1.8[d]	---	0.7[d]
B3153 (*yhbO*)	*E. coli* HP (44%)	---	---	0.7, 1.2[e]	---
Bsu0785 (*yfkM*)	*B. subtilis* HP (47%)	---	---	---	5.9[f]

[a] ID = gene identifier

[b] HP = hypothetical protein; amino acid sequence identity to PfpI is in parentheses

[c] Column '0-B' represents fold-change upon heat-shock temperature increase (zero time point), column '5-B' is 5 min after temperature change, column '7-B' is 7 min after temperature change, column '10-B' is 10 min after temperature change.

[d] Numbers for PfpI represent fold-changes occurring at time points after the *P. furiosus* culture temperature was heat shocked from 90ºC to 105ºC; unpublished data.

[e] Data from http://www.genome.wisc.edu/, *E. coli* genome project; first number is heat shock using microarray data, second number is heat shock using radioactive labeling (37).

[f] Data from Petersohn et al. (35). The *yfkM* gene is homologous to a general stress response protein in *Bacillus halodurans*, BH2238 (35).

VII. FIGURE CAPTIONS

Figure 4.1: Differential scanning calorimetry of concentrated PfpI hexamer from HIC pool 2. MHC represents molar heat capacity.

Figure 4.2: Effect of temperature on specific activity (AAF-MCA) for the three forms of PfpI. Units are a fluorescence reading of 100 µl of the cooled products at a sensitivity of 3. Legend: diamond, dodecamer; square, hexamer; and triangle, trimer**.**

Figure 4.3: Northern analysis of the *P. furiosus pfpI* transcript. *P. furiosus* was grown with either 10 mM cellobiose (C), or 5 g/L tryptone with 1% S° (T) as the primary carbon and energy source.

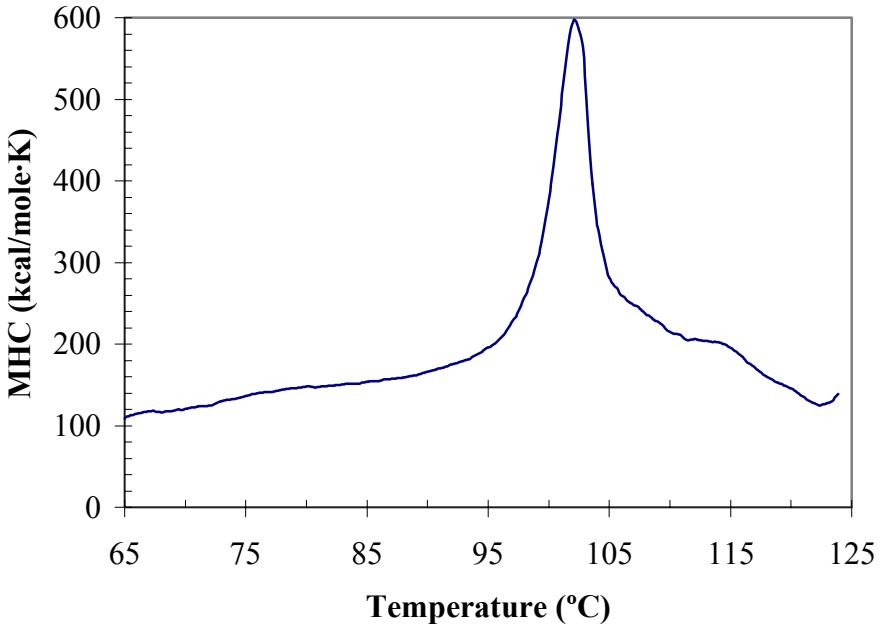


Figure 4.1: Differential scanning calorimetry of concentrated PfpI hexamer from HIC pool 1, showing the melting point at 102ºC. MHC represents molar heat capacity.

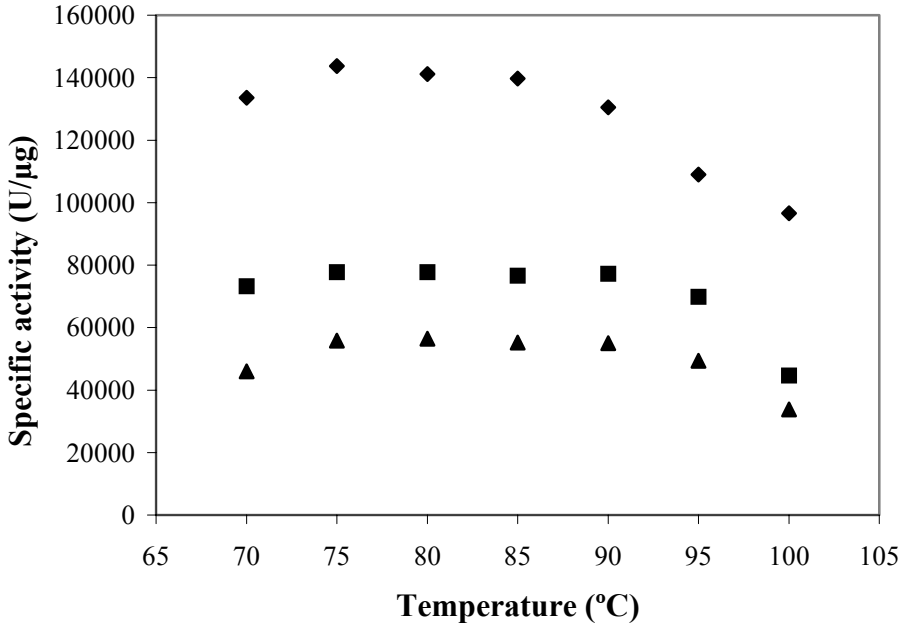


Figure 4.2: Effect of temperature on specific activity (AAF-MCA) for the three forms of PfpI. Units are fluorescence values for 100 µl of the cooled products at a sensitivity of 3. Legend: diamond, dodecamer; square, hexamer; and triangle, trimer**.**

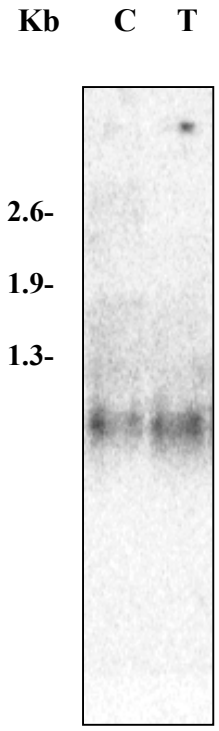


Figure 4.3: Northern analysis of the *P. furiosus pfpI* transcript. *P. furiosus* was grown with either 10 mM cellobiose (C), or 5 g/L tryptone with 1% S°(T) as the primary carbon and energy source.

CHAPTER 5:

An introduction to the proteasome: a self-compartmentalizing protein-destroying machine

I. INTRODUCTION

The proteasome is a protease found in all three domains of life (Dahlmann *et al*., 1989), including eukaryotes such as *Saccharomyces cerevisiae* (Groll *et al*., 1997; Hilt *et al*., 1993a) and humans (Tanaka *et al*., 1988), archaea such as *Thermoplasma* sp. (Dahlmann *et al.*, 1989), *Methanosarcina* sp. (Maupin-Furlow *et al*., 1998; Maupin-Furlow & Ferry, 1995), *Methanococcus* sp. (Wilson *et al*., 2000), and *Pyrococcus* sp. (Bauer *et al*., 1997), and bacteria such as *Escherichia* sp. (Chuang *et al*., 1993), *Rhodococcus* sp. (Tamura *et al*., 1995), and *Streptomyces* sp. (Nagy *et al*., 1998). The first observations of this protease, initially called "cylindrin," were reported in the late 1960's (Harris, 1968). However, it was not until the early 1980's that it was named the 'multicatalytic proteinase' (MCP) (Dahlmann *et al*., 1985; Wilk & Orlowski, 1980; Wilk & Orlowski, 1983) because of its ability to catalyze more than one distinct proteolytic reaction, including chymotrypin-like, trypsin-like, and peptidylglutamyl-peptide hydrolase (PGPH) activities (Wilk & Orlowski, 1983). Several years later, Arrigo et al. (1988) gave this protease its most recently known name, the 'proteasome,' in order to emphasize its role in the cell as a protein-destroying machine (Baumeister *et al*., 1998). Furthermore, it is important to realize that, since separate research groups observed many different forms of the proteasome around the same time, it actually has ~21 different names in literature (Gerards *et al*., 1998), including the prosome or 19S ribonucleoprotein (RNP) (Schmid *et al*., 1984), neutral protease (DeMartino & Goldberg, 1979), and MCP (Dahlmann *et al.*, 1985; Wilk & Orlowski, 1980; Wilk & Orlowski, 1983).

II. THE 20S PROTEASOME

The 20S proteasome is a 700-750-kDa cylinder-shaped protease arranged into four heptameric rings that are stacked in an axial fashion (Baumeister *et al.*, 1998; De Mot *et al*., 1999; DeMartino & Slaughter, 1999; Gerards *et al.*, 1998; Tanahashi *et al*., 1999; Tanaka, 1998). This barrel-like structure is a stack of four rings, made of two outer α rings and two inner β rings associated in the order $\alpha_7\beta_7\beta_7\alpha_7$. Depending on the source of the protease, the α and β rings are each made up of seven structurally similar or identical α and β subunits, respectively. The overall size of the 20S proteasome is 14.8 nm in length and 11.3 nm in diameter (Groll *et al.*, 1997; Lowe *et al*., 1995). It is hollow from end to end and contains three large inner cavities bounded by four narrow constrictions (Groll *et al.*, 1997; Hegerl *et al*., 1991; Lowe *et al.*, 1995). Substrates of this protease most likely enter at one end of the barrel structure through the central channel (Wenzel & Baumeister, 1995), whose opening is regulated by the N-terminal regions of the flanking α rings (Groll *et al*., 2000) and a separate protein complex, including one or more nucleotide triphosphatases (NTPases) (Gottesman *et al*., 1997; Larsen & Finley, 1997; Schmidt *et al*., 1999a). The hydrolysis of the substrate is achieved by the active sites buried in the central cavity and located on the β subunits (Groll *et al.*, 1997; Lowe *et al.*, 1995), a common characteristic of not only the proteasome, but several other self-compartmentalizing protease structures (Baumeister *et al.*, 1998; Lupas *et al*., 1997a) such as tricorn (Tamura *et al*., 1996) and the Clp family (Chung *et al*., 1997; Gottesman, 1996; Suzuki *et al*., 2002). In general, this compartmentalization of proteinases is hypothesized to be a protective mechanism for the cell, through careful localization and control of proteolytic activity. Although it is known that only peptides and unfolded

polypeptides can enter the catalytic channel (Wenzel & Baumeister, 1993), the precise mechanism of translocation and hydrolysis of these substrates remains to be determined. A schematic of the overall 20S proteasome structure is shown in Figure 5.1.

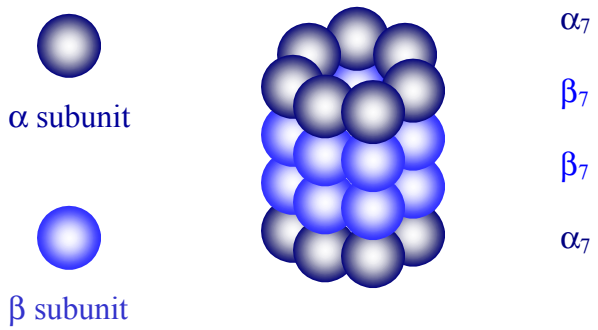


Figure 5.1: Schematic drawing of the 20S proteasome showing the characteristic $\alpha_7\beta_7\beta_7\alpha_7$ barrel structure conserved through all domains of life (Dahlmann *et al.*, 1989).

In all forms of the proteasome studied to date, the active site for hydrolysis of peptides and proteins is located on one or more of the β subunits located in the two inner β_7 rings of the barrel-like structure (Seemuller *et al.*, 1995). As found in several studies of the proteasome, the bacterial (Zuhl *et al.*, 1997a), archaeal (Maupin-Furlow *et al.*, 1998; Maupin-Furlow & Ferry, 1995; Seemuller *et al.*, 1996; Wilson *et al.*, 2000; Zwickl *et al.*, 1994), yeast (Arendt & Hochstrasser, 1999; Chen & Hochstrasser, 1996; Ditzel *et al.*, 1998), and mammalian (Schmidtke *et al.*, 1996) forms all contain at least one β subunit that initially exists as a prosubunit and is later processed, through autocatalytic removal of the N-terminal residues immediately upstream of the active-site Thr (Kisselev *et al.*, 2000; Lowe *et al.*, 1995; Seemuller *et al.*, 1995), during assembly with the α subunit. The side chain amino

groups Lys-33 and Asp/Glu-17 may accept the side chain Thr proton in a charge-relay system through a salt bridge, as found in *T. acidophilum* (Seemuller *et al.*, 1996). The active-site Thr, which is exposed to become the first N-terminal residue (Thr^1) upon removal of the prosequence, then acts as the primary catalytic site during proteolysis. Specifically, the γ-oxygen acts as the nucleophile and the α-amino group most likely acts as the proton acceptor (Fenteany *et al*., 1995; Groll *et al.*, 1997; Lowe *et al.*., 1995; Maupin-Furlow *et al.*, 1998; Seemuller *et al.*, 1995), with the precise mechanism hypothesized to be similar to that of serine proteases (Maupin-Furlow *et al*., 2001). Because of this observed activity by the β subunits, the proteasome has been classified to the superfamily of N-terminal nucleophile hydrolases, or Ntn-hydrolases, which characteristically contain a residue (Ser, Cys, or Thr) at one terminus that acts as the catalytic nucleophile (Brannigan *et al*., 1995; Dodson & Wlodawer, 1998).

The propeptide hydrolyzed from the Thr^1, which can be as small as 8-10 residues as in *T. acidophilum* (Seemuller *et al.*, 1996) or >50 residues as in human LMP7 (Schmidt *et al*., 1999b; Witt *et al*., 2000), was found to be important, though not essential, for proper assembly of the active sites of the proteasome. However, assembly and role of β subunits has been found to differ among the proteasomes from the three domains of life (Gerards *et al.*, 1998). In the case of the bacterial *Rhodococcus erythropolis* proteasome, absence of the propeptide significantly lowered the assembly efficiency of the subunits (Zuhl *et al.*, 1997a). In the archaeal *T. acidophilum* proteasome, the propeptide was less critical for proper assembly (Grziwa *et al*., 1994), and in yeast proteasomes, the propeptides of certain β subunits were critical in folding and incorporation of these subunits into the mature proteasome, though they had no role in correct positioning within the structure (Arendt &

Hochstrasser, 1999; Baker *et al*., 1993; Chen & Hochstrasser, 1996; Ramos *et al*., 1998; Shinde & Inouye, 1994).

Proteasomes from all domains of life are multicatalytic proteases characterized by three primary activities with distinct specificities against short synthetic peptides. The first is a chymotrypsin-like activity with a preference for tyrosine or phenylalanine (bulky hydrophobic groups) at the P_1 position. The second is a trypsin-like activity with a preference for arginine or lysine (basic groups) at the P_1 position. The third is a postglutamyl or PGPH activity with a preference for glutamate or other acidic residues at the P_1 position (Wilk & Orlowski, 1983). In only the eukaryotic forms, two other peptidase activities were found, as well, including hydrolysis between small neutral amino acids (SNAAP) and after branched chain residues (BrAAP) (Mykles, 1996; Orlowski *et al*., 1993). In the bacterial and archaeal versions of the proteasome, similar types of activity compared to the eukaryotic proteasome have been reported, however, it was found that the chymotrypsin-like activity was significantly higher than trypsin or PGPH activity for several of these proteasomes. In particular, this preferential activity was found for the bacterial proteasomes from *R. erythropolis* (Tamura *et al.*, 1995) and *S. coelicolor* (Nagy *et al.*, 1998), and the archaeal proteasomes from *T. acidophilum* (Dahlmann *et al*., 1992) and *Haloferax volcanii* (Wilson *et al*., 1999). Unlike the other archaeal proteasomes, however, the methanoarchaeal proteasomes characterized to date were found to have high levels of both chymotrypin-like and PGPH activities (Maupin-Furlow *et al.*, 1998; Wilson *et al.*, 2000).

III. THE 26S PROTEASOME

Although the 20S proteasome can be activated *in vitro*, it may never act as an isolated enzyme in cells (DeMartino & Slaughter, 1999). Instead, the 20S proteasome must be activated by regulatory proteins, many of which are NTPases in the AAA family (ATPases associated with various cellular activities) or AAA^+ superfamily (Neuwald *et al*., 1999) and a few of which are ATP-independent regulators. In mammalian and yeast cells, the primary NTPase regulatory complex (RC) is known as ball (Hoffman *et al*., 1992), μ-particle (Udvardy, 1993), PA700 (Chu-Ping *et al*., 1994), or 19S cap (Peters *et al*., 1994), while the ATP-dependent regulator is known as PA28 (Ma *et al*., 1992) or 11S (Dubiel *et al*., 1992). The NTPase RC, most commonly referred to as the 19S regulator (Voges *et al*., 1999), is a 700-1000-kDa, 18-20 subunit complex (most recent literature is conflicting on the exact number) that can bind to one or both of the terminal α rings in the 20S proteasome in an ATP-dependent and cooperative manner (Armon *et al*., 1990; DeMartino *et al*., 1994; Hoffman & Rechsteiner, 1994; Peters *et al.*, 1994; Tanahashi *et al.*, 1999). The reaction in which the 19S RC associates itself with the 20S proteasome requires ATP hydrolysis, an inherent function of the 19S regulator (DeMartino & Slaughter, 1999). Each subunit of the RC is in the range of 25-110 kDa in size (Tanahashi *et al.*, 1999; Tanaka, 1998). Once a regulatory complex is attached to each end of the 20S proteasome, the resulting complex is called the 26S proteasome. A schematic of the overall structure of the eukaryotic 26S proteasome is illustrated in Figure 5.2.

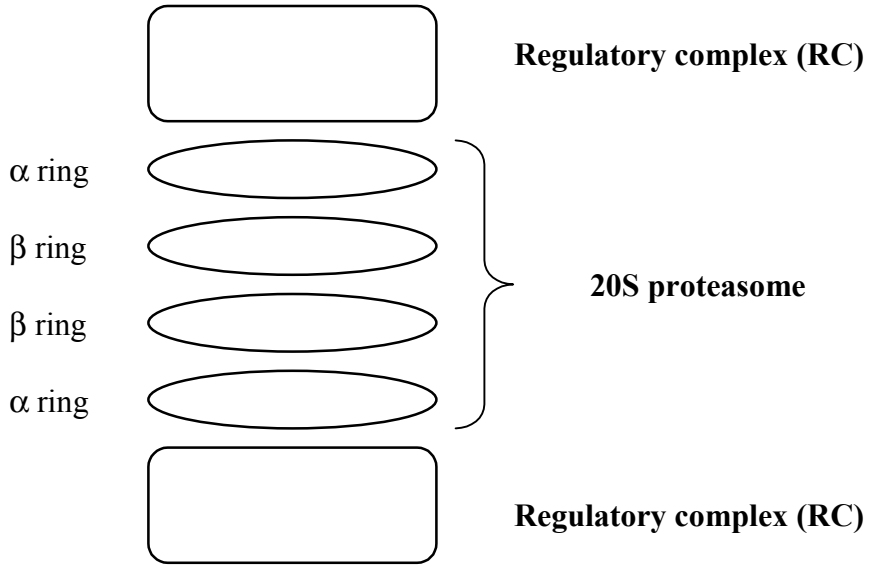


Figure 5.2: Schematic drawing of the 26S proteasome.

In bacterial and archaeal cells, the regulation of the proteasome is not clear, though several candidates for energy-dependent regulation are currently being studied. Structures have been viewed in transmission electron micrographs that indicate interactions occurring between the *M. jannaschii* 20S proteasome and an ATPase that was found to stimulate proteolytic activity (Wilson *et al.*, 2000). However, these interactions did not appear in a high percentage of the population but were instead transitory. There are no other reports of prokaryotic 26S proteasomes with regulatory proteins associated in a stable fashion, either purified or assembled from recombinant subunits.

In the 2.5-MDa, 45-nm long, 20-nm wide (Fujinami *et al*., 1994; Peters *et al*., 1993; Yoshimura *et al*., 1993) eukaryotic 26S proteasomes with doubly-bound 19S RC's (Voges *et al.*, 1999), the RC's are faced in the opposite direction, which implies that the contacts

between the RC's and the α rings of the 20S proteasome are specific (Gorbea *et al*., 1999). Each 19S RC is composed of two parts, a "base" and a "lid," with the base acting as the connector between the 19S and 20S portions of the 26S structure (Glickman *et al*., 1998). However, this connection is relatively flexible, as the entire structure composed of 19S – 20S – 19S linked end-to-end has been observed to undergo a "wagging" motion, with the movements of each of the caps appearing random in comparison to each other (Walz *et al*., 1998). In humans and yeast, the 19S "base" is made up of 6 distinct and highly homologous ATPases (Glickman *et al.*, 1998), which are all in the AAA family (Beyer, 1997; Confalonieri & Duguet, 1995) and separately named with an S/Rpn nomenclature. These ATPases assemble into a heteromeric ring that may act as the interface between the 20S proteasome core and the RC (Baumeister *et al.*, 1998; Gorbea *et al.*, 1999; Tanahashi *et al.*, 1999). In addition, the base contains two other large subunits (S1/Rpn2 and S2/Rpn1) as well as a third subunit (S5a/Rpn10) (Glickman *et al.*, 1998). One role of the ATPase is to continuously supply energy for the degradation of target proteins, since the association of the 19S base with the 20S proteasome allows ATP-dependent degradation of non-ubiquitinated proteins (Zwickl *et al*., 1999b). However, it is unknown why the 26S proteasome complex contains multiple homologous ATPases (Tanahashi *et al.*, 1999). The remaining RC subunits, including the eight contained in the "lid" of the 19S RC, are non-ATPases and, other than one that binds polyubiquitin (polyUb) and a second that may be an isopeptidase, have unknown function (Glickman *et al.*, 1998; Gorbea *et al.*, 1999). It is assumed that some of these must have polyUb-binding capabilities, as the single subunit that has been identified with this ability is not essential for protein degradation and growth of yeast (Gorbea *et al.*, 1999). A three-dimensional model of the 26S proteasome, illustrating the 19S – 20S – 19S

linkage and the observed "wagging" motion, along with a list of corresponding subunits for the 19S RC are shown in Figure 5.3.

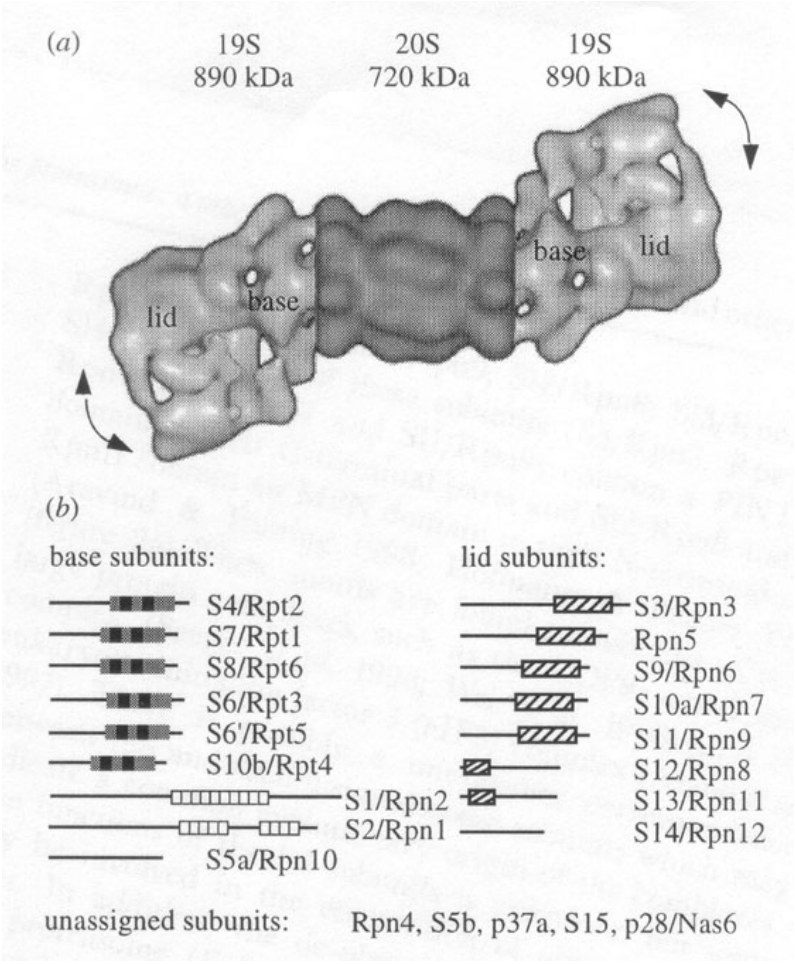


Figure 5.3: (a) Model of the three-dimensional eukaryotic 26S proteasome based on electron microscopy and the crystal structure of the *T. acidophilum* 20S proteasome. The masses of the 19S and 20S structures are listed directly above each structure. Arrows represent the "wagging" motion of the 19S RC's when attached to the 20S portion.
(b) Listed subunits of the base and lid complexes that make up the 19S RC. Figure from Zwickl et al. (1999b).

The eukaryotic 26S proteasome exhibits at least five distinct biochemical activities: 1) polyubiquitin (polyUb) chain recognition and binding, 2) nucleotidase, 3) isopeptidase, 4) unfoldase, and 5) endoproteinase (Armon *et al.*, 1990; Deveraux *et al*., 1994; Eytan *et al*., 1993; Hoffman & Rechsteiner, 1996; Kanayama *et al*., 1992; Lam *et al*., 1997; Ugai *et al*., 1993). The free 20S proteasome may have some proteinase/peptidase activity (Gorbea *et al.*, 1999), though this activity is relatively low without a regulatory protein to aid in displacing the α N-terminal gates that block the central channel leading to the active sites (Groll *et al.*, 1997). Within the cell, the proteasome is a primary participant in the non-lysosomal pathway of both ATP-dependent degradation of naturally occurring unstable regulatory proteins and the unregulated degradation of long-lived proteins (Gerards *et al.*, 1998; Tanaka, 1998). In addition, the proteasome is essential for the rapid removal of proteins that have become unassembled or misfolded as a result of mutations or environmental stresses, including heat stress, oxidation, and exposure to heavy metals (Tanaka, 1998). In eukaryotic cells, the selective proteasomal breakdown of these inactive and harmful proteins is mediated by a regulated pathway, termed the Ub (ubiquitin)-proteasome pathway or the ubiquitin (Ub)-pathway (Ciechanover, 1994; DeMartino & Slaughter, 1999; Hershko, 1996). Furthermore, it is believed that this Ub-pathway may be triggered by the initial phosphorylation of inactive proteins or other similar protein modifications. At this point, the literature lacks any proposed mechanisms to explain how phosphorylation leads to the Ub-pathway. Although Wolf et al. (1993) and Durner et al. (1995) reported the presence of ubiquitin in the archaeon *T. acidophilum* and the bacterium *Anabaena variabilis*, respectively, these reports have never been confirmed. In addition, ubiquitin or homologs of ubiquitin-conjugating enzymes have not been found in the archaea and bacteria studied to date. As such, it is widely accepted that

ubiquitin only exists in the eukaryotes. From personal communication with Dr. Baumeister (April, 2003), one possible explanation for the reported finding of ubiquitin in *T. acidophilum* in his laboratory (Wolf *et al.*, 1993) was the possibility that the identified ubiquitin came from the growth medium containing the cell extracts.

The primary component of the Ub-pathway is ubiquitin (Ub), an 8.6-kDa polypeptide that eukaryotic cells use to modify proteins that must be targeted for degradation by the proteasome. A full history of the Ub-pathway is reviewed by Hershko (1996). Normally, a chain of several Ub molecules is covalently attached to these targeted proteins (e.g., unassembled, misfolded, or mutated proteins) in an ATP-dependent fashion and, subsequently, serves as a proteasome recognition signal. In fact, the recognition of the poly-Ub chain attached to the protein is a key process in the selective degradation by the 26S proteasome (Tanahashi *et al.*, 1999). The addition of several Ub polypeptides to a target protein, often termed polyubiquitination, occurs through the sequential activation of three enzymes. The first is ubiquitin-activating enzyme (E1), the second is ubiquitin-conjugating enzyme (E2), and the third is ubiquitin-protein ligase (E3). Together, these three enzymes create a cascade of events that end in the complete ubiquitination of proteins targeted for degradation. Tanaka (1998) and Tanahashi et al. (1999) provide more detailed descriptions of this cascade. Figure 5.4 illustrates the basic scheme of the Ub-pathway, which ends in the degradation of the substrate-tagged poly-Ub chain by the 26S proteasome.

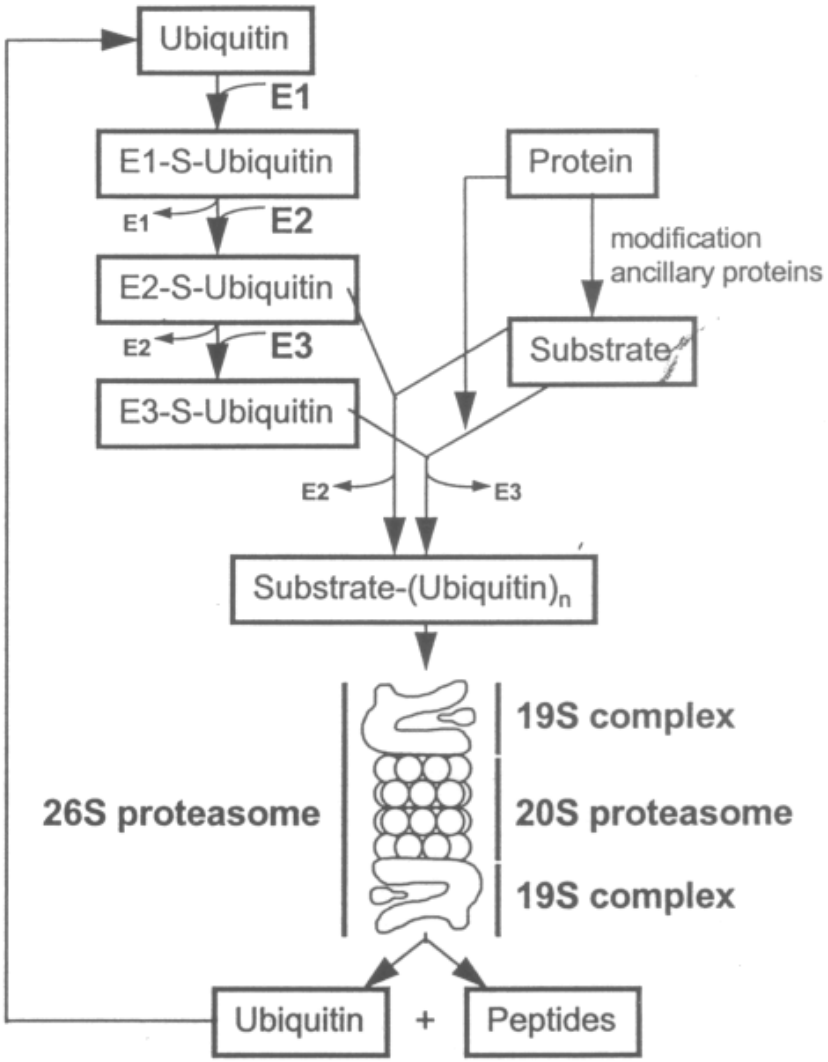


Figure 5.4: Schematic representation of the Ub (ubiquitin)-proteasome pathway, also called the ubiquitin (Ub) pathway. Ub is initially activated by enzyme E1 (Ub-activating), and passed on to enzyme E2 (Ub-conjugating). Ub is then linked directly, or with the help of enzyme E3 (Ub-ligase), through an isopeptide bond to a lysine residue of the substrate protein. The polyubiquitinated proteins are recognized and selectively degraded by the 26S proteasome. The product of degradation includes reusable Ub molecules and peptides that contain 5-15 amino acids each. Reconstructed from Gerards et al. (1998).

In addition to protein ubiquitination, it appears that there is an equally complex system for the de-ubiquitination of enzymes (Tanaka, 1998). In particular, as many as 17 genes were found to encode proteins with conserved catalytic sites for de-ubiquitinase in yeast (Tanaka, 1998). These de-Ub enzymes are used for trimming of abnormal poly-Ub structures, proofreading of incorrectly ubiquitinated proteins, or other similar functions (Tanaka, 1998). However, the reason for the presence of so many of these types of enzymes in the cell is not yet understood.

IV. ARCHAEAL PROTEASOMES

The first characterized archaeal proteasome, which was that isolated from *T. acidophilum* (Dahlmann *et al.*, 1992), became the "prototype" for the quaternary structure and topology of this protease (DeMartino & Slaughter, 1999). In fact, it is still used as a comparative basis for nearly every proteasome isolated from all three domains of life, both in terms of 20S and 26S structures. The α and β subunits that make up the *T. acidophilum* 20S proteasome are, literally, the true α and β subunits of the 'general 20S proteasome.' Therefore, all homologous proteasomal structures that have been characterized in the other domains are described as 'α- and β-like.' This terminology is used throughout the literature as an assumed comparison to the α and β subunits of the *T. acidophilum* proteasome.

The 20S proteasome has been isolated from several archaeal species, including *T. acidophilum* (Dahlmann *et al.*, 1992; Lowe *et al.*, 1995), *Methanosarcina thermophila* (Maupin-Furlow *et al.*, 1998; Maupin-Furlow & Ferry, 1995), *Methanococcus jannaschii*

(Wilson *et al.*, 2000), *H. volcanii* (Wilson *et al.*, 1999), and *Pyrococcus furiosus* (Bauer *et al.*, 1997). Although there is extensive sequence and structural information about the *T. acidophilum* proteasome (Akopian *et al.*, 1997; Dahlmann *et al.*, 1992; Grziwa *et al.*, 1991; Grziwa *et al.*, 1994; Hegerl *et al.*, 1991; Kisselev *et al.*, 1998; Kisselev *et al.*, 2000; Lowe *et al.*, 1995; Puhler *et al.*, 1992; Ruepp *et al.*, 1998; Seemuller *et al.*, 1995; Stock *et al.*, 1995; Wenzel & Baumeister, 1993; Zwickl *et al.*, 1992a; Zwickl *et al.*, 1994; Zwickl *et al.*, 1992b) along with important characterization reports on the other archaeal proteasomes, the physiological role and the physiological form of this protease remains unknown. Preliminary experiments have shown that the proteasome is essential for *T. acidophilum* under stress conditions (Ruepp *et al.*, 1998). Additionally, more recent genomic information has shown the presence of proteasome-like genes in several more archaeal species, including other pyrococcal species, methanogens, and sulfolobus species. A full list of hypothetical and characterized archaeal proteasome subunits, their gene names and numbers, and their similarity to the *P. furiosus* proteasome, is illustrated in Table 6.1. Generally, the archaeal forms tend to contain two to three forms of the α and β subunits, with the *H. volcanii* form containing two α, and several other forms containing two β subunits. In the cases where more than one α or β exists, the roles of the various subunit forms remain to be determined.

Like all 20S proteasomes, the basic structure of the archaeal form is the barrel shape, created by sets of subunits. In many of the characterized the archaea, there are two sets of subunits, termed α and β. Between the two species *T. acidophilum* and *M. thermophila*, there is a 46-60% sequence identity in encoding genes for the α and β subunits (Maupin-Furlow & Ferry, 1995). These subunits assemble as heptameric rings, which then assemble into stacks of four, as in Figure 5.1: $\alpha_7\beta_7\beta_7\alpha_7$. The rings are tightly packed, with no

openings in the side walls of the cylindrical structure. The only path from the outside of the structure to the inner cavities is through a 13-Å-diameter gate at both ends of the cylinder (Gerards *et al.*, 1998). The archaeal *T. acidophilum* (Seemuller *et al.*, 1996; Zwickl *et al.*, 1994), *M. thermophila* (Maupin-Furlow *et al.*, 1998), and *M. jannaschii* (Wilson *et al.*, 2000) forms (all containing one version of α and one version of β) were found to have an assembly order that started with fully assembled α_7 rings. These rings then provided the matrix onto which folding and processing of the β subunits occurred. This was the most likely scenario since, in all three archaeal proteasomes, the β subunits were unable to process the prosequence or produce full β_7 rings without the presence of the α subunit.

As is the case with the eukaryotic proteasome, several archaeal versions can expand their proteolytic capacity through the use of ATP, mediated through at least one known ATPase, referred to as PAN (proteasome-activating nucleotidase) (Benaroudj & Goldberg, 2000; Benaroudj *et al*., 2003; Navon & Goldberg, 2001; Wilson *et al.*, 2000; Zwickl *et al*., 1999a). To date, the only archaeal PAN to be characterized is that from *M. jannaschii* (Benaroudj & Goldberg, 2000; Benaroudj *et al.*, 2003; Navon & Goldberg, 2001; Wilson *et al.*, 2000; Zwickl *et al.*, 1999a). In general, the archeal PAN is hypothesized to play some of the same roles as the 19S, PA28, and other proteasome regulatory complexes from the eukaryotes because of its predicted protein sequence similarity to the eukaryotic 19S regulatory complex (Bult *et al*., 1996; Voges *et al.*, 1999) and its observed characteristics that are reminiscent of the 26S regulatory complexes. Since the archaea and bacteria have not been found to contain ubiquitin or homologs of ubiquitin-conjugating enzymes (Zwickl *et al.*, 1999a), the most likely alternative to the ubiquitin-dependent system is the presence of the ATP-dependent regulatory proteins such as PAN and FtsH / cell-division control (CDC48)

proteins. It is unknown whether a separate system with a series of participating proteins like the ubiquitin pathway exists in the archaea and bacteria or if, instead, the PAN, VAT (valosine-containing protein-like ATPase) (Zwickl *et al*., 2000), CDC48 (Maupin-Furlow *et al.*, 2001; Zwickl *et al.*, 2000), and other ATP-dependent proteins are in control of the proteasome degradation pathway, including substrate targeting. It has already been shown that the *M. jannaschii* PAN is able to associate with the ends of the *M. jannaschii* 20S proteasome in the presence of ATP (Wilson *et al.*, 2000). Furthermore, it has the ability to stimulate 20S proteolysis of polypeptides (but not small peptides (Zwickl *et al.*, 1999a), unlike the eukaryotic versions), such as β-casein (Wilson *et al.*, 2000; Zwickl *et al.*, 1999a), α-lactalbumin (Zwickl *et al.*, 1999a), and green fluorescent protein containing an 11-residue ssrA recognition peptide "tag" (Benaroudj & Goldberg, 2000) in the presence of both ATP and CTP (Benaroudj & Goldberg, 2000; Wilson *et al.*, 2000; Zwickl *et al.*, 1999a). Interestingly, the *M. jannashcii* PAN was found to interact and stimulate both the *M. jannaschii* (Benaroudj & Goldberg, 2000; Wilson *et al.*, 2000) and *T. acidophilum* (Zwickl *et al.*, 1999a) proteasomes. It was also found to unfold protein substrates for the proteasome (Navon & Goldberg, 2001), regulate the N-terminal "gate" that is present in the outer α ring of the proteasome, and control translocation of the substrate into the catalytic center of the proteasome (Benaroudj *et al.*, 2003), similar to the mechanism for translocation used by the eukaryotic PA28 (Whitby *et al*., 2000). It was initially believed that the N-terminal "gate" of the α ring was always in an open state in the archaeal versions of the proteasome, based on crystal structure data (Lowe *et al.*, 1995) and the observation that these versions were able to digest small peptides without the presence of ATPases (Zwickl *et al.*, 1999a). However, it was recently found that the archaeal proteasomes do contain a gate that is not necessarily in a

completely open state under normal conditions, proven by the ability of the proteasome to block entry of unfolded proteins in the absence of the PAN ATPase (Navon & Goldberg, 2001). As shown in Table 6.1, many of the sequenced archaea contain 1-2 genes that encode for sequences similar to PAN. However, several have not been found to contain a PAN, specifically the *Thermoplasma* species (Ruepp *et al*., 2000) and *Pyrobaculum aerophilum* (Zwickl *et al.*, 1999a). Though unpublished on the NCBI site, it was also recently found that *H. volcanii*, like the other listed halophilic archaeon, *Halobacterium* sp. NRC-1 (Ng *et al*., 2000), does encode two PAN paralogs (Maupin-Furlow *et al.*, 2001). In the case of the archaeal species that do not appear to contain PAN, other similar proteins may be key players in the proteasome degradation pathway. These other proteins include VAT (valosine-containing protein-like ATPase) (Pamnani *et al*., 1997; Zwickl *et al.*, 2000) and FtsH/CDC48 (Maupin-Furlow *et al.*, 2001; Zwickl *et al.*, 2000), which have been found through gene analysis or biochemical studies in *Thermoplasma* (Ruepp *et al.*, 2000) and *P. aerophilum* (Zwickl *et al.*, 1999a). These proteins have also been found in other archaeal species that contain PAN (Kawarabayasi *et al*., 1999; Kawarabayasi *et al*., 1998). It remains to be determined exactly how these regulating enzymes interact with the archaeal proteasome and with each other and, furthermore, how closely related this system of regulation is to the ubiquitin system of the eukaryotes.

V. EUKARYOTIC PROTEASOMES

The 26S proteasome is confined to the cytosolic and nuclear compartments of all eukaryotic cells (Palmer *et al*., 1996; Reits *et al*., 1997). The eukaryotic 20S proteasome has a three-dimensional structure that is similar to the basic proteasome described in Figure 5.1. However, unlike the bacterial and archaeal proteasomes, there are at least seven different α- and seven different β-type subunits that make up the eukaryotic 20S proteasome. The lower eukaryotes such as the yeast *S. cerevisiae* generally contain 7 α-type and 7 β-type subunits, as determined from sequence (Clayton *et al*., 1997) and biochemical information (Heinemeyer *et al*., 1994; Hilt *et al.*, 1993a). The higher eukaryotes such as mammals and *Arabidopsis thaliana* (Fu *et al*., 1998) have been found to encode for as many as 13 α-type and 10 β-type subunits. Within the stack of four heptameric rings, each ring contains a complete complement of the seven α- or β-type subunits (DeMartino & Slaughter, 1999). The proposed model of the human 20S proteasome, based on the structure of the yeast 20S proteasome (Groll *et al.*, 1997), is shown in Figure 5.5.

The α-type subunits chaperone the folding, processing, and assembly of the β-type subunits and they define the polypeptide channel that controls access to the interior of the proteasome (Zuhl *et al.*, 1997a). However, the assembly of the proteasome is still not fully understood (Gerards *et al.*, 1998). For the eukaryotic 20S proteasome, the proposed assembly is similar to that of the archaeal proteasomes: a single α/β double ring is initially formed by co-assembly of the α- and β-type subunits. Then, once two sets of these half-proteasomes are formed, they associate to make a fully-stacked, four-level proteasome

(Gerards *et al.*, 1998; Gerards *et al*., 1997). The primary difference between the eukaryotic and archaeal proteasome assembly patterns, however, is during the formation of the half-proteasomes. While the archaeal α subunits assemble into isolated rings before associating with the β subunits (Seemuller *et al.*, 1996; Zwickl *et al.*, 1994), the eukaryotic α-type subunits appear to associate with the β-type subunits while assembling into rings. Furthermore, although processing of the β propeptide appears to be a dispensable step in the archaeal *T. acidophilum* proteasome assembly pathway (Zwickl *et al.*, 1992b), it has been found to be a critical step in the proper assembly of eukaryotic versions of the proteasome (Chen & Hochstrasser, 1996). It was found that this critical β subunit processing is most likely a late event in the assembly of the eukaryotic proteasome, as structures that were most likely the half-proteasome were found to contain β subunits with their prosequence still attached (Frentzel *et al*., 1994; Schmidtke *et al.*, 1996). A schematic of the eukaryotic assembly pattern is illustrated in Figure 5.6.

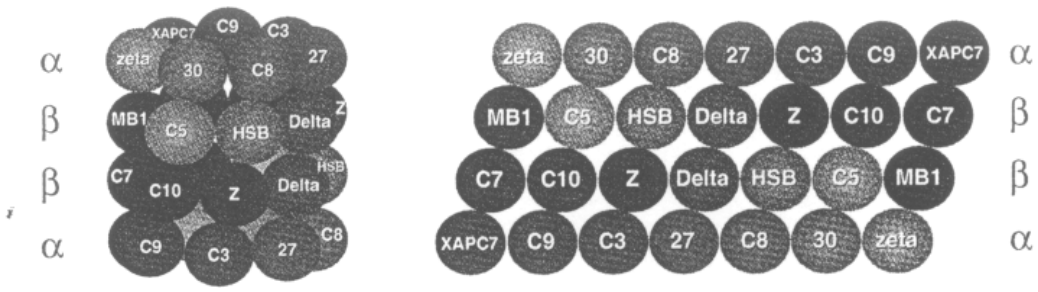


Figure 5.5: Proposed schematic representation of the human 20S proteasome, based on the yeast crystal structure (Groll *et al.*, 1997). On the left is the three-dimensional form of the proteasome. On the right is the rolled-out cylinder envelope, which shows the proteasomal subunit topology. The label on each sphere represents an abbreviated form of each gene encoding for each α- or β-type subunit. Recreated from Gerards et al. (1998).

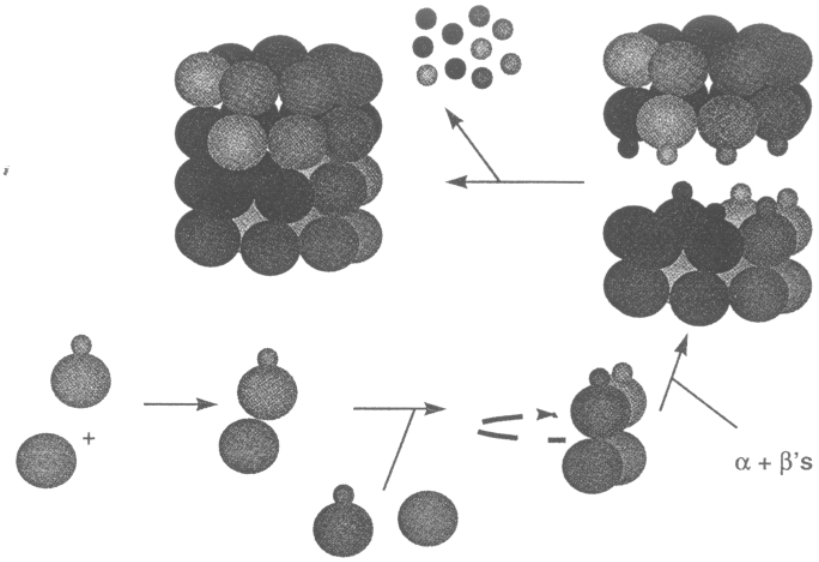


Figure 5.6: A proposed model for the assembly pattern of the eukaryotic 20S proteasome. The α/β double ring, or half-proteasome, is formed in a coordinated fashion by co-assembly of the α- and β-type subunits. The two half-proteasomes then form into a four-level 20S proteasome. The single large spheres represent the α-type subunits and the large spheres with attached small spheres represent the β-type subunits. The small spheres represent possible chaperones used to keep the β-type subunits at the interface of the two half-proteasomes in the proper conformation until the complete proteasome is formed. These chaperones may dissociate from the β subunits once the full proteasome is assembled. Recreated from Gerards et al. (1998).

From comparisons between organisms both from different domains and within the same domain, the complexity of the proteasome appears to mirror the complexity of the organism. For example, yeasts contain seven genes encoding β-type subunits (Clayton *et al.*, 1997), but higher eukaryotes, such as mammals have ten β-subunit encoding genes (DeMartino & Slaughter, 1999; Fu *et al.*, 1998; Gerards *et al.*, 1998). The three extra genes in the mammalian cells are homologous to three of the β-type genes whose products help to form the typical 20S proteasome. However, these genes are γ-interferon-inducible and their products can each replace their corresponding β-type 20S proteasome subunit to create a proteasome with altered proteolytic activity. These γ-interferon-inducible subunits may function in the generation of antigenic peptides (DeMartino & Slaughter, 1999; Gerards *et al.*, 1998).

Another example of increased proteasome complexity with increased cell complexity is found in the regulation of the proteasome. Mammalian cells, but not yeast cells, contain, in addition to the 19S cap, a second proteasome activator termed PA28 or the 11S regulator (Dubiel *et al.*, 1992; Ma *et al.*, 1992). This activator is composed of two 28-kDa subunits, α- and β-type, that are 50% identical in primary structure. These subunits combine further to form a ring-shaped molecule that is approximately 180 kDa in size. The exact quaternary structure is not known, but is proposed to be heteroheptameric or, more likely, heterohexameric (Ahn *et al*., 1996; Kuehn & Dahlmann, 1996; Song *et al*., 1996).

Like the 19S regulatory complex, PA28 binds to the α-subunit ends of the 20S proteasome (Ahn *et al.*, 1996; Koster *et al*., 1995; Kuehn & Dahlmann, 1996). However, it does not require ATP or any other cofactor to do so. Furthermore, it does not activate the 20S proteasomal degradation of large proteins, whether the protein is ubiquitinated or not.

The proposed reason for this is the probable inability of PA28 to unfold and/or translocate these larger substrates in an ATP-dependent manner. Although recombinant PA28α (the α subunit of PA28) is sufficient for proteasome activation (Dick *et al*., 1996; Knowlton *et al*., 1997), it is thought that neither the PA28α nor PA28β (the β subunit of PA28) exist alone in cells. It has been found that the isolated PA28α activation is much less efficient than the native heteromeric PA28 activation. After binding of the PA28 complex to the 20S proteasome, the regulatory complex may work to activate the proteasome by opening the channel at the terminal rings. This increases the access of the substrates to the inner catalytic sites and may allosterically activate the catalytic sites. The regulation of proteasome activity in mammalian cells is illustrated in Figure 5.7. As shown in the figure, the PA28 complex can activate the 20S proteasome alone or with the 19S regulatory complex.

The physiological role of PA28 is unknown at this point. Since it is absent in yeast, it may not be essential for general ubiquitin-dependent proteolysis. However, it may participate in a related reaction that supplements the ubiquitin-dependent pathway or play a role completely unrelated to the ubiquitin-dependent pathway (Gorbea *et al.*, 1999).

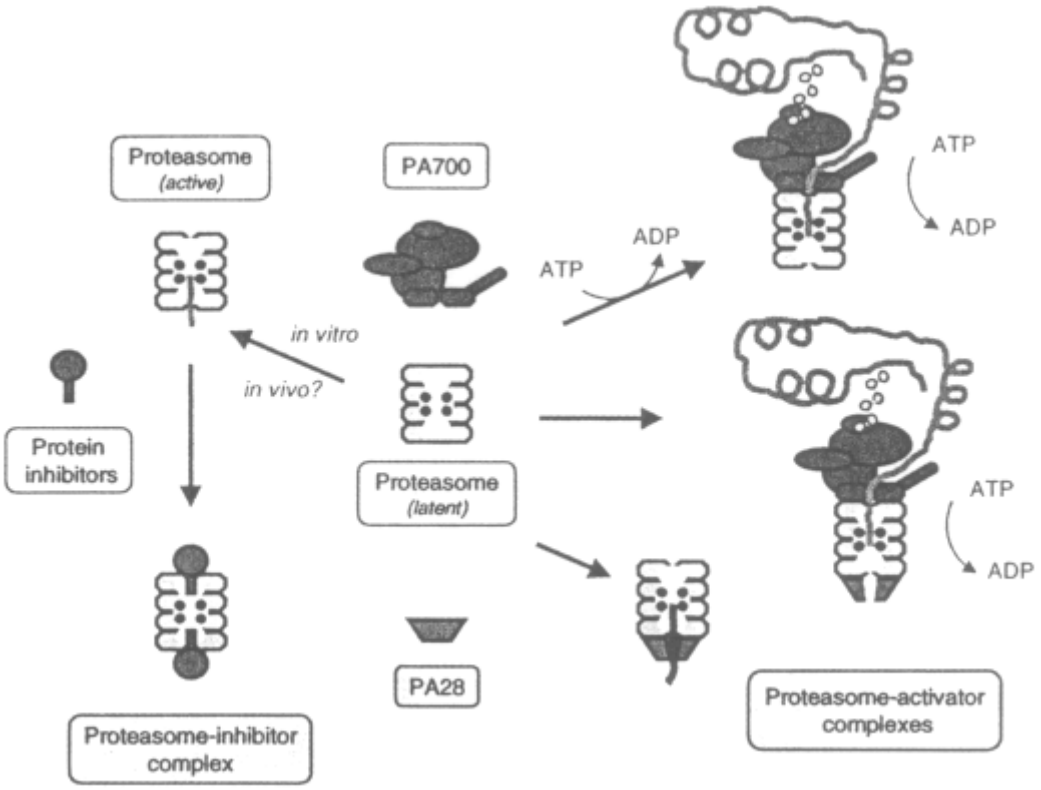


Figure 5.7: Regulation of mammalian 20S proteasome activity by regulatory proteins. The 20S proteasome is essentially inactive because its structure isolates the catalytic sites (depicted by black circles (●)) from substrates outside of the complex. Binding of the 20S proteasome to regulatory proteins, such as PA28 or the 19S regulatory complex (PA700), activates degradation of ubiquitinated peptides and proteins, respectively. Part of the regulation may involve opening of channels at the terminal rings to increase substrate access to the catalytic sites. For clarity, the figure depicts proteasomes capped by a single PA28 or PA700 RC, however, each activator can doubly cap one 20S proteasome particle. Protein inhibitors can inactivate the 20S proteasome that may become activated in the cell and may also inhibit formation of proteasome-activator complexes. Recreated from DeMartino and Slaughter (1999).

VI. THE YEAST PROTEASOME

In the yeast *S. cerevisiae*, there are 14 different known genes that may code for the complete set of 20S proteasomal units (Heinemeyer *et al.*, 1994). These can further be divided into 7 α- and 7 β-type subunits. This division was made according to the degree of homology of these subunits to the α and β subunits of the *T. acidophilum* proteasome (Hilt *et al*., 1993b). The yeast proteasome was discovered in 1984 and originally named yscE (Achstetter *et al*., 1984) until, in 1988, it showed homology to a 20S cylindrical proteasome from another organism, *Xenopus laevis* (Kleinschmidt *et al*., 1988). It exhibits the typical proteasomal activity, including trypsin-like, chymotrypsin-like, and PGPH activities (Achstetter *et al.*, 1984). The 7 α-type genes encode for members of a subfamily of proteasomal subunits that have strongly conserved primary structures (35% identity), while the β-type subunits are less similar in primary structure (25% identity). These genes for the individual subunits are not grouped on the yeast chromosome, but instead are randomly distributed over the yeast genome. The chromosomal deletion of each of the yeast proteasomal genes is lethal for the cell, except for the deletion of one of the α-type subunits (Hilt *et al.*, 1993b). The molecular masses of the 14 subunits range from 21.2 to 31.6 kDa (Heinemeyer *et al.*, 1994).

As with all studied forms of the proteasome, it is the β-type subunits that appear to hold the proteolytic activities for the yeast proteasome. Certain of the β-type subunits must be present and intact for activity. Furthermore, different subunits are needed to exhibit either the chymotrypsin- and trypsin-like activities (Pre2 gene) or the PGPH activity (Pre3 gene) (Groll *et al.*, 1997; Hilt *et al.*, 1993b). The architecture of the yeast 20S proteasome was

elucidated through use of the information obtained from structural studies of the *T. acidophilum* proteasome. It was found that the basic architecture of the yeast 20S proteasome was exactly the same as that of the archaeal and other eukaryotic proteasomes: a barrel-shaped particle made of two inner rings composed of seven β-type subunits each, and two outer disks, each composed of seven α-type subunits (Groll *et al.*, 1997). Furthermore, as with the mammalian proteasomes, each distinct subunit has a particular placement within the ordered structure (Kopp *et al*., 1993; Kopp *et al*., 1997) and the overall structure is identical to the three-dimensional structure of the human 20S proteasome shown in Figure 5.5. Although the overall architecture of the yeast 20S proteasome is similar to that of the archaeal proteasomes (barrel-shaped complex containing 3 inner cavities), it lacks the 13-Å wide entry ports at its ends. Instead, it has 10-Å wide passages in the side wall at the interface between the α and β rings of the complex (Groll *et al.*, 1997). Binding of the 19S regulatory complex may trigger rearrangement of the α rings and, therefore, open entry ports in the 20S complex (Gerards *et al.*, 1998).

VII. BACTERIAL PROTEASOMES

It was thought that the occurrence of the proteasome was restricted to archaeal and eukaryotic cells, until the first true bacterial proteasome was isolated in 1995 from *R. erythropolis* (Tamura *et al.*, 1995). Although there are many similarities between the bacterial proteasomes and proteasomes from other domains, there are also some distinct differences. It has been observed, so far, that the 20S quaternary structure is conserved

across all domains. However, the subunit composition of the bacterial proteasome is much less complex compared to eukaryotic proteasomes and, in some cases, may also differ from archaeal proteasomes. Instead of having 14 different (but related) subunits, or sets containing just one type of α and one type of β subunit, the bacterial 20S proteasome from *R. erythropolis* strain NI86/21 has two α-type and two β-type subunits (Tamura *et al.*, 1995). However, it is probable that the presence of four subunit types may be uncommon in bacteria, since only one α and one β subunit were found in other *R. erythropolis* strains (Lupas *et al*., 1997b), *Mycobacterium* species, and *Streptomyces coelicolor* (Nagy *et al.*, 1998).

Unlike the archaeal proteasomes, but like the eukaryotic proteasomes, the α-type subunits of the bacterial 20S proteasome do not assemble spontaneously into seven-membered rings, thereby creating a template for the formation of the catalytic β-type subunits (Zuhl *et al*., 1997b). Instead, the formation of a complete α ring is not a prerequisite for bacterial proteasome assembly. It follows, then, that the α-type subunits may not chaperone the 20S assembly in the same manner as the eukaryotic and archaeal α subunits. For instance, though the α subunits are critical for active-site formation in the β subunits, the *R. erythropolis* proteasome contains α subunits that cannot form the 7-membered rings without the presence of β and vice-versa (Lupas *et al.*, 1997b). From this knowledge, it was hypothesized that α/β heterodimers are initially created and quickly form into half-proteasomes, which then associate with each other to create the full structure (Mayr *et al*., 1998; Zuhl *et al.*, 1997a; Zuhl *et al.*, 1997b). Furthermore, the outer and inner rings are thought to have a random distribution of the α- and β-type subunits, respectively (Gerards *et*

al., 1998; Zuhl *et al.*, 1997a). Finally, at this point, it is thought that proteasomes only reside in actinomycetes (gram (+) bacteria with high G+C content) and are not found in other types of bacterial cells (Lupas *et al.*, 1997b; Nagy *et al.*, 1998).

VIII. REFERENCES

Achstetter, T., Ehmann, C., Osaki, A. & Wolf, D. H. (1984). Proteolysis in eukaryotic cells. Proteinase yscE, a new yeast peptidase. *J Biol Chem* **259**(21), 13344-8.

Ahn, K., Erlander, M., Leturcq, D., Peterson, P. A., Fruh, K. & Yang, Y. (1996). In vivo characterization of the proteasome regulator PA28. *J Biol Chem* **271**(30), 18237-42.

Akopian, T. N., Kisselev, A. F. & Goldberg, A. L. (1997). Processive degradation of proteins and other catalytic properties of the proteasome from Thermoplasma acidophilum. *J Biol Chem* **272**(3), 1791-8.

Arendt, C. S. & Hochstrasser, M. (1999). Eukaryotic 20S proteasome catalytic subunit propeptides prevent active site inactivation by N-terminal acetylation and promote particle assembly. *Embo J* **18**(13), 3575-85.

Armon, T., Ganoth, D. & Hershko, A. (1990). Assembly of the 26 S complex that degrades proteins ligated to ubiquitin is accompanied by the formation of ATPase activity. *J Biol Chem* **265**(34), 20723-6.

Arrigo, A. P., Tanaka, K., Goldberg, A. L. & Welch, W. J. (1988). Identity of the 19S 'prosome' particle with the large multifunctional protease complex of mammalian cells (the proteasome). *Nature* **331**(6152), 192-4.

Baker, D., Shiau, A. K. & Agard, D. A. (1993). The role of pro regions in protein folding. *Curr Opin Cell Biol* **5**(6), 966-70.

Bauer, M. W., Bauer, S. H. & Kelly, R. M. (1997). Purification and characterization of a proteasome from the hyperthermophilic archaeon Pyrococcus furiosus. *Appl Environ Microbiol* **63**(3), 1160-1164.

Baumeister, W., Walz, J., Zuhl, F. & Seemuller, E. (1998). The proteasome: paradigm of a self-compartmentalizing protease. *Cell* **92**(3), 367-80.

Benaroudj, N. & Goldberg, A. L. (2000). PAN, the proteasome-activating nucleotidase from archaebacteria, is a protein-unfolding molecular chaperone. *Nat Cell Biol* **2**(11), 833-9.

Benaroudj, N., Zwickl, P., Seemuller, E., Baumeister, W. & Goldberg, A. L. (2003). ATP hydrolysis by the proteasome regulatory complex PAN serves multiple functions in protein degradation. *Mol Cell* **11**(1), 69-78.

Beyer, A. (1997). Sequence analysis of the AAA protein family. *Protein Sci* **6**(10), 2043-58.

Brannigan, J. A., Dodson, G., Duggleby, H. J., Moody, P. C., Smith, J. L., Tomchick, D. R. & Murzin, A. G. (1995). A protein catalytic framework with an N-terminal nucleophile is capable of self-activation. *Nature* **378**(6555), 416-9.

Bult, C. J., White, O., Olsen, G. J., Zhou, L., Fleischmann, R. D., Sutton, G. G., Blake, J. A., FitzGerald, L. M., Clayton, R. A., Gocayne, J. D., Kerlavage, A. R., Dougherty, B. A., Tomb, J. F., Adams, M. D., Reich, C. I., Overbeek, R., Kirkness, E. F., Weinstock, K. G., Merrick, J. M., Glodek, A., Scott, J. L., Geoghagen, N. S. & Venter, J. C. (1996). Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii. *Science* **273**(5278), 1058-73.

Chen, P. & Hochstrasser, M. (1996). Autocatalytic subunit processing couples active site formation in the 20S proteasome to completion of assembly. *Cell* **86**(6), 961-72.

Chuang, S. E., Burland, V., Plunkett, G., 3rd, Daniels, D. L. & Blattner, F. R. (1993). Sequence analysis of four new heat-shock genes constituting the hslTS/ibpAB and hslVU operons in Escherichia coli. *Gene* **134**(1), 1-6.

Chung, C. H., Yoo, S. J., Seol, J. H. & Kang, M. S. (1997). Characterization of energy-dependent proteases in bacteria. *Biochem Biophys Res Commun* **241**(3), 613-6.

Chu-Ping, M., Vu, J. H., Proske, R. J., Slaughter, C. A. & DeMartino, G. N. (1994). Identification, purification, and characterization of a high molecular weight, ATP-dependent activator (PA700) of the 20 S proteasome. *J Biol Chem* **269**(5), 3539-47.

Ciechanover, A. (1994). The ubiquitin-proteasome proteolytic pathway. *Cell* **79**(1), 13-21.

Clayton, R. A., White, O., Ketchum, K. A. & Venter, J. C. (1997). The first genome from the third domain of life. *Nature* **387**(6632), 459-62.

Confalonieri, F. & Duguet, M. (1995). A 200-amino acid ATPase module in search of a basic function. *Bioessays* **17**(7), 639-50.

Dahlmann, B., Kopp, F., Kuehn, L., Niedel, B., Pfeifer, G., Hegerl, R. & Baumeister, W. (1989). The multicatalytic proteinase (prosome) is ubiquitous from eukaryotes to archaebacteria. *FEBS Lett* **251**(1-2), 125-31.

Dahlmann, B., Kuehn, L., Grziwa, A., Zwickl, P. & Baumeister, W. (1992). Biochemical properties of the proteasome from Thermoplasma acidophilum. *Eur J Biochem* **208**(3), 789-97.

Dahlmann, B., Kuehn, L., Rutschmann, M. & Reinauer, H. (1985). Purification and characterization of a multicatalytic high-molecular- mass proteinase from rat skeletal muscle. *Biochem J* **228**(1), 161-70.

De Mot, R., Nagy, I., Walz, J. & Baumeister, W. (1999). Proteasomes and other self-compartmentalizing proteases in prokaryotes. *Trends Microbiol* **7**(2), 88-92.

DeMartino, G. N. & Goldberg, A. L. (1979). Identification and partial purification of an ATP-stimulated alkaline protease in rat liver. *J Biol Chem* **254**(10), 3712-5.

DeMartino, G. N., Moomaw, C. R., Zagnitko, O. P., Proske, R. J., Chu-Ping, M., Afendis, S. J., Swaffield, J. C. & Slaughter, C. A. (1994). PA700, an ATP-dependent activator of the 20 S proteasome, is an ATPase containing multiple members of a nucleotide-binding protein family. *J Biol Chem* **269**(33), 20878-84.

DeMartino, G. N. & Slaughter, C. A. (1999). The proteasome, a novel protease regulated by multiple mechanisms. *J Biol Chem* **274**(32), 22123-6.

Deveraux, Q., Ustrell, V., Pickart, C. & Rechsteiner, M. (1994). A 26 S protease subunit that binds ubiquitin conjugates. *J Biol Chem* **269**(10), 7059-61.

Dick, T. P., Ruppert, T., Groettrup, M., Kloetzel, P. M., Kuehn, L., Koszinowski, U. H., Stevanovic, S., Schild, H. & Rammensee, H. G. (1996). Coordinated dual cleavages induced by the proteasome regulator PA28 lead to dominant MHC ligands. *Cell* **86**(2), 253-62.

Ditzel, L., Huber, R., Mann, K., Heinemeyer, W., Wolf, D. H. & Groll, M. (1998). Conformational constraints for protein self-cleavage in the proteasome. *J Mol Biol* **279**(5), 1187-91.

Dodson, G. & Wlodawer, A. (1998). Catalytic triads and their relatives. *Trends Biochem Sci* **23**(9), 347-52.

Dubiel, W., Pratt, G., Ferrell, K. & Rechsteiner, M. (1992). Purification of an 11 S regulator of the multicatalytic protease. *J Biol Chem* **267**(31), 22369-77.

Durner, J. & Boger, P. (1995). Ubiquitin in the prokaryote Anabaena variabilis. *J Biol Chem* **270**(8), 3720-5.

Eytan, E., Armon, T., Heller, H., Beck, S. & Hershko, A. (1993). Ubiquitin C-terminal

hydrolase activity associated with the 26 S protease complex. *J Biol Chem* **268**(7), 4668-74.

Fenteany, G., Standaert, R. F., Lane, W. S., Choi, S., Corey, E. J. & Schreiber, S. L. (1995). Inhibition of proteasome activities and subunit-specific amino-terminal threonine modification by lactacystin. *Science* **268**(5211), 726-31.

Frentzel, S., Pesold-Hurt, B., Seelig, A. & Kloetzel, P. M. (1994). 20 S proteasomes are assembled via distinct precursor complexes. Processing of LMP2 and LMP7 proproteins takes place in 13-16 S preproteasome complexes. *J Mol Biol* **236**(4), 975-81.

Fu, H., Doelling, J. H., Arendt, C. S., Hochstrasser, M. & Vierstra, R. D. (1998). Molecular organization of the 20S proteasome gene family from Arabidopsis thaliana. *Genetics* **149**(2), 677-92.

Fujinami, K., Tanahashi, N., Tanaka, K., Ichihara, A., Cejka, Z., Baumeister, W., Miyawaki, M., Sato, T. & Nakagawa, H. (1994). Purification and characterization of the 26 S proteasome from spinach leaves. *J Biol Chem* **269**(41), 25905-10.

Gerards, W. L., de Jong, W. W., Boelens, W. & Bloemendal, H. (1998). Structure and assembly of the 20S proteasome. *Cell Mol Life Sci* **54**(3), 253-62.

Gerards, W. L., Enzlin, J., Haner, M., Hendriks, I. L., Aebi, U., Bloemendal, H. & Boelens, W. (1997). The human alpha-type proteasomal subunit HsC8 forms a double ringlike structure, but does not assemble into proteasome-like particles with the beta-type subunits HsDelta or HsBPROS26. *J Biol Chem* **272**(15), 10080-6.

Glickman, M. H., Rubin, D. M., Coux, O., Wefes, I., Pfeifer, G., Cjeka, Z., Baumeister, W.,

Fried, V. A. & Finley, D. (1998). A subcomplex of the proteasome regulatory particle required for ubiquitin-conjugate degradation and related to the COP9-signalosome and eIF3. *Cell* **94**(5), 615-23.

Gorbea, C., Taillandier, D. & Rechsteiner, M. (1999). Assembly of the regulatory complex of the 26S proteasome. *Mol Biol Rep* **26**(1-2), 15-9.

Gottesman, S. (1996). Proteases and their targets in Escherichia coli. *Annu Rev Genet* **30**, 465-506.

Gottesman, S., Wickner, S. & Maurizi, M. R. (1997). Protein quality control: triage by chaperones and proteases. *Genes Dev* **11**(7), 815-23.

Groll, M., Bajorek, M., Kohler, A., Moroder, L., Rubin, D. M., Huber, R., Glickman, M. H. & Finley, D. (2000). A gated channel into the proteasome core particle. *Nat Struct Biol* **7**(11), 1062-7.

Groll, M., Ditzel, L., Lowe, J., Stock, D., Bochtler, M., Bartunik, H. D. & Huber, R. (1997). Structure of 20S proteasome from yeast at 2.4 A resolution. *Nature* **386**(6624), 463-71.

Grziwa, A., Baumeister, W., Dahlmann, B. & Kopp, F. (1991). Localization of subunits in proteasomes from Thermoplasma acidophilum by immunoelectron microscopy. *FEBS Lett* **290**(1-2), 186-90.

Grziwa, A., Maack, S., Puhler, G., Wiegand, G., Baumeister, W. & Jaenicke, R. (1994). Dissociation and reconstitution of the Thermoplasma proteasome. *Eur J Biochem* **223**(3), 1061-7.

Harris, J. R. (1968). Release of a macromolecular protein component from human erythrocyte ghosts. *Biochim Biophys Acta* **150**(3), 534-7.

Hegerl, R., Pfeifer, G., Puhler, G., Dahlmann, B. & Baumeister, W. (1991). The three-dimensional structure of proteasomes from Thermoplasma acidophilum as determined by electron microscopy using random conical tilting. *FEBS Lett* **283**(1), 117-21.

Heinemeyer, W., Trondle, N., Albrecht, G. & Wolf, D. H. (1994). PRE5 and PRE6, the last missing genes encoding 20S proteasome subunits from yeast? Indication for a set of 14 different subunits in the eukaryotic proteasome core. *Biochemistry* **33**(40), 12229-37.

Hershko, A. (1996). Lessons from the discovery of the ubiquitin system. *Trends Biochem Sci* **21**(11), 445-9.

Hilt, W., Enenkel, C., Gruhler, A., Singer, T. & Wolf, D. H. (1993a). The PRE4 gene codes for a subunit of the yeast proteasome necessary for peptidylglutamyl-peptide-hydrolyzing activity. Mutations link the proteasome to stress- and ubiquitin-dependent proteolysis. *J Biol Chem* **268**(5), 3479-86.

Hilt, W., Heinemeyer, W. & Wolf, D. H. (1993b). Studies on the yeast proteasome uncover its basic structural features and multiple in vivo functions. *Enzyme Protein* **47**(4-6), 189-201.

Hoffman, L., Pratt, G. & Rechsteiner, M. (1992). Multiple forms of the 20 S multicatalytic and the 26 S ubiquitin/ATP- dependent proteases from rabbit reticulocyte lysate. *J Biol Chem* **267**(31), 22362-8.

Hoffman, L. & Rechsteiner, M. (1994). Activation of the multicatalytic protease. The 11 S regulator and 20 S ATPase complexes contain distinct 30-kilodalton subunits. *J Biol Chem* **269**(24), 16890-5.

Hoffman, L. & Rechsteiner, M. (1996). Nucleotidase activities of the 26 S proteasome and its

regulatory complex. *J Biol Chem* **271**(51), 32538-45.

Kanayama, H. O., Tamura, T., Ugai, S., Kagawa, S., Tanahashi, N., Yoshimura, T., Tanaka, K. & Ichihara, A. (1992). Demonstration that a human 26S proteolytic complex consists of a proteasome and multiple associated protein components and hydrolyzes ATP and ubiquitin-ligated proteins by closely linked mechanisms. *Eur J Biochem* **206**(2), 567-78.

Kawarabayasi, Y., Hino, Y., Horikawa, H., Yamazaki, S., Haikawa, Y., Jin-no, K., Takahashi, M., Sekine, M., Baba, S., Ankai, A., Kosugi, H., Hosoyama, A., Fukui, S., Nagai, Y., Nishijima, K., Nakazawa, H., Takamiya, M., Masuda, S., Funahashi, T., Tanaka, T., Kudoh, Y., Yamazaki, J., Kushida, N., Oguchi, A., Kikuchi, H. & et al. (1999). Complete genome sequence of an aerobic hyper-thermophilic crenarchaeon, Aeropyrum pernix K1. *DNA Res* **6**(2), 83-101, 145-52.

Kawarabayasi, Y., Sawada, M., Horikawa, H., Haikawa, Y., Hino, Y., Yamamoto, S., Sekine, M., Baba, S., Kosugi, H., Hosoyama, A., Nagai, Y., Sakai, M., Ogura, K., Otsuka, R., Nakazawa, H., Takamiya, M., Ohfuku, Y., Funahashi, T., Tanaka, T., Kudoh, Y., Yamazaki, J., Kushida, N., Oguchi, A., Aoki, K. & Kikuchi, H. (1998). Complete sequence and gene organization of the genome of a hyper- thermophilic archaebacterium, Pyrococcus horikoshii OT3. *DNA Res* **5**(2), 55-76.

Kisselev, A. F., Akopian, T. N. & Goldberg, A. L. (1998). Range of sizes of peptide products generated during degradation of different proteins by archaeal proteasomes. *J Biol Chem* **273**(4), 1982-9.

Kisselev, A. F., Songyang, Z. & Goldberg, A. L. (2000). Why does threonine, and not serine, function as the active site nucleophile in proteasomes? *J Biol Chem* **275**(20), 14831-7.

Kleinschmidt, J. A., Escher, C. & Wolf, D. H. (1988). Proteinase yscE of yeast shows homology with the 20 S cylinder particles of Xenopus laevis. *FEBS Lett* **239**(1), 35-40.

Knowlton, J. R., Johnston, S. C., Whitby, F. G., Realini, C., Zhang, Z., Rechsteiner, M. & Hill, C. P. (1997). Structure of the proteasome activator REGalpha (PA28alpha). *Nature* **390**(6660), 639-43.

Kopp, F., Dahlmann, B. & Hendil, K. B. (1993). Evidence indicating that the human proteasome is a complex dimer. *J Mol Biol* **229**(1), 14-9.

Kopp, F., Hendil, K. B., Dahlmann, B., Kristensen, P., Sobek, A. & Uerkvitz, W. (1997). Subunit arrangement in the human 20S proteasome. *Proc Natl Acad Sci U S A* **94**(7), 2939-44.

Koster, A. J., Walz, J., Lupas, A. & Baumeister, W. (1995). Structural features of archaebacterial and eukaryotic proteasomes. *Mol Biol Rep* **21**(1), 11-20.

Kuehn, L. & Dahlmann, B. (1996). Proteasome activator PA28 and its interaction with 20 S proteasomes. *Arch Biochem Biophys* **329**(1), 87-96.

Lam, Y. A., Xu, W., DeMartino, G. N. & Cohen, R. E. (1997). Editing of ubiquitin conjugates by an isopeptidase in the 26S proteasome. *Nature* **385**(6618), 737-40.

Larsen, C. N. & Finley, D. (1997). Protein translocation channels in the proteasome and other proteases. *Cell* **91**(4), 431-4.

Lowe, J., Stock, D., Jap, B., Zwickl, P., Baumeister, W. & Huber, R. (1995). Crystal structure of the 20S proteasome from the archaeon T. acidophilum at 3.4 A resolution. *Science* **268**(5210), 533-9.

Lupas, A., Flanagan, J. M., Tamura, T. & Baumeister, W. (1997a). Self-compartmentalizing

proteases. *Trends Biochem Sci* **22**(10), 399-404.

Lupas, A., Zuhl, F., Tamura, T., Wolf, S., Nagy, I., De Mot, R. & Baumeister, W. (1997b). Eubacterial proteasomes. *Mol Biol Rep* **24**(1-2), 125-31.

Ma, C. P., Slaughter, C. A. & DeMartino, G. N. (1992). Identification, purification, and characterization of a protein activator (PA28) of the 20 S proteasome (macropain). *J Biol Chem* **267**(15), 10515-23.

Maupin-Furlow, J. A., Aldrich, H. C. & Ferry, J. G. (1998). Biochemical characterization of the 20S proteasome from the methanoarchaeon Methanosarcina thermophila. *J Bacteriol* **180**(6), 1480-7.

Maupin-Furlow, J. A. & Ferry, J. G. (1995). A proteasome from the methanogenic archaeon Methanosarcina thermophila. *J Biol Chem* **270**(48), 28617-22.

Maupin-Furlow, J. A., Kaczowka, S. J., Ou, M. S. & Wilson, H. L. (2001). Archaeal proteasomes: proteolytic nanocompartments of the cell. *Adv Appl Microbiol* **50**, 279-338.

Mayr, J., Seemuller, E., Muller, S. A., Engel, A. & Baumeister, W. (1998). Late events in the assembly of 20S proteasomes. *J Struct Biol* **124**(2-3), 179-88.

Mykles, D. L. (1996). Differential effects of bovine PA28 on six peptidase activities of the lobster muscle proteasome (multicatalytic proteinase). *Arch Biochem Biophys* **325**(1), 77-81.

Nagy, I., Tamura, T., Vanderleyden, J., Baumeister, W. & De Mot, R. (1998). The 20S proteasome of Streptomyces coelicolor. *J Bacteriol* **180**(20), 5448-53.

Navon, A. & Goldberg, A. L. (2001). Proteins are unfolded on the surface of the ATPase ring before transport into the proteasome. *Mol Cell* **8**(6), 1339-49.

Neuwald, A. F., Aravind, L., Spouge, J. L. & Koonin, E. V. (1999). AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. *Genome Res* **9**(1), 27-43.

Ng, W. V., Kennedy, S. P., Mahairas, G. G., Berquist, B., Pan, M., Shukla, H. D., Lasky, S. R., Baliga, N. S., Thorsson, V., Sbrogna, J., Swartzell, S., Weir, D., Hall, J., Dahl, T. A., Welti, R., Goo, Y. A., Leithauser, B., Keller, K., Cruz, R., Danson, M. J., Hough, D. W., Maddocks, D. G., Jablonski, P. E., Krebs, M. P., Angevine, C. M., Dale, H., Isenbarger, T. A., Peck, R. F., Pohlschroder, M., Spudich, J. L., Jung, K. W., Alam, M., Freitas, T., Hou, S., Daniels, C. J., Dennis, P. P., Omer, A. D., Ebhardt, H., Lowe, T. M., Liang, P., Riley, M., Hood, L. & DasSarma, S. (2000). Genome sequence of Halobacterium species NRC-1. *Proc Natl Acad Sci U S A* **97**(22), 12176-81.

Orlowski, M., Cardozo, C. & Michaud, C. (1993). Evidence for the presence of five distinct proteolytic components in the pituitary multicatalytic proteinase complex. Properties of two components cleaving bonds on the carboxyl side of branched chain and small neutral amino acids. *Biochemistry* **32**(6), 1563-72.

Palmer, A., Rivett, A. J., Thomson, S., Hendil, K. B., Butcher, G. W., Fuertes, G. & Knecht, E. (1996). Subpopulations of proteasomes in rat liver nuclei, microsomes and cytosol. *Biochem J* **316**(Pt 2), 401-7.

Pamnani, V., Tamura, T., Lupas, A., Peters, J., Cejka, Z., Ashraf, W. & Baumeister, W. (1997). Cloning, sequencing and expression of VAT, a CDC48/p97 ATPase homologue from the archaeon Thermoplasma acidophilum. *FEBS Lett* **404**(2-3), 263-8.

Peters, J. M., Cejka, Z., Harris, J. R., Kleinschmidt, J. A. & Baumeister, W. (1993).

Structural features of the 26 S proteasome complex. *J Mol Biol* **234**(4), 932-7.

Peters, J. M., Franke, W. W. & Kleinschmidt, J. A. (1994). Distinct 19 S and 20 S subcomplexes of the 26 S proteasome and their distribution in the nucleus and the cytoplasm. *J Biol Chem* **269**(10), 7709-18.

Puhler, G., Weinkauf, S., Bachmann, L., Muller, S., Engel, A., Hegerl, R. & Baumeister, W. (1992). Subunit stoichiometry and three-dimensional arrangement in proteasomes from Thermoplasma acidophilum. *Embo J* **11**(4), 1607-16.

Ramos, P. C., Hockendorff, J., Johnson, E. S., Varshavsky, A. & Dohmen, R. J. (1998). Ump1p is required for proper maturation of the 20S proteasome and becomes its substrate upon completion of the assembly. *Cell* **92**(4), 489-99.

Reits, E. A., Benham, A. M., Plougastel, B., Neefjes, J. & Trowsdale, J. (1997). Dynamics of proteasome distribution in living cells. *Embo J* **16**(20), 6087-94.

Ruepp, A., Eckerskorn, C., Bogyo, M. & Baumeister, W. (1998). Proteasome function is dispensable under normal but not under heat shock conditions in Thermoplasma acidophilum. *FEBS Lett* **425**(1), 87-90.

Ruepp, A., Graml, W., Santos-Martinez, M. L., Koretke, K. K., Volker, C., Mewes, H. W., Frishman, D., Stocker, S., Lupas, A. N. & Baumeister, W. (2000). The genome sequence of the thermoacidophilic scavenger Thermoplasma acidophilum. *Nature* **407**(6803), 508-13.

Schmid, H. P., Akhayat, O., Martins De Sa, C., Puvion, F., Koehler, K. & Scherrer, K. (1984). The prosome: an ubiquitous morphologically distinct RNP particle associated with repressed mRNPs and containing specific ScRNA and a characteristic set of proteins. *Embo J* **3**(1), 29-34.

Schmidt, M., Lupas, A. N. & Finley, D. (1999a). Structure and mechanism of ATP-dependent proteases. *Curr Opin Chem Biol* **3**(5), 584-91.

Schmidt, M., Zantopf, D., Kraft, R., Kostka, S., Preissner, R. & Kloetzel, P. M. (1999b). Sequence information within proteasomal prosequences mediates efficient integration of beta-subunits into the 20 S proteasome complex. *J Mol Biol* **288**(1), 117-28.

Schmidtke, G., Kraft, R., Kostka, S., Henklein, P., Frommel, C., Lowe, J., Huber, R., Kloetzel, P. M. & Schmidt, M. (1996). Analysis of mammalian 20S proteasome biogenesis: the maturation of beta- subunits is an ordered two-step mechanism involving autocatalysis. *Embo J* **15**(24), 6887-98.

Seemuller, E., Lupas, A. & Baumeister, W. (1996). Autocatalytic processing of the 20S proteasome. *Nature* **382**(6590), 468-71.

Seemuller, E., Lupas, A., Stock, D., Lowe, J., Huber, R. & Baumeister, W. (1995). Proteasome from Thermoplasma acidophilum: a threonine protease. *Science* **268**(5210), 579-82.

Shinde, U. & Inouye, M. (1994). The structural and functional organization of intramolecular chaperones: the N-terminal propeptides which mediate protein folding. *J Biochem (Tokyo)* **115**(4), 629-36.

Song, X., Mott, J. D., von Kampen, J., Pramanik, B., Tanaka, K., Slaughter, C. A. & DeMartino, G. N. (1996). A model for the quaternary structure of the proteasome activator PA28. *J Biol Chem* **271**(42), 26410-7.

Stock, D., Ditzel, L., Baumeister, W., Huber, R. & Lowe, J. (1995). Catalytic mechanism of the 20S proteasome of Thermoplasma acidophilum revealed by X-ray crystallography. *Cold Spring Harb Symp Quant Biol* **60**, 525-32.

Suzuki, T., Iwasaki, T., Uzawa, T., Hara, K., Nemoto, N., Kon, T., Ueki, T., Yamagishi, A. & Oshima, T. (2002). Sulfolobus tokodaii sp. nov. (f. Sulfolobus sp. strain 7), a new member of the genus Sulfolobus isolated from Beppu Hot Springs, Japan. *Extremophiles* **6**(1), 39-44.

Tamura, T., Nagy, I., Lupas, A., Lottspeich, F., Cejka, Z., Schoofs, G., Tanaka, K., De Mot, R. & Baumeister, W. (1995). The first characterization of a eubacterial proteasome: the 20S complex of Rhodococcus. *Curr Biol* **5**(7), 766-74.

Tamura, T., Tamura, N., Cejka, Z., Hegerl, R., Lottspeich, F. & Baumeister, W. (1996). Tricorn protease--the core of a modular proteolytic system. *Science* **274**(5291), 1385-9.

Tanahashi, N., Kawahara, H., Murakami, Y. & Tanaka, K. (1999). The proteasome-dependent proteolytic system. *Mol Biol Rep* **26**(1-2), 3-9.

Tanaka, K. (1998). Proteasomes: structure and biology. *J Biochem (Tokyo)* **123**(2), 195-204.

Tanaka, K., Yoshimura, T., Kumatori, A., Ichihara, A., Ikai, A., Nishigai, M., Kameyama, K. & Takagi, T. (1988). Proteasomes (multi-protease complexes) as 20 S ring-shaped particles in a variety of eukaryotic cells. *J Biol Chem* **263**(31), 16209-17.

Udvardy, A. (1993). Purification and characterization of a multiprotein component of the Drosophila 26 S (1500 kDa) proteolytic complex. *J Biol Chem* **268**(12), 9055-62.

Ugai, S., Tamura, T., Tanahashi, N., Takai, S., Komi, N., Chung, C. H., Tanaka, K. & Ichihara, A. (1993). Purification and characterization of the 26S proteasome complex catalyzing ATP-dependent breakdown of ubiquitin-ligated proteins from rat liver. *J Biochem (Tokyo)* **113**(6), 754-68.

Voges, D., Zwickl, P. & Baumeister, W. (1999). The 26S proteasome: a molecular machine

designed for controlled proteolysis. *Annu Rev Biochem* **68**, 1015-68.

Walz, J., Erdmann, A., Kania, M., Typke, D., Koster, A. J. & Baumeister, W. (1998). 26S proteasome structure revealed by three-dimensional electron microscopy. *J Struct Biol* **121**(1), 19-29.

Wenzel, T. & Baumeister, W. (1993). Thermoplasma acidophilum proteasomes degrade partially unfolded and ubiquitin-associated proteins. *FEBS Lett* **326**(1-3), 215-8.

Wenzel, T. & Baumeister, W. (1995). Conformational constraints in protein degradation by the 20S proteasome. *Nat Struct Biol* **2**(3), 199-204.

Whitby, F. G., Masters, E. I., Kramer, L., Knowlton, J. R., Yao, Y., Wang, C. C. & Hill, C. P. (2000). Structural basis for the activation of 20S proteasomes by 11S regulators. *Nature* **408**(6808), 115-20.

Wilk, S. & Orlowski, M. (1980). Cation-sensitive neutral endopeptidase: isolation and specificity of the bovine pituitary enzyme. *J Neurochem* **35**(5), 1172-82.

Wilk, S. & Orlowski, M. (1983). Evidence that pituitary cation-sensitive neutral endopeptidase is a multicatalytic protease complex. *J Neurochem* **40**(3), 842-9.

Wilson, H. L., Aldrich, H. C. & Maupin-Furlow, J. (1999). Halophilic 20S proteasomes of the archaeon Haloferax volcanii: purification, characterization, and gene sequence analysis. *J Bacteriol* **181**(18), 5814-24.

Wilson, H. L., Ou, M. S., Aldrich, H. C. & Maupin-Furlow, J. (2000). Biochemical and physical properties of the Methanococcus jannaschii 20S proteasome and PAN, a homolog of the ATPase (Rpt) subunits of the eucaryal 26S proteasome. *J Bacteriol* **182**(6), 1680-92.

Witt, E., Zantopf, D., Schmidt, M., Kraft, R., Kloetzel, P. M. & Kruger, E. (2000).

Characterisation of the newly identified human Ump1 homologue POMP and analysis of LMP7(beta 5i) incorporation into 20 S proteasomes. *J Mol Biol* **301**(1), 1-9.

Wolf, S., Lottspeich, F. & Baumeister, W. (1993). Ubiquitin found in the archaebacterium Thermoplasma acidophilum. *FEBS Lett* **326**(1-3), 42-4.

Yoshimura, T., Kameyama, K., Takagi, T., Ikai, A., Tokunaga, F., Koide, T., Tanahashi, N., Tamura, T., Cejka, Z., Baumeister, W. & et al. (1993). Molecular characterization of the "26S" proteasome complex from rat liver. *J Struct Biol* **111**(3), 200-11.

Zuhl, F., Seemuller, E., Golbik, R. & Baumeister, W. (1997a). Dissecting the assembly pathway of the 20S proteasome. *FEBS Lett* **418**(1-2), 189-94.

Zuhl, F., Tamura, T., Dolenc, I., Cejka, Z., Nagy, I., De Mot, R. & Baumeister, W. (1997b). Subunit topology of the Rhodococcus proteasome. *FEBS Lett* **400**(1), 83-90.

Zwickl, P., Baumeister, W. & Steven, A. (2000). Dis-assembly lines: the proteasome and related ATPase-assisted proteases. *Curr Opin Struct Biol* **10**(2), 242-50.

Zwickl, P., Grziwa, A., Puhler, G., Dahlmann, B., Lottspeich, F. & Baumeister, W. (1992a). Primary structure of the Thermoplasma proteasome and its implications for the structure, function, and evolution of the multicatalytic proteinase. *Biochemistry* **31**(4), 964-72.

Zwickl, P., Kleinz, J. & Baumeister, W. (1994). Critical elements in proteasome assembly. *Nat Struct Biol* **1**(11), 765-70.

Zwickl, P., Lottspeich, F. & Baumeister, W. (1992b). Expression of functional Thermoplasma acidophilum proteasomes in Escherichia coli. *FEBS Lett* **312**(2-3), 157-60.

Zwickl, P., Ng, D., Woo, K. M., Klenk, H. P. & Goldberg, A. L. (1999a). An archaebacterial

ATPase, homologous to ATPases in the eukaryotic 26 S proteasome, activates protein breakdown by 20 S proteasomes. *J Biol Chem* **274**(37), 26008-14.

Zwickl, P., Voges, D. & Baumeister, W. (1999b). The proteasome: a macromolecular assembly designed for controlled proteolysis. *Philos Trans R Soc Lond B Biol Sci* **354**(1389), 1501-11.

CHAPTER 6:

Role of β-1 subunit in 20S proteasome from the hyperthermophilic archaeon *Pyrococcus furiosus* relates to thermal stress response

Lara S. Chang, Keith R. Shockley, Shannon B. Conners,
Kevin L. Epting, Matthew R. Johnson and Robert M. Kelly*

Department of Chemical Engineering, North Carolina State University,
Raleigh, NC 27695-7905

Submitted to: **Journal of Biological Chemistry** (June, 2003)

Running Title: *Pyrococcus furiosus* proteasome

*Address inquiries to: **Robert M. Kelly**
Department of Chemical Engineering
North Carolina State University
Raleigh, NC 27695-7905

Phone: (919) 515-6396
Fax: (919) 515-3465
Email: rmkelly@eos.ncsu.edu

I. ABSTRACT

The roles of the two (48% identical) β subunits were examined for the 20S proteasome from the hyperthermophilic archaeon *Pyrococcus furiosus*. To date, most of the archaeal proteasomes studied contain only one α-type and one β-type subunit (Maupin-Furlow *et al*., 1998; Seemuller *et al*., 1996; Wilson *et al*., 2000; Zwickl *et al*., 1994). *Haloferax volcanii* contains two α-type subunits, whose individual roles have not yet been elucidated (Maupin-Furlow *et al*., 2001). In the case of *P. furiosus*, it was found from targeted cDNA microarray experiments that the proteasome α-type subunits were down-regulated upon heat shock (2- to 3-fold after 60 minutes of heat shock), while the two β subunits (β1 and β2) were both up-regulated (1.5- to 2-fold after 60 min). The proteasome-activating nucleotidase (PAN) from *P. furiosus* was up-regulated as high as 15-fold during the first 10 min of heat shock, with this up-regulation disappearing after 60 min. Biochemical and biophysical studies were done on a native proteasome from *P. furiosus* cells grown under normal conditions and a native heat-shocked proteasome (NHS) from cells that were heat-shocked. Additionally, the individual proteasome subunits were expressed and assembled in all possible combinations under several conditions. These various proteasome forms were then characterized. It was found that the α subunit was able to form large stable structures spontaneously, while the β subunits were not able to process their pro-sequences or assemble into organized structures. It is likely that the α subunit plays a role in initiating the processing and assembly of the β subunits. As well, it was found that the β2 was as catalytically active as the native proteasome, which contained both β1 and β2 subunits.

From this, it was concluded that the presence or absence of β1 had no noticeable effect on the catalytic ability of the proteasome, but instead may have played a stabilizing role as seen in activity decay studies. The PAN ATPase activity (from expressed *P. furiosus* PAN) was able to stimulate polypeptide degradation by the native and several recombinant proteasomes, but inhibited the activities of the proteasome from heat-shocked *P. furiosus* and the recombinant proteasomes that were assembled under high-temperataure stress conditions. PAN was most likely a regulator of the proteasome from *P. furiosus* under normal growth conditions, but did not play a role in regulating the activity of the proteasome assembled during stress conditions. This stress form of the proteasome appeared to include a higher level of the stabilizing, non-catalytic β1 subunit.

II. INTRODUCTION

The protein degradation pathway, beginning with a substrate protein that is targeted for breakdown and ending with single amino acids that can be used to build new proteins, is an important part of the cellular survival mechanism (Rao *et al*., 1998). Often proteins are targeted for degradation when they become misfolded or aggregated in response to a stressful environmental condition, or when they have completed their specific regulatory cellular role (Baumeister *et al*., 1998; Voges *et al*., 1999). Although the proteases that participate in this pathway are a vital part of both routine and stress response actions in the cell, they could also be detrimental if their proteolytic activity were to become unregulated. As a result, an important part of proteolytic regulation relates to compartmentalizing structural aspects (Baumeister *et al.*, 1998; Lupas *et al*., 1997a): certain proteases are composed of self-assembling small subunits forming barrel-like structures, which limit substrate access to the active sites. Often, these active sites are only accessible through a carefully controlled sequence of events involving one or more energy-dependent regulatory enzymes, such as ATPases (Gottesman *et al*., 1997; Larsen & Finley, 1997).

The proteasome is one of several self-compartmentalizing proteases studied to date, with the rapid advances made in elucidating its assembly, mechanism of action, and regulation becoming a defining factor in the concept of self-compartmentalization (Baumeister *et al.*, 1998). It has been found in all domains of life, ranging from bacteria to mammals (Dahlmann *et al*., 1989; Tanaka *et al*., 1988), and is characterized by a conserved core "20S" structure, consisting of 28 subunits arranged in a stack of 4 rings containing 7 subunits each. The outer two rings contain α-type, non-catalytic subunits, while the inner

two rings contain β-type, catalytic subunits, giving the characteristic $\alpha_7\beta_7\beta_7\alpha_7$ structure (Grziwa *et al*., 1991; Lowe *et al*., 1995). The eukaryotic form of the proteasome, in conjunction with the accompanying ubiquitin-mediated substrate targeting system (Ciechanover, 1994; Hershko, 1996), has been extensively looked at because of its human therapeutic relevance; proteasome inhibitors have become effective tools in the treatment of cancer and inflammatory disorders (Goldberg & Rock, 2002; Kisselev & Goldberg, 2001). The archaeal proteasome is considerably simpler than the eukaryotic form, as it contains only 2 to 3 forms of the α and β subunits as compared to as many as 14 different (eukaryotic) subunit versions. As such, the archaeal proteasome, whose structural organization and proteolytic mechanism are very similar to those of eukaryotic 20S (Baumeister *et al.*, 1998; De Mot *et al*., 1999; Zwickl *et al*., 2000), presents a simpler system for structural (Lowe *et al.*, 1995) and mechanistic studies (Akopian *et al*., 1997; Kisselev *et al*., 1998), the results of which can be extrapolated to the more complex eukaryotic forms.

Sequenced genomes of thermophilic archaea have confirmed the presence of elements of the 20S and 26S proteasome in these microorganisms. A summary of those genes and their similarity to the *P. furiosus* 20S proteasome is listed in Table 6.1. Native and recombinant proteasomes from several thermophilic archaea have been examined with respect to structural and biochemical properties, including *Thermoplasma acidophilum* (Akopian *et al.*, 1997; Dahlmann *et al*., 1992; Grziwa *et al.*, 1991; Grziwa *et al*., 1994; Hegerl *et al*., 1991; Kisselev *et al.*, 1998; Kisselev *et al*., 2000; Lowe *et al.*, 1995; Puhler *et al*., 1992; Ruepp *et al*., 1998; Seemuller *et al*., 1995; Stock *et al*., 1995; Wenzel & Baumeister, 1993; Zwickl *et al*., 1992a; Zwickl *et al.*, 1994; Zwickl *et al*., 1992b), *Methanosarcina thermophila* (Maupin-Furlow *et al.*, 1998; Maupin-Furlow & Ferry, 1995),

and *Methanococcus jannaschii* (Wilson *et al.*, 2000). However, while archaeal and bacterial 20S proteasomes typically involve 2 to 3 different types of subunits arranged in the characteristic $\alpha_7\beta_7\beta_7\alpha_7$ conformation (one important exception is *Rhodococcus erythropolis*, which has 2 α- and 2 β-type subunits, with the paralogs over 80% identical; (Tamura *et al*., 1995; Zuhl *et al*., 1997b)), the lower eukaryotes such as the yeast *Saccharomyces cerevisiae* generally contain 7 α-type and 7 β-type subunits, as deduced from genome sequence (Clayton *et al*., 1997) and biochemical information (Heinemeyer *et al*., 1994; Hilt *et al*., 1993). Genomes of higher eukaryotes, such as *Arabidopsis thaliana* (Fu *et al*., 1998), have been found to encode for as many as 13 α-type and 10 β-type subunits. However, many of these are more than 90% identical at the amino acid sequence level, possibly representing some degree of redundancy (Fu *et al*., 1999). Contrary to the eukaryotic species, the archaeal species with more than one α- or β-type subunit contain paralogs that are generally less than 60% identical, suggesting a distinctive role for each subunit.

Bacterial (Zuhl *et al*., 1997a), archaeal (Maupin-Furlow *et al.*, 1998; Maupin-Furlow & Ferry, 1995; Seemuller *et al.*, 1996; Wilson *et al.*, 2000; Zwickl *et al.*, 1994), yeast (Arendt & Hochstrasser, 1999; Chen & Hochstrasser, 1996; Ditzel *et al*., 1998), and mammalian (Schmidtke *et al*., 1996) proteasomes all contain at least one β subunit that initially exists as a pro-subunit. This subunit is later processed, through autocatalytic removal of the amino-terminal residues, during assembly with the α subunit. The pro-peptide, immediately upstream of the active-site threonine (Thr) (Kisselev *et al.*, 2000; Lowe *et al.*, 1995; Seemuller *et al.*, 1995) can be as small as 8-10 residues, as in *T. acidophilum* (Seemuller *et al.*, 1996), or >50 residues, as in human LMP7 (Schmidt *et al*., 1999; Witt *et al*., 2000). During hydrolysis, the side chain amino groups Lys-33 and Asp/Glu-17 may

accept the side chain Thr proton in a charge-relay system through a salt bridge, as found in *T. acidophilum* (Seemuller *et al.*, 1996). The active-site Thr, which is exposed to become the first N-terminal residue (Thr^1) upon removal of the pro-sequence, then acts as the primary catalytic site during proteolysis. Specifically, the γ-oxygen acts as the nucleophile and the α-amino group most likely acts as the proton acceptor (Fenteany *et al*., 1995; Groll *et al*., 1997; Lowe *et al.*, 1995; Maupin-Furlow *et al.*, 1998; Seemuller *et al.*, 1995), with the precise mechanism hypothesized to be similar to that of serine proteases (Maupin-Furlow *et al.*, 2001).

In many cases, the pro-peptide was found to be important, though not essential, for proper proteasome assembly. However, the assembly and role of β subunits has been found to differ among the proteasomes from the three domains of life (Gerards *et al*., 1998b). In the case of the bacterial *Rhodococcus erythropolis* proteasome, absence of the pro-peptide significantly lowered the assembly efficiency of the subunits (Zuhl *et al.*, 1997a). In the archaeal *T. acidophilum* proteasome, the pro-peptide was not critical for proper assembly (Grziwa *et al.*, 1994). In yeast proteasomes, the pro-peptides of certain β subunits were critical for folding and incorporation of these subunits into the mature proteasome, though they had no role in correct positioning within the structure (Arendt & Hochstrasser, 1999; Baker *et al*., 1993; Chen & Hochstrasser, 1996; Ramos *et al*., 1998; Shinde & Inouye, 1994). Assuming that the archaeal proteasome was an evolutionary precursor to the eukaryotic version, it is interesting to consider how the expansion of the α- and β-subunit inventories relates to functional attributes of protein turnover.

As is the case with the eukaryotic proteasome, several archaeal versions can expand their proteolytic capacity through the use of ATP, mediated through at least one known

ATPase, referred to as PAN (proteasome-activating nucleotidase) (Benaroudj & Goldberg, 2000; Benaroudj *et al*., 2003; Navon & Goldberg, 2001; Wilson *et al.*, 2000; Zwickl *et al*., 1999a). To date, the only archaeal PAN to be characterized is that from *M. jannaschii* (Benaroudj & Goldberg, 2000; Benaroudj *et al.*, 2003; Navon & Goldberg, 2001; Wilson *et al.*, 2000; Zwickl *et al.*, 1999a). In general, the archaeal PAN is hypothesized to play some of the same roles as the 19S, PA28, and other proteasome regulatory complexes from the eukaryotes because of its predicted protein sequence similarity to the eukaryotic 19S regulatory complex (Bult *et al*., 1996; Voges *et al.*, 1999) and its observed characteristics that are reminiscent of the 26S regulatory complexes. Since the archaea and bacteria have not been found to contain ubiquitin, or homologs of ubiquitin-conjugating enzymes (Zwickl *et al.*, 1999a), the most likely alternative to the ubiquitin-dependent system is the presence of the ATP-dependent regulatory proteins, such as PAN and FtsH / cell-division control (CDC48) proteins. Note that the presence of ubiquitin has been reported in *T. acidophilum* (Wolf *et al*., 1993) and the bacterium *Anabaena variabilis* (Durner & Boger, 1995). These findings were not confirmed nor has ubiquitin, or any of the related enzymes, been found in the archaea and bacteria studied to the present, such that it is widely accepted that ubiquitin only exists in the eukaryotes. Therefore, a system for targeted proteolysis has not been determined, for the archaea or bacteria, that involves labeling by covalent attachment of ubiquitin. In the eukaryotes, the ubiquitin system is quite sophisticated, with the ubiquitin "tag" binding to the ε-amino group of Lys residues on the substrate protein through an energy-dependent reaction (Hershko & Ciechanover, 1998; Hochstrasser, 1996), followed by activation of several steps, each involving one or more separate enzymes, that ultimately lead to proteasome hydrolysis of the substrate protein. It is also not known whether a separate

system with a series of participating proteins, akin to the ubiquitin pathway, exists in prokaryotes or if, instead, the PAN, VAT (valosine-containing protein-like ATPase) (Zwickl *et al.*, 2000), CDC48 (Maupin-Furlow *et al.*, 2001; Zwickl *et al.*, 2000), and other ATP-dependent proteins are in control of the proteasome degradation pathway, including substrate targeting.

It has already been shown that the *M. jannaschii* PAN is able to associate with the ends of the *M. jannaschii* 20S proteasome in the presence of ATP (Wilson *et al.*, 2000). It also has the ability to stimulate 20S proteolysis of polypeptides (but not small peptides (Zwickl *et al.*, 1999a), unlike the eukaryotic versions), including β-casein (Wilson *et al.*, 2000; Zwickl *et al.*, 1999a), α-lactalbumin (Zwickl *et al.*, 1999a), and green fluorescent protein containing an 11-residue ssrA recognition peptide "tag" (Benaroudj & Goldberg, 2000), in the presence of both ATP and CTP (Benaroudj & Goldberg, 2000; Wilson *et al.*, 2000; Zwickl *et al.*, 1999a). Interestingly, this *M. jannaschii* PAN was able to stimulate polypeptide degradation by both the *M. jannaschii* (Benaroudj & Goldberg, 2000; Wilson *et al.*, 2000) and *T. acidophilum* (Zwickl *et al.*, 1999a) proteasomes. Furthermore, it was found to unfold protein substrates for the proteasome (Navon & Goldberg, 2001) and regulate the N-terminal "gate" that is present in the outer α ring of the proteasome. This α "gate," located at the N-termini of the α subunits, was initially found through deletion studies (Benaroudj *et al.*, 2003) and appeared to block substrate entry to the β subunit active sites buried within the structure. It is analogous to the α "gate" located through deletions in the yeast proteasome (Groll *et al*., 2000), which were found to be regulated by 19S (Kohler *et al*., 2001) or PA28-related activators (Whitby *et al*., 2000). Finally, the *M. jannaschii* PAN was found to control translocation of the substrate into the catalytic center of the proteasome

(Benaroudj *et al.*, 2003), similar to the mechanism for translocation used by the eukaryotic PA28 (Whitby, 2000). It was initially believed that the N-terminal "gate" of the α ring was always in an open state in the archaeal versions of the proteasome, based on crystal structure data (Lowe *et al.*, 1995) and the observation that these versions were able to digest small peptides without the presence of ATPases (Zwickl *et al.*, 1999a). However, it was recently found that the archaeal proteasomes do contain a gate that is not necessarily in a completely open state under normal conditions, proven by the ability of the proteasome to block entry of unfolded proteins in the absence of the PAN ATPase (Navon & Goldberg, 2001).

As shown in Table 6.1, many of the sequenced archaea contain one or two genes that encode for sequences similar to PAN. However, several have not been found to contain a PAN, specifically the *Thermoplasma* species (Ruepp *et al*., 2000) and *Pyrobaculum aerophilum* (Zwickl *et al.*, 1999a). Though unpublished on the National Center for Biotechnology Information web site (NCBI: http://www.ncbi.nlm.nih.gov/), it was also recently found that, like the other listed halophilic archaeon *Halobacterium* sp. NRC-1 (Ng *et al*., 2000), *Haloferax volcanii* does encode two PAN paralogs (Maupin-Furlow *et al.*, 2001). In the case of the archaeal species that do not contain PAN, other similar proteins may be key players in the proteasome degradation pathway. These proteins include VAT (Pamnani *et al*., 1997; Zwickl *et al.*, 2000) and FtsH / CDC48 (Maupin-Furlow *et al.*, 2001; Zwickl *et al.*, 2000), which were found through gene analysis or biochemical studies in *Thermoplasma* (Ruepp *et al.*, 2000) and *P. aerophilum* (Zwickl *et al.*, 1999a). The VAT and FtsH / CDC48 proteins have also been found in other archaeal species that contain PAN, including *Pyrococcus horikoshii* (Kawarabayasi *et al*., 1998) and *Aeropyrum pernix* (Kawarabayasi *et al*., 1999). It remains to be determined exactly how these regulating enzymes interact with

the archaeal proteasome and with each other and, furthermore, how closely related this system of regulation is to the ubiquitin system of the eukaryotes.

Pyrococcus furiosus is a hyperthermophilic archaeon with an optimal growth temperature of approximately 100°C (Fiala & Stetter, 1986). A native version of the proteasome, purified from *P. furiosus* grown at 95°C, was found to contain one α- and one β-subunit (referred to here as β2, encoded by the *psmB-2* gene PF0159), according to N-terminal sequencing (Bauer *et al*., 1997). However, the *P. furiosus* genome sequence (Robb *et al*., 2001), revealed the presence of a second β-subunit, referred to as β1 (*psmB-1* gene PF1404). Similarly, genome sequences from two other members of this genus, *P. horikoshii* (Kawarabayasi *et al.*, 1998) and *Pyrococcus abyssi* (Heilig, NCBI site: http://www.ncbi.nlm.nih.gov/), along other archaeal species such as *Sulfolobus solfataricus* (She *et al*., 2001) and *Sulfolobus tokodaii* (Kawarabayasi *et al*., 2001), *A. pernix* (Kawarabayasi *et al.*, 1999), and *P. aerophilum* (Fitz-Gibbon *et al*., 2002), also indicate the presence of both β1 and β2 subunits (Ward *et al*., 2002). The fact that two β subunits are encoded in the pyrococcal and other genomes yet only one form was detected in the native version (Bauer *et al.*, 1997) raises questions about the role of the second subunit. This issue is addressed here through transcriptional analysis of *P. furiosus* under normal and stressed growth conditions in conjunction with biochemical and biophysical analyses of various forms of the 20S and 26S proteasome.

III. EXPERIMENTAL PROCEDURES

Peptidase assays. Proteasome peptidase activity was determined with a microtiter plate reader (Model HTS 7000 Plus Bio Assay Reader, Perkin-Elmer, Wellesley, MA) by detection of 7-amino-4-methylcoumarin (MCA) released from the carboxyl terminus of N-terminally blocked peptides (Bauer *et al.*, 1997). Fluorogenic peptides and polypeptide substrates were purchased from Sigma-Aldrich (St. Louis, MO). Endpoint assays were performed in 50 mM sodium phosphate buffer pH 7.2 (unless otherwise noted), with 5 μM substrate, from a 10 mM stock in DMSO (stored at –20ºC for no longer than 2 months), and 200-300 ng enzyme used per 100-μl reaction. Substrate and buffer were combined in a 96-well microtiter plate on ice. When included in the assay, nucleotide (final concentration of 1 mM) and PAN (4:1 molar ratio PAN:proteasome) were then added. Finally, enzyme was combined and the plate was incubated at 95ºC in a hybridization oven (Continental Lab Products, San Diego, Ca.) for 15 minutes and cooled to room temperature for 10 min before fluorescence was obtained at λ_{em} = 360 nm and λ_{ex} = 465 nm.

Specific activity values were determined by scale-up of the 100-μl microtiter plate assays to a total volume of 250 μl in locking tubes. Substrate and buffer were initially combined and preheated to 95ºC for 3 min before addition of preheated (1 min at 95ºC) enzyme. In assays containing PAN and ATP, 1 mM ATP and a 4:1 molar ratio (PAN:proteasome) were added to substrate and buffer before the 3-min preheating at 95ºC. The mixture was incubated at 95ºC for 45 seconds and quenched on ice. Aliquots of 100 μl were then added to pre-chilled microtiter plates before fluorescence was obtained. Each assay was done in duplicate, with two 100-μl aliquots per assay analyzed in the microtiter

plate. For all assays, negative controls (no enzyme added) were run in triplicate or higher and subtracted from enzymatic assays to account for thermal degradation of the substrates.

Polypeptide and protein degradation assays. Degradation of polypeptides and proteins by the proteasome was analyzed by mixing polypeptides of varying lengths and sequences with the proteasome and viewing separation of the degradation products using reverse-phase high-performance liquid chromatography (RP-HPLC) (Waters, Milford, MA) with a C_{18} Nucleosil column in a triflouroacetic acid (TFA)/acetonitrile solvent system (Dib *et al*., 1998; Kannan *et al*., 2001). Polypeptide substrates were initially mixed on ice with 50 mM SPB pH 7.2 + 100 mM NaCl + 1 mM $MgCl_2$ + 1 mM DTT (DTT added immediately before use). Final substrate mass per assay varied from 18 µg for smaller chains (neurotensin, adrenocorticotropic 1-24) to 30 µg for medium-length chains (insulin chain B) and 100 µg for longer chains (β-casein). When included in the assay, nucleotide (final concentration of 1 mM) and PAN (specified molar ratio proteasome:PAN) were then added. Finally, proteasome (600 ng – 1.2 µg) was combined and the 200-µl mixtures were incubated at 80-90ºC for 2-4 hours. Mixtures were quenched on ice, followed by the addition of 1% TFA. The reactions were passed through Microcon centrifugal concentrators (30,000 - 100,000 MWCO) for filtration and the flow-through (50 µl) was injected onto the C_{18} column for degradation products to be separated in a linear gradient of 12-48% acetonitrile in 0.1% TFA over 15 min. The flow rate was 1 ml/min and peptide bonds were detected at an absorbance of 214 nm. Disappearance of insulin was quantified by comparing the concentration of insulin remaining per sample to that in the negative control (insulin without enzyme after exposure to assay conditions) through use of insulin peak areas (µV•sec)

calculated using the accompanying HPLC software. Percent hydrolysis was then calculated and normalized to the assay incubation time. These values were then normalized to the assay condition of 90ºC incubation temperature with enzyme only (no added ATP or PAN) in order to compare effects of temperature and PAN ATPase activity.

Nucleotidase assays. PAN activity on nucleotides was determined by release and detection of inorganic phosphate (P_i). Since phosphate buffer cannot be used in this type of nucleotidase assay, the initial buffer was 50 mM Tris pH 8.9 + 100 mM NaCl + 10 mM $MgCl_2$ + 1 mM DTT (nucleotidase Tris). However, with heating to 90ºC, the final pH of the assay mixture was approximately 6.9 (pKa/ºC = -0.031 for Tris), which was assumed acceptable based on the reported optimum pH range of 7-8 for the *M. jannaschii* version of PAN (Wilson *et al.*, 2000). The assay mixtures were similar to those reported for activity detection by other nucleotidases (Schirmer *et al*., 1998). The 1 mM DTT was added to the nucleotidase Tris from a 1 M stock solution immediately before the assay was performed. A premix of buffer and 1 mM (final concentration) nucleotide was initially combined, aliqotted to locking tubes, and chilled on ice. In the case of assays with added polypeptide, β-casein was then added at a 1,000-fold molar excess over PAN (3 µg from a 10 mg/ml stock in nucleotidase Tris). Finally, PAN (0.75 µg) was added to the cold mixture and the assays (final volume of 50 µl) were heated at 90ºC for 5 min. The reactions were quenched on ice for 5-10 min and P_i was detected using the modified molybdate/malachite green method (Lanzetta *et al*., 1979), with Triton X-100 substituted for Sterox (Bornemann *et al*., 1995; Kodama *et al*., 1986). The malachite green color reagent (0.034% malachite green, 10.5 g/L ammonium molybdate, 0.1% Triton X-100 in 1 M HCl; filtered through a Whatman #5 filter)

was added at a volume of 800 µl and, after 1 min, 100 µl of 34% sodium citrate was added. The reactions were mixed by vortexing and incubated at room temperature for 45 min - 1 hr for color development. Absorbance at 660 nm was then recorded and amount of P_i was calculated from a phosphate standard curve based on dilutions of KH_2PO_4. For all assays, enzymatic reactions and negative controls (heated nucleotide with no enzyme) were done in triplicate. In the case of added polypeptide, negative controls with β-casein were used to subtract the non-enzymatic degradation of the nucleotides.

Total protein assays. Total protein concentrations were determined using the Coomassie blue dye-binding method (Bradford, 1976) (Bio-Rad, Hercules, Ca.) in microtiter plates with bovine serum albumin (Sigma-Aldrich, St. Louis, Mo.) as the standard.

Purification of native P. furiosus 20S proteasome. *P. furiosus* (DSM 3638) biomass for purification of the native proteasome derived from cell growth under normal conditions (Bauer *et al.*, 1997) was kindly provided by Michael W. W. Adams, Department of Biochemistry, University of Georgia. *P. furiosus* was grown as previously described (Verhagen *et al*., 2001) at 95ºC, harvested, and rapidly frozen in liquid N_2 and stored at -70ºC before and after transport. Approximately 45 g of the frozen cells were thawed and re-suspended in 90 ml of 50 mM sodium phosphate buffer pH 8 (SPB) and disrupted in a French-pressure cell (Thermo Spectronic, Waltham, MA) (6 passes) at 18,000 psi. The cells were then centrifuged (10,000 x g, 4ºC) for 30 min and the supernatant separated before storage at 4ºC. The proteasome was purified by FPLC (Amersham Biosciences, San Francisco, CA), through four column steps, with identification after each step achieved using

activity assays on the flourogenic peptide N-carbobenzyloxy-Val-Lys-Met 7-amido-4-methylcoumarin (VKM-MCA) (see "*Peptidase assays*") supplemented with viewing on SDS-PAGE gels (10-12.5%). Pooled fractions were combined and concentrated using Centriplus centrifugal filter units of 10,000 MWCO (Millipore, Bedford, MA) and filtered through 0.45 μm membranes. In the first purification step, one-quarter of the cell extract (20 ml) was applied to a 250-ml DEAE (weak anion exchange) column using a 1-L linear gradient from 0-1 M NaCl in SPB, with proteasome elution occurring between 0.4-0.5 M NaCl. The fractions with highest VKM-MCA activity were then combined and applied to a HiLoad 16/10 Phenyl Sepharose HP (hydrophobic interaction) column and eluted during a 400-ml linear gradient from 30-0% ammonium sulfate in SPB. The sample was initially prepared by gradual addition of, and equilibration in, 30% ammonium sulfate after salt removal through dialysis (three 12-hr treatments in SPB). Elution of the proteasome occurred between 6-0% ammonium sulfate. After dialysis into SPB, concentration, and filtration, the fractions with the highest VKM-MCA were applied to an 80-ml hydroxylapatite (Calbiochem, San Diego, CA) XK 16 column. A 300-ml linear gradient from 50 mM – 0.5 M sodium phosphate buffer pH 8 allowed elution of the proteasome between 220-265 mM sodium phosphate. The dialyzed, concentrated, and filtered VKM-MCA active pool was then run through a HiPrep 26/60 Sephacryl S-200 HR gel filtration column in SPB + 150 mM NaCl. The proteasome primarily eluted in the void volume. After this column, the proteasome was relatively pure, with the exception of the glutamate dehydrogenase (PF1602), which tends to characteristically co-purify with the proteasome.

Identification of the proteasome and other proteins, such as the dehydrogenase, was performed at Research Triangle Institute Mass Spectrometry Institute (RTI, Research

Triangle Park, NC). The sample was separated on a 10% native PAGE gel and bands were individually digested using trypsin or endoproteinase Lys-C and denatured using reduction and alkylation. The resulting peptide digests were run via MALDI-TOF after being cleaned with a C_{18} zip tip and spotted onto plates. The peptide maps were compared to the *P. furiosus* protein database for identification.

Purification of native P. furiosus heat-shocked 20S proteasome. *P. furiosus* (DSM 3638) was grown on sea salts media (40 g/L sea salts (Sigma-Aldrich, St. Louis, Mo.), 3.3 g/L PIPES buffer, 1 ml/L trace elements (VSM), 5 g/L tryptone, 1 g/L yeast extract, 120 grams total sulfur). The culture (12 L) was grown in a 14-L fermentor (New Brunswick, Edison, NJ) at an agitation rate of 600 rpm, pressure of 0.5 bar, and sparge rate (N_2) of 0.5 L/min. The cells grew at 90ºC (pH of 6.2) for 8.75 hr and were heat shocked to 105.4 ± 0.3ºC (within 5 min) for 1 hr before being harvested and stored for 3 weeks at –20ºC. The 16-g cell pellet was re-suspended in 60 ml of 20 mM Tris pH 8 (Tris), passed (4x) through a French-pressure cell at 18,000 psi, and centrifuged (10,000 x g, 4ºC) for 30 min. The supernatant was applied to a 40-ml Q-Sepharose (strong anion exchange) XK 26/20 column and eluted between 0.5-0.7 M NaCl during a 600-ml linear gradient from 0-1 M NaCl in Tris. After many of the smaller contaminating proteins were cleared from the resulting VKM-MCA active pool in a Microcon centrifugal concentrator of 100,000 MWCO, the heat-shocked proteasome in the retentate was mostly pure, except for the presence of the same glutamate dehydrogenase (PF1602) that co-purified with the non-heat-shocked proteasome.

Cloning of the P. furiosus psm genes in Escherichia coli. The genes for the three subunits associated with the proteasome, including *psmA* (PF1571, "proteasome, subunit alpha (multicatalytic endopeptidase complex alpha subunit)"), *psmB-1* (PF1404, "proteasome, subunit beta (multicatalytic endopeptidase complex beta subunit)"), and *psmB-2* (PF0159, "proteasome, subunit beta (multicatalytic endopeptidase complex beta subunit)"), were separately cloned into the pET-24d(+) vector carrying the T7*lac* gene (Novagen, Madison, WI). Each of the *psm* genes were individually PCR-amplified (PfuTurbo DNA polymerase, Stratagene, La Jolla, CA) from *P. furiosus* genomic DNA with oligonucleotides (Integrated DNA Technologies, Coralville, IA), designed for creation of cut sites appropriate for ligation between the NcoI and BamHI sites within the cloning/expression region of the pET-24d(+) vector. The α and β2 genes were cloned using the open reading frames (ORFs) exactly as quoted on the NCBI website (http://www.ncbi.nlm.nih.gov/). The β1 gene was cloned using an ORF that started with the fourth amino acid from the quoted N-terminus, a more probable start site for expression of the *psmB-2* gene based on start codon and location of likely ribosomal binding site. The *psmA* and *psmB-1* PCR products, along with the pET-24d(+) vector, were cut with NcoI and BamHI (New England Biolabs, Beverly, MA; Promega, Madison, WI) before ligation. Since *psmB-2* has an internal NcoI cut site, this gene was cut with BspHI, which creates compatible cohesive ends with NcoI, and BamHI before ligation into the pET-24d(+) vector. For *psmA*, the PCR primers used were: forward (5' - TGA ACG CCA TGG CAT TTG TTC CAC CTC A - 3') and reverse (5' - ATA AAA ATT GGA TCC AAG TCA GTA GTT GCT ATC CA - 3'). For *psmB-1*, the PCR primers used were: forward (5' - TGT TGC CCA TGG AAG AGA AAC TTA AGG GAA - 3') and reverse (5' - AAA TTG TCG GAT CCT TGG ACT ACT TTA ACA TTT T - 3'). For *psmB-*

2, the PCR primers used were: forward (5' - TTA GGT GGT GCT CAT GAA GAA AAA GAC TGG AA - 3') and reverse (5' - TAA GGA AGC CTG GAT CCT TCA TAC TAC AAA CTC TT - 3'). All ligation products were transformed into an *E. coli* NovaBlue non-expression cloning host (Novagen, Inc., Madison, WI) and purified with Qiagen columns (Valencia, CA).

Expression of the P. furiosus psm genes in E. coli. All three *psm* genes were expressed in *E. coli* hosts containing the T7 RNA polymerase from the *lacUV5* promoter in λDE3 lysogens. The *psmA* gene was expressed in the *E. coli* general expression host BL21(DE3), while the two *psmB* genes were separately expressed in the *E. coli* expression host BL21-CodonPlus®(DE3)-RIL (Stratagene, La Jolla, CA) because of rare codon issues. This strain contains the *argU*, *ileY*, and *leuW* genes encoding $tRNA_{AGA/AGG}$, $tRNA_{AUA}$, and $tRNA_{CUA}$, respectively. Though *psmB-1* does not contain a high percentage of rare codons, it was found to express more efficiently in the CodonPlus strain. The *psmB-2* strain does have a high number of rare codons adjacent to each other, particularly those encoding for Arg and Ile, which makes expression in the CodonPlus strain much more efficient. Expression was induced with 0.4 mM IPTG at an OD_{595} of 0.60, as the pET-24d(+) vector is under control of the T7*lac* promoter. The cells were harvested 3-5 hr after induction (37ºC) and frozen overnight at –20ºC. Cells were thawed on ice and re-susupended in 4-6 volumes (wt/vol) of 50 mM sodium phosphate pH 7.2 + 1 mM DTT (SPB+DTT), a final concentration of 1 mg/ml lysozyme (from a fresh 10-mg/ml stock in water) was added, and the mixture was incubated at 30ºC and 100-150 rpm for 30 min. The cells were split into 8-ml aliquots and sonicated (Misonix, Inc., Farmingdale, NY) at a setting of 3.5 for 6-8 min (20 sec on / 20 sec

off) on ice. They were then centrifuged (18,000 x g, 4ºC) for 30 min and the supernatant was separated. After two 20-min heat treatments, the first at 85ºC and the second at 90ºC, each followed by cooling on ice for 30 min, spinning (18,000 x g, 4ºC) for 30 min, and re-suspension of resulting pellets in 1/3 the original volume (2 volumes original cell pellet) of SPB+DTT, each expressed subunit was at a different level of purity. In all cases where column purification of the recombinant subunits was necessary, screening for the subunits was done by first combining the subunit of interest with one of the other subunits directly in a microtiter plate (see "*Combining the expressed psm gene products to create the P. furiosus proteasome*") and then conducting the usual assay for VKM-MCA activity. Fractions of interest were then re-checked on 12.5% SDS-PAGE gels. The α subunit, expressed from the *psmA* gene, was most pure in the supernatant from the second heat treatment of the supernatant (first heat treatment) from the sonicated pellet. This was applied to a MonoQ (strong anion exchange) HR 5/5 in a 40-ml linear gradient from 0-1 M NaCl in SPB+DTT. The α subunit eluted in the range 0.3-0.5 M NaCl. The pooled fractions were then dialyzed into SPB+DTT with 150 mM NaCl, concentrated, and run through a HiPrep 26/60 Sephacryl S-200 HR gel filtration column, where the α primarily eluted at, and slightly after, the void volume. The β1 and β2 subunits, expressed from the *psmB-1* and *psmB-2* genes, respectively, were heat treated in the same way as the α and, after optimization, were pure without the need for column purification. In the case of β1, the purest and most concentrated form was located in the supernatant fractions after the second heat treatments that were derived from the pellets of both the pellet and supernatants from sonication. In most cases, β2 was only heat treated once at 85ºC for 20 min, cooled on ice for 30 min, and centrifuged (18,000 x g, 4ºC) for 30 min. The supernatant from the pellet after sonication was pure and

concentrated. In earlier expression runs, the β1 was passed through a 40-ml Q-Sepharose (strong anion exchange) XK 26/20 column and eluted between 0.2-0.3 M NaCl during a 600-ml linear gradient from 0-1 M NaCl in SPB+DTT. The fractions containing β1 were then dialyzed into SPB+DTT with 150 mM NaCl and eluted from a HiPrep 26/60 Sephacryl S-200 HR gel filtration column at several molecular weights, with the primary range near 25 kDa. There was some elution at higher molecular weights (approximately 45 and 63 kDa), indicating the presence of dimer and trimer forms of β1.

Testing for expression of correct subunits through TCA precipitation and N-terminal sequencing. Each expressed subunit was precipitated with 10% trichloroacetic acid (TCA), with the resulting protein pellet washed in ice-cold acetone 2X, and re-suspended in buffer. The resulting denatured subunits were then run on 12.5% SDS-PAGE gels to check for size. Each subunit was also electroblotted (FisherBiotech semi-dry blotting apparatus, Fisher Research, Pittsburgh, PA) onto a PVDF membrane (Bio-Rad, Hercules, CA) and the first 5 amino acid residues at the N-terminus of each subunit were determined. Amino-terminal sequencing of expressed subunits was performed by The University of Georgia Molecular Genetics Instrumentation Facility (Athens, Ga.).

Combining the expressed psm gene products to create the P. furiosus proteasome. For all experiments that tested stability and activity of the different proteasome forms, the subunits were combined in equimolar ratios in a final total protein concentration of 0.5 mg/ml. In the case where all three subunits were combined, each β subunit was added in an equimolar ratio with the α subunit, making the final total protein concentration 0.7 mg/ml for

these samples. Combined subunits were then incubated at the indicated temperature for 1 hr, cooled on ice for 1 hr, and precipitated material was removed through centrifugation (16,000 x g, 4ºC) for 30 min. Total protein concentration was determined for each sample and these values were used for subsequent assays. For screening during column purification, an initial assay had to be performed on crude expressed material to determine which combinations of subunits yielded VKM-MCA activity (see "*Peptidase assays*"). The α was combined with β1 and β2 in two separate samples to a final total protein concentration of 1 mg/ml. Each tube was heated to 85ºC for 1 hr, cooled on ice for 1 hr, and VKM-MCA activity levels checked. Once this was found, the subunit being purified (10 µl of each fraction from a column) was combined with a compatible active subunit (150-200 ng) directly in a microtiter plate. The plate was heated to 95ºC for 20 min to allow for assembly, VKM-MCA (4-5 µM) was added, and the plate was heated to 95ºC for another 20 min before fluorescence was determined.

Cloning and expression of the P. furiosus PAN gene in E. coli. The PAN gene (PF0115), also known as the "ATP-dependent 26S protease regulatory subunit" based on sequence data, was cloned into the pET-21b(+) vector (Novagen, Madison, WI). The PAN gene was PCR-amplified (PfuTurbo DNA polymerase) from *P. furiosus* genomic DNA using the following forward and reverse primers, respectively, so that the PCR product could be ligated into the pET-21b(+) vector at the NdeI and BamHI cut sites within the cloning/expression region: forward (5'- GGT GAT ACA TAT GAG TGA GGA CGA AGC TCA ATT T – 3') and reverse (5' - TAA AAA TTA GGA TCC TCA GCC GTA AAT GAC TTC A - 3'). After ligation, clones were amplified in an *E. coli* NovaBlue cloning host

(Novagen, Inc., Madison, WI) and purified with Qiagen columns (Valencia, CA). Because of rare codon issues (particularly with Arg and Ile), the clones were expressed in the *E. coli* expression host BL21-CodonPlus®(DE3)-RIL (Stratagene, La Jolla, CA). This strain contains the T7 RNA polymerase from the *lacUV5* promoter in λDE3 lysogens, and the *argU*, *ileY*, and *leuW* genes encoding $tRNA_{AGA/AGG}$, $tRNA_{AUA}$, and $tRNA_{CUA}$, respectively. Expression was induced with 0.4 mM IPTG at an OD_{595} of 0.60, as the pET-21b(+) vector is under control of the T7*lac* promoter, and cells were harvested 3-5 hr after induction (37ºC) and frozen overnight at –20ºC. Cells were thawed on ice and re-suspended in 4-6 volumes (wt/vol) of 20 mM Tris pH 8 + 0.5% CHAPS + 1 mM DTT; a final concentration of 1 mg/ml lysozyme (from a fresh 10-mg/ml stock in water) was added, and the mixture was incubated at 30ºC and 100-150 rpm for 30 min. The cells were split into 8-ml aliquots and sonicated at a setting of 3.5 for 6-8 min (20 sec on / 20 sec off) on ice. The resulting extract was then centrifuged (18,000 x g, 4ºC) for 30 min and the supernatant was separated. The pellets were re-suspended in 1/3 the original volume (2 volumes original cell pellet) of 20 mM Tris pH 8 + 0.5% CHAPS + 1 mM DTT and heat-treated at 85ºC for 20 min. Re-suspensions were cooled on ice for at least 30 min and centrifuged at (18,000 x g, 4ºC) for 30 min. Pellets were then re-suspended as before the heat treatment and all fractions were heat treated again at 90ºC for 20 min with cooling and centrifugation as before. With this type of gradual heat treatment (the first heat treatment at a lower temperature than the second) combined with the use of CHAPS, a high concentration of expressed PAN (10-15 mg/ml) is successfully purified without the need for column purification. Without the CHAPS, it was found that the PAN remains associated with the cell wall debris (Loe *et al*., 1989; Zhao *et al*., 1991). The highest concentration of PAN, and the one used in all experimental procedures, was that in

the supernatant from the heat treatment of the pellet that was formed during the first heat treatment.

Differential scanning calorimetry of recombinant P. furiosus proteasome and PAN. The melting temperatures of all expressed proteins were determined using a CSC nano differential scanning calorimeter (DSC; Calorimetry Sciences Corp., American Fork, UT). Samples of 0.5-1 mg/ml were degassed and heated from 25-125ºC at a rate of 1ºC/min in two heating and cooling steps. Resulting heat capacity versus temperature curves were compared to baseline buffer scans in the software program accompanying the DSC instrument to determine melting temperatures. After each sample was analyzed on the DSC, activity assays and native gels were run on the proteasomes to determine if complete or irreversible denaturation had occurred. Samples were centrifuged (16,000 x g, 4ºC) to remove aggregates, total protein concentrations were obtained, and activity assays were run simultaneously against the corresponding mixture in the "unmelted" state to obtain relative loss of activity. Activity of melted proteasomes were monitored every ~5 days for at least 30 and up to 60 days following DSC analysis.

Targeted cDNA microarray analysis of dynamic heat shock response. A targeted cDNA microarray containing 201 genes, including those corresponding to *P. furiosus* proteolysis, was used to follow transient transcriptional response after a temperature shift from 90ºC to 105ºC. *P. furiosus* (DSM 3638) was cultured anaerobically at 90ºC on Sea Salts Medium (SSM), as described previously (Shockley *et al*., 2003). Tryptone (Sigma, St. Louis, MO) was added to SSM (final concentration 5 g/L) as a carbon source prior to

inoculation, along with elemental sulfur (10 g/L). A 50 ml batch culture was used to inoculate 500 ml of SSM medium supplemented with 5 g/L tryptone and 10 g/L sulfur in a 1-L pyrex bottle. Two hundred fifty ml of this culture was added to 13 L of media in a 14-L fermentor (New Brunswick Scientific, Edison, NJ). The fermentor contained an internal temperature controller, and the pH was maintained by a Chemcadet controller (Cole Parmer, Vernon Hills, IL). High purity N_2 was used to reduce the medium and to sparge during inoculation. The culture was grown to mid-log phase at 90ºC, after which a sample was collected. The temperature set point was then shifted to 105ºC, with the culture taking approximately 2 min to reach the set point temperature. Once the culture reached 105ºC, samples were taken at 0, 5, 10, 60 and 90 min. Approximately 20 ml of culture were collected prior to sampling at each time point to eliminate pre-existing fluid in the sampling lines. At each time point, 500 ml of culture were withdrawn and immediately put on ice until processed for RNA extraction (see below). One ml of sample was removed for cell density enumeration by epifluorescent microscopy with acridine orange stain (Hobbie *et al*., 1977).

RNA was extracted from each 500-ml sample culture as described previously (Shockley *et al.*, 2003). In short, the 500-ml samples from the fermentor were centrifuged for 20 min (10,000 x g, 4ºC) and the pellets were rinsed twice in 300 mM NaCl (Fisher Scientific, Pittsburgh, PA). After treatment with RNA lysis buffer, the samples were stored at –70ºC. Extractions proceeded with ethanol precipitation and purification using Promega Total SV RNA kits (Promega, Madison, WI). Concentrations and degree of purity were determined by optical density at 260 nm and 280 nm, as well as with gel electrophoresis (1% agarose gel, 60 V). Procedures for reverse transcription reactions, aminoallyl-labeling with Cy3 and Cy5, and hybridization reactions are reported in Chhabra et al. (2003).

A targeted cDNA microarray including 402 open reading frames (ORFs) was printed, following protocols described previously (Chhabra *et al.*, 2003). A loop experimental design incorporated reciprocal labeling of most time point samples with both Cy3 and Cy5. The 10-min sample was labeled only with Cy3. Mixed model analysis was used to evaluate differential expression data using approaches presented elsewhere (Chhabra *et al.*, 2003). Briefly, least squares estimates of gene-specific treatment effects, corrected for global and gene-specific sources of error, were used to construct pair-wise contrasts analogous to fold changes for each gene between all pairs of conditions. The statistical significance of these fold changes was determined and a Bonferroni correction was used to establish an experiment-wide false positive rate of $\alpha = 0.05$ by dividing α by 2,821, the number of comparisons performed for all genes over all possible treatment pairs. The corrected false positive rate was 1.77×10^{-5} (corresponding to a $-\log_{10}(\text{p-value}) > 4.8$). Least squares estimates of gene-specific treatment effects were also used to perform hierarchical clustering in JMP 5.0 (SAS Institute, Cary, NC).

IV. RESULTS

Transcriptional analysis of proteasome gene expression during P. furiosus heat shock response. Shockley et al. (2003) used a targeted cDNA microarray to compare *P. furiosus* differential gene expression before and after a temperature shift from 90ºC to 105ºC. Here, transient expression of the genes encoding proteasome-related components and other proteases was examined for the same temperature shift (see Table 6.2). Differential gene expression data, acquired using a targeted cDNA microarray, showed that the genes encoding the thermosome (HSP60), small heat shock protein (HSP20), and VAT1/2 were induced 3.7-fold or higher at some point during the 60-min period following a temperature shift from 90ºC to 105ºC, which was consistent with the results previously reported (Shockley *et al.*, 2003). The gene encoding proteasome β1 was not affected during the first 5 min of heat shock, but was induced two-fold after 10 min and remained this way until at least 60 min after heat shock. However, the gene encoding proteasome β2 subunit was not affected during the initial stages of heat shock, with a two-fold induction appearing after 60 min. Interestingly, the α subunit was down-regulated approximately 2.5- to 4-fold during the entire heat shock period. The gene encoding PAN was initially stimulated up to 15-fold in the first 10 min following the temperature shift but returned to 90ºC expression levels after 60 min. These transcription data showed that no significant stimulation of the proteasome components (α, β1, and β2 subunits) was observed during thermal stress response, which brings into question the thermal stability of the proteasome at physiologically-elevated temperatures. Native and recombinant versions of the proteasome were then examined from this perspective.

Expression products for the P. furiosus psm genes in E. coli. To examine the thermal stability and assembly of the proteasome at varying temperatures, the recombinant forms of the α, β1, and β2 subunits were produced. Based on amino acid sequence data (ExPASy Compute pI/MW tool; http://us.expasy.org/tools/pi_tool.html), the hypothetical molecular masses for the *psm* gene products were as follows: *psmA* (α) = 29,009, *psmB-1* (β1) = 22,002, and *psmB-2* (β2) = 21,649. For both of the β subunits, which were cloned in full form, with the apparent pro-peptide region of 6-7 N-terminal residues upstream of the active-site Thr tripeptide (TTT) present, these hypothetical molecular masses included the pro-peptide. Archaeal proteasomes are known to be resistant to, and even stimulated by, a variety of denaturing agents, such as heat, detergents, guanidine hydrochloride, and other more specific inhibitors (Akopian *et al.*, 1997; Goldberg *et al.*, 1997; Lupas *et al.*, 1998). Therefore, once the subunits were expressed, they were examined to verify molecular weights following full denaturation via TCA precipitation. In addition, the 5 amino acids at the N-terminus of each subunit were determined after electroblotting of each subunit onto a PVDF membrane. When the purified gene products were run on a 12.5% denaturing SDS-PAGE gel, both before and after TCA precipitation, the observed sizes were within range of the expected values. As shown in Figure 6.1, the α subunit ran at 30,000, β1 ran at approximately 23,000, and β2 ran slightly large at 26,000 Da. In addition, the N-terminal sequences were as expected for all three of the expressed subunits, with both β subunits containing the pro-peptide regions after expression and before combination with the other subunits.

Characterization of the individual P. furiosus subunits (native and recombinant versions). The native *P. furiosus* proteasome was originally purified from biomass grown under the usual *P. furiosus* growth conditions by Bauer et al. (1997) and was found to be 640 kDa, consisting of multiple copies of one α-type and one β-type subunit. The N-terminal amino acid sequences from these subunits matched those of the α and the β2 subunit lacking the pro-peptide region. However, the significant presence of the β1 subunit was not reported during that study, which preceded the report of the *P. furiosus* full genome sequence (Robb *et al.*, 2001). During the in-gel digestion by Lys-C and MALDI-TOF characterization of the native proteasome from *P. furiosus* cells grown under normal conditions, the presence of all three subunits was verified. As well, the analysis of a separate band within the same sample showed that the α subunit was present as an individual structure. This band can be viewed in gel 2 of Figure 6.4, where lane 1 is the native proteasome and lanes 2-3 are the recombinant α subunit. The band analyzed by MALDI-TOF and found to be the α subunit is represented by the lighter band immediately below the top band (the proteasome), which runs in-line with the recombinant α in the next two lanes. In addition, it was found that the recombinant form of α appeared large enough that it was unable to migrate much further than the full proteasome in a 10% native PAGE gel (see Fig. 6.4, gel 2). It was able to remain stable and in a larger assembly after heating for 1 hr at 105ºC. Furthermore, its melting temperature appeared to be above 120-125ºC (DSC limit), such that it did not show any transition, even when tested in two separate scans and at the upper level of the concentration range used for all samples. At the same time, when the α samples after DSC analysis were combined with β2 and heated to 90ºC, very little activity was recovered (~3% activity). In this case, it is

possible that the melting point of α is very close to the upper DSC limit of 125ºC, high enough that it could not be analyzed with confidence (within the range 120-125ºC).

The purified forms of the recombinant β1 and β2 subunits appeared in as many as 3 individual bands per subunit on a 10% native PAGE, suggesting that they were able to form assembled structures. However, the β subunits were not able to process their own pro-sequences, as is noted during proteasome assembly, since their N-terminal analyses showed the presence of the pro-sequences on the recombinant expressed versions. Additionally, neither β subunit exhibited peptidase or proteinase activity as individual subunits or when combined with each other at any of the tested assembly temperatures and conditions (see below). Although these two β subunits are 48% identical and 69% similar to each other on the amino acid level and both contain the characteristic sequence of 3 Thr residues (active-site "TTT") beyond each of their pro-sequences, they also exhibited many differences. One particular difference that was clearly noted was their stability. During storage, the β2 subunit tended to fall out of solution over time. In addition, its melting temperature was found to be 93.1ºC during DSC analysis as shown in Figure 6.2a. This was relatively low for an enzyme from a hyperthermophilic microorganism, such as *P. furiosus* (growth T_{opt} = 100ºC, (Fiala & Stetter, 1986)) and for a subunit from an enzyme that exhibits the highest peptidase activity at a temperature of 95ºC (Bauer *et al.*, 1997). In contrast, the melting temperature of β1 was found to be 104.4ºC (Fig. 6.2a). Finally, when each of the β subunits was heated in assembly conditions at 90ºC, the β1 was clearly visible in various forms on a native gel, whereas the β2 was too unstable to withstand the heating and/or electrophoretic conditions.

Combining the recombinant P. furiosus subunits. Bauer et al. (1997) found that the native *P. furiosus* 20S proteasome was most active on the fluorogenic amyloid A4-generating enzyme substrate VKM-MCA (Ishiura *et al*., 1990) compared to the other tested MCA-linked peptides. Therefore, this substrate was used, along with comparisons against the native purified proteasome, to test the activity of the various recombinant forms. Each individual subunit did not exhibit activity against VKM-MCA when initially expressed and stored at 4ºC or when heated at 85ºC, 90ºC, or 105ºC under the usual assembly conditions used in this study. As well, when β1 and β2 were combined in a 1:1 molar ratio, they were not active against VKM-MCA at incubation temperatures of 4ºC, 85ºC, 90ºC, and 105ºC. Figure 6.3 shows the relative activity of the combinations that did exhibit activity (compared to the native proteasome). The combinations α + β2 (1:1 molar ratio) and α + β1+ β2 (1:1 molar ratio of each α to β) both exhibited full or nearly full activity after incubation at 90-105ºC. The α + β2 combination tended to decrease in activity with increased assembly temperature, while the α + β1+ β2 combination was most active after assembly at 105ºC. This trend leads to the conclusion that the presence of β1 in the mixture has an affect on the assembly of the entire structure, with a positive affect occurring upon increase in temperature. Both of these combinations were also marginally active without the high-temperature incubation, however, strong activity from these forms was only observed after the incubation step. The α + β1 (1:1 molar ratio) combination did not show a high level of activity, even after the incubation step. When higher levels of β1 were added so that α:β1 molar ratios were 1:5 and 1:10, activity increased slightly with increased β1, but was still extremely low in comparison to the native form and the other recombinant forms of the proteasome (unpublished data). Beyond

the comparative high stability of β1 compared to β2, it was also observed that β1 combined with α in a 1:1 ratio increased in VKM-MCA activity levels as the assembly incubation temperature was increased. Although the highest activity of this assembly was still significantly below that of the native and other active recombinant forms, the increase in activity was consistent and noticeable when moving from an incubation temperature of 4ºC to 105ºC. The $\alpha + \beta 1$ combinations at 4ºC, 90ºC, and 98ºC had undetectable activity, however, at an incubation temperature of 105ºC, there was some positive activity (Fig. 6.3). Although the relative activity of the $\alpha + \beta 1$ combinations assembled at 4-98ºC and 105ºC went from 0% to ~1% compared to the native proteasome, respectively, this activity was still far below that of the $\alpha + \beta 2$ and $\alpha + \beta 1 + \beta 2$ combinations even at an incubation temperature of 4ºC, each of which gave a relative (to native proteasome) activity of ~15% (unpublished data).

Peptidase activities of the native and recombinant P. furiosus proteasomes. Since the $\alpha + \beta 1$ form of the proteasome exhibited relatively low activity on VKM-MCA, there remained the possibility that β1 exhibited a different type of activity compared to β2. Therefore, all of the recombinant proteasome subunits from *P. furiosus* were tested in endpoint assays for activity against 16 different N-terminally blocked MCA-linked substrates from 6 proteolytic categories, including chymotrypsin (C), trypsin (T), peptidyl-glutamyl peptide hydrolyzing (PGPH), small neutral amino acid peptide hydrolyzing (SNAAP), amyloid A4-generating enzyme (Amy), and 20S proteasome (20S). For all assays, a fixed mass of enzyme was added, based on the concentrations of the assemblies after each incubation at 90ºC or 105ºC. As with VKM-MCA, each individual subunit was inactive on

all of the substrates tested. The combinations β1+ β2, assembled at temperatures of 90ºC and 105ºC, were also inactive on all substrates. As with tests done on VKM-MCA (Amy substrate), the α + β1 combination, assembled at 90ºC and 105ºC, was inactive on almost all tested substrates, with some residual activity on VKM-MCA and the 20S substrate N-carbobenzyloxy-Leu-Leu-Leu MCA (LLL-MCA) (unpublished data). The combinations α + β2 and α + β1+ β2, at assembly temperatures of 90ºC and 105ºC, were inactive against the SNAAP substrate N-succinyl-Ile-Ala MCA (IA-MCA), the C substrates N-succinyl-Ala-Ala-Pro-Phe MCA (AAPF-MCA) and N-succinyl-Leu-Tyr MCA (LY-MCA), the T substrates N-carbobenzyloxy-Phe-Arg MCA (FR-MCA), N-*tert*-butyloxycarbonyl-Phe-Ser-Arg MCA (FSR-MCA), and N-carbobenzyloxy-Gly-Gly-Arg MCA (GGR-MCA), and the PGPH substrate N-acetyl-Val-Glu-His-Asp MCA (VEHD-MCA). Both combinations, at both assembly temperatures, had some residual activity on the C substrate N-succinyl-Ile-Ile-Trp MCA (IIW-MCA), the T substrates N-benzoyl-Phe-Val-Arg MCA (FVR-MCA) and N-*tert*-butyloxycarbonyl-Val-Leu-Lys MCA (VLK-MCA), and the PGPH substrate N-acetyl-Try-Val-Ala-Asp MCA (YVAD-MCA). Table 6.3 shows the relative activities of the two combinations against the other tested substrates, including 2 C, a PGPH, a 20S, and an Amy substrate. As was found in Bauer et al. (1997), the proteasome was most active on VKM-MCA. Approximately 50% activity was found for all four combinations on the 20S substrate LLL-MCA. There were differences in the types of activities exhibited by the proteasomes assembled at the two temperatures, implying that there may be differences in the composition of each assembly. These differences were noticeable on VKM-MCA, the two C substrates N-succinyl-Ala-Ala-Phe MCA (AAF-MCA) and N-succinyl-Leu-Leu-Val-Tyr MCA (LLVY-MCA), and the PGPH substrate N-carbobenzyloxy-Leu-Leu-Glu MCA (LLE-MCA).

Specific activity values were determined for the 90ºC and 105ºC assemblies $\alpha + \beta 2$ and $\alpha + \beta 1 + \beta 2$, as well as for these combinations assembled at an intermediate temperature of 98ºC; values are listed in Table 6.4. These specific activity values were within the range of those found for the *M. jannaschii* proteasome on similar peptide substrates (Maupin-Furlow *et al.*, 1998), though they were about 10-fold less than the reported values for the native *P. furiosus* proteasome (Bauer *et al.*, 1997). These differences could have been a combination of several factors: 1) underestimation of the mass of proteasome per assay (mass was based on total protein, which in this study included full proteasome along with partially formed proteasome and individual subunits); 2) the use of pH that was outside of the reported optimum for the proteasome (pH selected was slightly higher than the reported 6.5 (Bauer *et al.*, 1997), since the optimum pH of the PAN, used in later studies, is above neutral (Wilson *et al.*, 2000)); and 3) simple differences in assay technique which, with the extreme sensitivity of the fluorescence assays to all conditions, including temperature fluctuations, can alter final readings. In general, the specific activity values for $\alpha + \beta 2$ tended to drop with increasing assembly temperature, whereas the opposite occurred for the $\alpha + \beta 1 + \beta 2$ proteasomes. And, in fact, the specific activity of the 90ºC $\alpha + \beta 2$ assembly was approximately equal to that of the 105ºC $\alpha + \beta 1 + \beta 2$ assembly. The $\alpha + \beta 2$ mixture assembled at 105ºC may have had reduced activity due to the instability of the β2 subunit, particularly noticeable in both the endpoint and specific activity values on VKM-MCA. Much of the β2 may have denatured at the higher temperature (12ºC above the melting temperature), even in the presence of the α rings, before being incorporated into the proteasome. However, full activity was restored at the same assembly temperature when the

β1 was added during assembly of the $\alpha + \beta1 + \beta2$ mixture. This restored activity was one piece of evidence that indicates that β1, although not active on peptides when it was alone in the $\alpha + \beta1$ proteasome, was able to increase the assembly efficiency through stabilization or was able to stimulate catalysis by the β2 through direct interaction with α and/or β2.

Direct viewing of the recombinant P. furiosus proteasomes. Figure 6.4 illustrates the differences between the proteasomes as a function of subunits present and assembly temperatures. In gel 1 (12.5% SDS-PAGE) the first two lanes show the native proteasome from heat-shocked biomass (NHS proteasome) and the native proteasome from *P. furiosus* cells grown under normal conditions. The largest band, which did not migrate once it reached the separating gel, was the proteasome and the smaller 50-kDa band was the glutamate dehydrogenase. The next 6 lanes were the $\alpha + \beta1$ (lanes 3-4), $\alpha + \beta2$ (lanes 5-6), and $\alpha + \beta1 + \beta2$ (lanes 7-8) assembled at 90ºC (odd-numbered lanes) and 105ºC (even-numbered lanes). The two versions of $\alpha + \beta1$, which were relatively inactive in peptidase studies, were also either not assembled in a high enough concentration to view, or were unstable enough to completely denature under the conditions used to prepare the sample for SDS-PAGE (equal masses of total protein were loaded onto the gels). As was seen with both native forms and the active recombinant forms $\alpha + \beta2$ and $\alpha + \beta1 + \beta2$, the active forms were all stable enough to remain partially intact during electrophoresis so that the full proteasome was concentrated enough to view with Coomassie staining (see bands at top edge of gel). Gel 2 (10% native PAGE) gives a view of the native proteasomes in lanes 1-2 (band at top edge of gel, verified by in-gel tryptic digestion and MALDI-TOF) compared to the

various recombinant proteasomes. The native proteasome in lane 1 was that purified from the heat-shocked *P. furiosus* (NHS), while that in lane 2 was purified from *P. furiosus* grown under normal conditions. Lanes 3-4, 5-6, and 7-8 were the α, β1, and β2 subunits after incubation at 90ºC (odd-numbered lanes) and 105ºC (even-numbered lanes). Both the α and β1 remained stable whereas β2 (no visible bands) was found to be too unstable for the conditions of the incubation and electrophoresis. Lanes 9-10 were the combinations of β1 + β2 after incubation at 90ºC (lane 9) and 105ºC (lane 10), which were previously found to be inactive on peptide substrates. At the 90ºC assembly, the β subunits were still visible, but not in a form similar to the full proteasome structure. At the 105ºC assembly, a proteasome-like structure was not visible either, as was expected, since without the presence of α, the β subunits were found to be incapable of processing the pro-sequence and assembling into full rings. In addition, the β subunits were not visible in any of the smaller forms seen in lane 6, even though the isolated β1 was previously stable enough to remain visible on a gel after incubation. It is likely that the two β subunits were interacting in solution, but without α, this interaction became a destabilizing factor. Lanes 11-12 contained the α + β1 mixtures after assembly at 90ºC (lane 11) and 105ºC (lane 12). The presence of proteasome was questionable in both lanes, however, the α subunits were clearly seen (top band) in both lanes and β1 was seen in the 90ºC assembly. As with the β1 + β2, however, the β1 was not visible at the higher incubation temperature. The α + β2 combinations were in lanes 13-14, with the 90ºC assembly in lane 13 and the 105ºC assembly in lane 14. Full proteasome structures were clearly visible, with more of the isolated α ring appearing at the 105ºC temperature. It is likely that this extreme temperature caused a high level of β2 to

destabilize before it had a chance to assemble with α. However, even at the higher temperature, the presence of α allowed β2 to process and assemble in conditions that were previously found to fully denature the isolated form. The last two lanes contained all three of the subunits in combination, with lanes 15 and 16 representing the 90ºC and 105ºC assemblies, respectively. At 90ºC, it was clear that a significant level of β1 was not incorporated into the proteasome, since this subunit was visible at a high concentration in the lower 3 bands. However, at 105ºC, β1 was no longer visible. In the case where the three subunits were assembled at 98ºC, an intermediate amount of β1, compared to the same mixtures assembled at 90ºC and 105ºC, was seen on a native gel (unpublished data). Therefore, as the assembly temperature was increased, the amount of β1 present in solution (outside of the proteasome structure) decreased, as was consistently seen after several separate assembly experiments.

Differential scanning calorimetry on recombinant P. furiosus proteasomes. Figure 6.2b summarizes the melting point curves for the active forms of the recombinant proteasomes, including $\alpha + \beta 2$ and $\alpha + \beta 1 + \beta 2$ assembled at 90ºC and 105ºC. All of the melting points fall within the range 110.5-112ºC, with the assemblies containing β1 at the higher end of the range. The one variation that was observed was the sizeable peak in the $\alpha + \beta 1 + \beta 2$ mixture assembled at 90ºC (curve C) at the melting point of β1, which was found to be 104.5ºC. This indicated that a high concentration of β1, unassociated with the proteasome, was present in that mixture. It is not, however, present in the same mixture that was assembled at 105ºC (curve D). When viewed on a gel (unpublished data), all of the

melted post-DSC samples still contained the full proteasome in a concentrated enough form to view with Coomassie staining. Furthermore, for all of the proteasome forms, activity on VKM-MCA remained at 25-50% after DSC analysis. The $\alpha + \beta 2$ forms retained approximately 25% activity, while the $\alpha + \beta 1 + \beta 2$ form assembled at 105ºC retained 35% activity. The $\alpha + \beta 1 + \beta 2$ assembled at 90ºC retained at least 50% activity, but only after approximately 25 days of storage. Within the first 25 days after analysis, only 25% activity was observed. However, over time, this sample regained a considerable amount of activity (beginning at 25 days and remaining at 50% until at least 60 days after the DSC scans were completed on this mixture). Whether the presence of excess β1 in this sample allowed for re-assembly of the proteasome structure during storage is unclear. The individual subunits were previously seen to have the ability to process the β subunits and assemble into active proteasomes to gain approximately 15% activity within only 1 hr of being combined and incubated at 4ºC, with the absence of the high-temperature incubation. It is likely that, given several weeks, any reversibly denatured subunits still in the DSC samples would also be able to create full proteasomes even without incubation at the optimum assembly temperature. However, the interactions that occurred in the 90ºC $\alpha + \beta 1 + \beta 2$ between the free β1 and the rest of the structure may have curbed the higher level of irreversible denaturation that occurred in the other three mixtures. The one point that makes it particularly difficult to elucidate a complete understanding of these observations is the fact that the reported melting points did not represent full unfolding of the proteasomes, since, with such a complex structure, several transitions must occur before complete degradation is achieved. Based on the most likely assembly pathway of archaeal proteasomes, the first transition may simply be creation of half-proteasomes from the full structures, followed by removal of β subunits from

α rings, and finally dissociation of the α subunits from the α_7 rings. From the observed results in this study, it is clear that all four forms of the proteasome were not fully denatured, even after 2 heating steps to 125ºC.

Comparing stability among the recombinant P. furiosus proteasome forms. A high-temperature incubation of the active recombinant proteasome forms ($\alpha + \beta2$ and $\alpha + \beta1 + \beta2$ assembled at 90ºC and 105ºC) was used to compare their stabilities. Each mixture was adjusted to a baseline concentration of 0.15 mg/ml and incubated at 115ºC in an oil bath for up to 12 hr. Aliquots were taken at time points from 0-12 hr and stored on ice until the end of the incubation period. The usual VKM-MCA microtiter plate assay was then used to compare the activities of the mixtures (300 ng enzyme, based on pre-incubation concentration, was mixed with 5 μM VKM-MCA and heated to 95ºC for 15 min). The resulting fluorescence scores, with average background values subtracted, are illustrated in Figure 6.5a for the four proteasome forms. The corresponding decay constants (k_{obs}) were calculated using the equation $\nu/\nu_o = \exp(-k_{obs}t)$ (Eqn. 1) (Copeland, 1996), where ν/ν_o is represented by the fluorescence value before incubation divided into each fluorescent value of the incubated samples. The natural log of both sides of Eqn. 1 were obtained, giving the new equation $\ln(\nu/\nu_o) = -k_{obs}t$ (Eqn. 2). Linear regression was then performed to obtain the value for k_{obs} (the negative slope) for each proteasome form. The calculated k_{obs} for 90ºC $\alpha + \beta2$, 105ºC $\alpha + \beta2$, and 90ºC $\alpha + \beta1 + \beta2$ were all relatively close, ranging from 0.15-0.19, as shown on Fig. 6.5a. In contrast, the calculated k_{obs} for the 105ºC $\alpha + \beta1 + \beta2$ was $k_{obs} = 0.025 \pm 0.004$, indicating a much more stable structure compared to the other three. The higher level of stability is also clear on the graph of fluorescence against pre-incubation

time (Fig. 6.5a), which drops only slightly for the 105ºC $\alpha + \beta 1 + \beta 2$ proteasome. In addition, all four mixtures were viewed on a 10% native gel (unpublished data) after the 12-hr incubation. The 105ºC $\alpha + \beta 1 + \beta 2$ structure was the only one of the four that contained a high enough concentration of proteasome to be visible with Coomassie staining.

Polypeptide degradation by the P. furiosus native, native heat-shocked, and recombinant proteasomes. Degradation of insulin chain B (oxidized from bovine pancreas, 3496 Da, Sigma #I6383) was tested with all native and recombinant forms of the proteasome and remaining insulin versus degradation products were viewed after separation through RP-HPLC. In addition, the levels of insulin degradation were quantified by integrating the insulin peaks. These were then compared to the negative control case, where insulin was incubated without enzyme at the assay conditions used. Since the native forms contained the glutamate dehydrogenase, the amount of protease per assay was an approximation. As well, with the recombinant forms, masses of active proteasome structures were unknown, since it was clear in all previous data that the subunits were not completely incorporated into the proteasome structures during assembly. However, levels of degradation were comparable between the recombinant forms, with the knowledge that the differences were not only representing catalytic ability of the proteasomes, but instead were representing the coupling of assembly efficiency (given a set amount of raw material, i.e. proteasome subunits, how much active proteasome was made during the given assembly period) with catalytic ability (how well did the assembled material degrade the substrate).

For all proteasomes, activity on insulin B was tested at two assay temperatures: 80ºC and 90ºC. As seen with the MCA-linked peptides, the $\alpha + \beta 1$ form was not active on insulin

B at either assay temperature. Although extent of proteolysis was variable, the degradation patterns, represented by the peaks on the HPLC chromatograms, were identical for the native and recombinant forms $\alpha + \beta 2$ and $\alpha + \beta 1 + \beta 2$. However, the NHS proteasome gave a very different degradation pattern, as shown in Figure 6.6. Part a) of Fig. 6.6 compares the degradation patterns of the NHS proteasome and the $\alpha + \beta 1 + \beta 2$ assembled at 105ºC, which produced the same peaks (though at different peak absorbance values) identical to that of the native and other recombinant forms. Particularly noticeable was the peak at 13 min that was present in the NHS, but not in the recombinant trace. As well, the peaks between 13.5 and 14 min were both shifted and at different absorbance units relative to other peaks within the same chromatograms. Beyond the degradation pattern, the activities of the native compared to the NHS also varied with incubation time and temperature. For the native proteasome, activity was higher at the lower incubation temperature of 80ºC. The NHS proteasome gave the opposite trend, with considerably lower activity at the lower incubation temperature of 80ºC (see Table 6.5a). When the incubation time was increased at the 80ºC incubation temperature, further degradation of insulin occurred in both the native (and recombinant) and NHS cases. As shown in part b) of Fig. 6.6, the difference in degradation patterns between the two proteasome forms became even more extreme after the longer incubation.

Although the degradation patterns were nearly identical for the native and recombinant forms, a closer look at the proportion of products and the effects of temperature showed that there were differences between all of these forms. The native form exhibited a clear difference in activity level, with an increase in the amount of degradation products and a decrease in insulin substrate occurring when the incubation temperature was decreased from 90ºC to 80ºC. Effects of temperature were tested for the $\alpha + \beta 2$ and $\alpha + \beta 1 + \beta 2$ forms

that were assembled at 90ºC and 98ºC. In contrast to the native proteasome, these recombinant forms did not exhibit dramatic differences in activity levels when the incubation temperature was altered. Instead, the activities of the recombinant forms either did not change or dropped when moving from 90ºC to 80ºC, similar to the NHS proteasome (see Table 6.5b). Variations also occurred among the recombinant forms. In particular, the $\alpha + \beta 2$ forms that were assembled at 98ºC and 105ºC (the two higher assembly temperatures) exhibited extremely low activity compared to the other proteasome forms (both native and recombinant), even though the same mass of enzyme previously had shown equivalent activity on the MCA-linked peptides. In addition, there were subtle, but consistent, differences between the amounts of certain degradation products that were formed by the $\alpha + \beta 2$ and $\alpha + \beta 1 + \beta 2$ forms, as shown in part c) of Fig. 6.6. Although the same peaks appeared for each assay, the levels of specific products were comparatively different between the recombinant proteasome forms. Furthermore, these two samples had previously been shown to have equivalent activities on the peptide substrates. In the case of insulin, however, the proteasome containing $\beta 1$ was more efficient in its degradation (notice the presence of less insulin in the peak for this form).

The expression product for the P. furiosus "PAN" gene in E. coli was a larger assembly based on a single subunit. The "ATP-dependent 26S protease regulatory subunit" protein product from the PF0115 gene in *P. furiosus* was termed "PAN" (Zwickl *et al.*, 1999a) because of its sequence similarity to other archaeal PANs and putative PANs, such as that from *M. jannaschii* (Benaroudj & Goldberg, 2000; Benaroudj *et al.*, 2003; Navon & Goldberg, 2001; Wilson *et al.*, 2000; Zwickl *et al.*, 1999a), which is, thus far, the only PAN-

like ATPase from the archaea to be characterized. Based on amino acid sequence data (ExPASy Compute pI/MW tool; http://us.expasy.org/tools/pi_tool.html), the hypothetical molecular weight for the *P. furiosus* PAN was 44,805 Da. After this gene product was expressed in *E. coli* and purified, it ran near the expected size, at approximately 43 kDa, on SDS-PAGE (Fig. 6.1b). However, on 10% native PAGE, the recombinant *P. furiosus* PAN did not migrate in the separating gel (data not shown). Although the single PAN gene product is ~45 kDa in size, it has been found in studies of the *M. jannaschii* version that the recombinant form assembles into much larger structures of 550 kDa (Wilson *et al.*, 2000) to 650 kDa (Benaroudj & Goldberg, 2000; Zwickl *et al.*, 1999a), containing ~12 subunits each (Wilson *et al.*, 2000). The native *M. jannaschii* form was also purified and found to be 550 kDa, however, it was not determined if this complex contained only the single PAN subunit in a homo-oligomeric structure or if there were other associating proteins (Wilson *et al.*, 2000). The assembled *M. jannaschii* PAN was also found to have an irregular ring structure (comma-shaped in transmission electron micrographs) with a diameter near the range of the 20S proteasome (Wilson *et al.*, 2000). From these studies on the *M. jannaschii* PAN, the inability of the *P. furiosus* PAN to migrate in a native gel was most likely due to the assembly of its subunits into a similarly large structure. Therefore, all subsequent studies were done assuming an approximate size of 550 kDa for PAN.

The recombinant P. furiosus PAN exhibited activity on several nucleotidases and this activity was stimulated by the presence of a polypeptide substrate. The nucleotidase activity by PAN on ATP, ADP, AMP-PNP, CTP, GTP, and UTP was tested and hydrolysis rates were reported in Table 6.6. Similar to reports on the *M. jannaschii* PAN, the *P. furiosus*

PAN was most active on ATP and CTP, less active on GTP and UTP, and not active on ADP and AMP-PNP (Wilson *et al.*, 2000; Zwickl *et al.*, 1999a), with rates of ATP hydrolysis similar to several reported values (Benaroudj *et al.*, 2003; Zwickl *et al.*, 1999a). There was a discrepancy between Zwickl et al. (1999a) and Wilson et al. (2000) as to which nucleotide PAN hydrolyzed most efficiently, with the former report claiming highest activity on CTP and the latter reporting highest activity on ATP. The observed activity of the *P. furiosus* PAN was found to agree with Wilson et al. (2000), with the highest hydrolysis rate on ATP rather than CTP.

It was found by Benaroudj et al. (2003) in their study of the *M. jannaschii* PAN that higher hydrolysis rates occurred on ATP in the presence of various types of polypeptide substrates such as β-casein, which contains little secondary or tertiary structure (Creamer *et al*., 1981), and green fluorescent protein ssrA (GFP with 11 residue extension at its carboxy-terminus), which is globular (Benaroudj & Goldberg, 2000). When 1,000-fold molar excess of β-casein over PAN (as was found to be optimum for stimulation (Benaroudj *et al.*, 2003)) was added to the *P. furiosus* PAN ATPase assay, the hydrolysis of ATP increased by 3-fold, similar to the 5-fold increase found with the *M. jannaschii* version (Benaroudj *et al.*, 2003). In addition, the *P. furiosus* PAN exhibited 2- to 4-fold increases in hydrolysis rates on CTP, GTP, and UTP, upon addition of β-casein (Table 6.6). The recombinant *P. furiosus* PAN was also analyzed on the DSC and its melting point in solution (in the absence of proteasome, nucleotides, or protein substrates) was found to be 94.2ºC in two separate scans (see Fig. 6.2c). This was quite low in comparison to the proteasome and two of its subunits, which all had melting temperatures in excess of 100ºC.

The P. furiosus PAN does not affect proteasome assembly, stablization, or peptidase activity. In the study of the *P. furiosus* PAN, the peptidase activities of all forms of the proteasome (native, NHS, and recombinant) were not stimulated by the presence of PAN and ATP. This was found after PAN and 1 mM ATP was combined with all forms of the proteasome in a 4:1 molar ratio of PAN over the proteasome (found to be ideal by Zwickl et al. (1999a) for polypeptide stimulation) during hydrolysis of VKM-MCA (see Table 6.4). All forms of the proteasome, including $\alpha + \beta1$, $\alpha + \beta2$, and $\alpha + \beta1 + \beta2$ assembled at 90ºC, 98ºC, and 105ºC were also checked on VKM-MCA and it was found that there was no difference in activity after the addition of PAN and ATP. These results correspond to the earlier findings on the *M. jannaschii* PAN, which also did not stimulate proteasome activity on small peptides (Zwickl *et al.*, 1999a).

All of the recombinant *P. furiosus* proteasome forms were also assembled in the presence of PAN, with 1 mM ATP and a 4:1 molar ratio of PAN over proteasome added to the subunit mixtures during assembly. The presence of PAN with ATP did not have any apparent affect on the efficiency of assembly or on the characteristics of the resulting proteasome forms when they were tested for stability and activity on both small peptides (MCA-linked) and the polypeptide insulin B. Specifically, all subunit combinations that previously exhibited some level of peptidase activity ($\alpha + \beta1$, $\alpha + \beta2$, and $\alpha + \beta1 + \beta2$) were assembled as usual with PAN and ATP present, precipitates were spun out, and a fixed mass of enzyme in each the remaining mixtures was tested for activity on VKM-MCA (peptidase activity) and insulin B (polypeptide activity). In the assays, these samples were tested as is, with ATP added (1 mM), and with PAN (4:1 molar ratio over proteasome) and ATP added and directly compared to the assemblies that were created without PAN present. In addition,

the same stability study that was done with the assemblies lacking PAN (see "*Comparing stability among the recombinant P. furiosus proteasome forms*") was performed on these recombinant assemblies with PAN to check for any stabilizing ability by PAN. There was no apparent affect by PAN to further stabilize the proteasomes. Furthermore, the presence of PAN (approximate 4:1 molar ratio PAN:proteasome) and ATP (1 mM) together with the NHS form during the 115ºC stability studies had no affect on its stability, as shown in Figure 6.5b.

The P. furiosus PAN had varying effects on the activity of each native P. furiosus proteasome on the polypeptide substrate insulin B. The recombinant *P. furiosus* PAN, ATP, and the combination of both were added in varying amounts and at different incubation temperatures to the proteolytic assays for degradation of insulin B by the *P. furiosus* proteasome forms. In all cases, the addition of only ATP or only PAN had no affect on the hydrolysis of insulin B by any of the proteasome forms (see Tables 6.5a-b, Table 6.5a shows native forms, 6.5b shows recombinant forms). However, each form, including the native, NHS, and various recombinant forms exhibited a different reaction to the presence of PAN and ATP in combination, and, therefore, to the ATPase activity of PAN. It was reported earlier that, compared to the other proteasome forms, the NHS proteasome forms exhibited different degradation patterns, as well as opposing changes in the level of activity when the assay temperature was changed. In addition to these differences, each native proteasome reacted to the presence of PAN/ATP in an opposite manner. As shown in Table 6.5a, the native proteasome had increased activity on insulin B as the temperature was decreased from 90ºC to 80ºC and with the presence of PAN/ATP. There was no apparent correlation

between activity level and amount of PAN added, nor did the presence of PAN/ATP have a dramatic affect at the higher assay temperature. This was most likely because insulin B was in a more denatured state at 90ºC and, therefore, unfoldase activity by PAN was not particularly effective. Figure 6.7 illustrates some examples of the differences caused by the presence of PAN/ATP. In part a), the level of insulin B remaining in the assay steadily dropped as levels of degradation products increased upon alteration of the assay conditions from 90ºC (green) to 80ºC (pink) to 80ºC with added PAN/ATP. Part b) shows an example of a slightly longer incubation time and clearly illustrates the increase in degradation products when PAN/ATP were added to the assay. The opposite effects were seen with the NHS proteasome, as shown in Table 6.5a. In addition to the decreased activity when temperature was changed from 90ºC to 80ºC, as the amount of PAN/ATP was increased (particularly from 1:1 to 1:10 ratio of proteasome:PAN), the activity of the proteasome decreased. Figure 6.8 shows an example of several of the chromatograms resulting from the NHS activity on insulin B. In part a), the inhibitory effect of increased levels of PAN on NHS activity at 80ºC is shown, with a close correlation between the ratio of PAN and the remaining amounts of intact insulin in the assay. Part b) illustrates the same trend at the 90ºC assay temperature, with a dramatic drop in NHS activity upon addition of PAN/ATP.

Each of the various active recombinant forms of the proteasome ($\alpha + \beta 2$ and $\alpha + \beta 1 + \beta 2$ assembled at the three different tested temperatures) was affected in a different way by the presence of PAN and ATP. These effects are listed in Table 6.5b and illustrated in Figure 6.9. The proteasomes assembled at 90ºC both gave a similar change in activity upon addition of PAN/ATP as compared to the native proteasome. However, the level of activity of these forms were relatively low without the PAN/ATP at both assay conditions,

with assay temperature having no noticeable affect on activity. Once PAN/ATP was added, particularly at the lower assay temperature of 80ºC, activity was greatly stimulated by both 90ºC assemblies (Fig. 6.9a). However, the amount of PAN had no affect on activity. The proteasomes assembled at 98ºC both exhibited relatively low activity, with minimal temperature effects and greater stimulation by PAN/ATP at the 80ºC assay temperature (Fig. 6.9b) as compared to the 90ºC assemblies. Again, the amount of PAN added was inconsequential. Interestingly, the $\alpha + \beta 2$ assembled at 105ºC was relatively inactive at 90ºC, and activity dropped upon addition of PAN/ATP. This same structure was reported as very active on the peptide substrate VKM-MCA as compared to the other proteasomes with the same relative amount of each proteasome used in both the peptidase and insulin B hydrolysis experiments. Furthermore, the $\alpha + \beta 1 + \beta 2$ assembled at 105ºC exhibited decreased activity upon addition of PAN/ATP at the 90ºC assay temperature, similar to the NHS form of the proteasome. This is clearly visible in the resulting chromatogram shown in Fig. 6.9c. In general, there were consistent differences in degradation of insulin and the effects by PAN/ATP, with no two recombinant assemblies showing identical trends. Furthermore, the two native forms had hydrolysis characteristics that exhibited opposite trends compared to each other, as well as differences when compared to the recombinant proteasome forms.

V. DISCUSSION

The native and recombinant P. furiosus α subunit assembles into larger structures and acts as a matrix for assembly of the proteasome. In versions of the proteasome that have been biochemically characterized, including the *T. acidophilum* form (Zwickl *et al.*, 1994), *M. thermophila* form (Maupin-Furlow *et al.*, 1998; Wilson *et al.*, 2000), α5 from *Trypanosoma brucei* (Yao *et al*., 1999), and the human HsC8 α-type subunit (Gerards *et al*., 1998a; Gerards *et al*., 1997), it was found that the α subunit spontaneously self-assembles into rings. It was also thought to form a scaffolding for the assembly of the β subunits, which cannot assemble into full rings without the presence of the α subunit. However, in other cases where multiple forms of the α subunit are incorporated into the proteasome, such as with the *R. erythropolis* form, the α subunits were found to depend on the presence of β subunits for proper assembly (Zuhl *et al.*, 1997a). This discrepancy among the different proteasomes relates to their assembly pathways, which depend on the source organism. For example, several archaeal forms and the bacterial *R. erythropolis* form exhibited clear differences between assembly pathways, based on experimental evidence from the individual characteristics of each subunit. In particular, the archaeal *T. acidophilum* (Seemuller *et al.*, 1996; Zwickl *et al.*, 1994), *M. thermophila* (Maupin-Furlow *et al.*, 1998), and *M. jannaschii* (Wilson *et al.*, 2000) forms (all containing one version of α and one version of β) were found to have an assembly order that started with fully assembled α_7 rings. These rings then provided the matrix onto which folding and processing of the β subunits occurred. This was the most likely scenario since, in all three archaeal proteasomes, the β subunits were unable

to process the pro-sequence or produce full β_7 rings without the presence of the α subunit. The *R. erythropolis* proteasome, which contains two different α-type and two different β-type subunits (Tamura *et al.*, 1995), contains α subunits that cannot form the 7-membered rings without the presence of β and vice-versa (Lupas *et al*., 1997b). From this knowledge, it was hypothesized that α/β heterodimers are initially created and quickly form into half-proteasomes (two 7-membered rings), which then associate with each other to create the full structure (Mayr *et al*., 1998; Zuhl *et al.*, 1997a; Zuhl *et al.*, 1997b).

In the case of the *P. furiosus* proteasome, it was found that the α subunit was able to assemble into a larger structure that was stable enough to co-purify with the proteasome during purification of the native form (isolated α subunit was found in the native sample based on MALDI-TOF analysis). Both the native and recombinant forms of this subunit appeared large enough that they could not migrate in a 10% native gel. The recombinant version was stable enough that it could withstand incubation for 1 hr at 105ºC and did not melt at or below a temperature of 120ºC on the DSC. These observations indicate that the *P. furiosus* version of the proteasome contains subunits that behave much more similarly to those of the other archaea, such as *T. acidophilum*, *M. thermophila*, and *M. jannaschii*, than to those of the *R. erythropolis*, even though the other archaea contain only one α- and one β-type subunit. In addition, it was found that the β subunits could not form larger structures, as seen in the native gels of the individual subunits and the β1 + β2 combinations (Fig. 6.4b), and could not process their own pro-sequences, as seen from the N-terminal sequencing results of the recombinant forms. This evidence indicates that the overall assembly pathway of the *P. furiosus* proteasome is most likely similar to that of the other archaeal proteasomes.

Specifically, the α subunit spontaneously forms the 7-membered rings, which are quite stable, and these rings become the matrix onto which the β subunits assemble to form the proteasome. This likely assembly pattern is consistent with microarray data, which indicated a decrease in α levels and a simultaneous increase in both β levels upon heat shock of the *P. furiosus* cells (Table 6.2). One possible reason for the opposing changes in subunit expression could be that stable α_7 rings are already present in excess within the cell under normal conditions. Upon heat shock, the levels of β could rise, with expressed β subunits immediately associating with the α rings, increasing the level of fully assembled proteasome and, therefore, the degradation of polypeptides and proteins that were denatured as a result of the stress condition.

The P. furiosus β1 subunit is not an active copy of the β2 subunit. The α + β2 (1:1 molar ratio) combination was relatively active compared to the native proteasome, particularly when comparing the forms assembled at the lower temperatures of 90ºC and 98ºC and when looking at activity on both peptides and polypeptides. Since α + β2 did not contain β1, yet the native form was found to contain β1, the β1 that was part of the native structure did not appear to affect the catalytic ability of the proteasome. The same level of activity occurred whether β1 was present in the structure or not. To further support this finding, it was also observed that the α + β1 (1:1 molar ratio) combination was not appreciably active, even after the high-temperature incubation steps. Though it was quite stable, exhibiting a melting temperature of 104.4ºC, the activity of β1 was consistently low on all of the tested peptides, including chymotrypsin, trypsin, PGPH, SNAAP, 20S, and

Amy, and was nonexistent on the tested polypeptides. Presence of detectable, though low, activity by $\alpha + \beta1$ suggested that some interaction was occurring between α and $\beta1$. The absence of fully formed proteasome in these gels suggested that the residual $\alpha + \beta1$ peptidase activity must have occurred as a result of a relatively low amount of active proteasome being created during assembly, rather than from a mixture of active proteasome containing a high level of partially active or inactive "proteasome-like" structures. Clearly the $\beta1$ did not assemble and incorporate into a proteasome with α in the $\alpha + \beta1$ combinations. Furthermore, in the high temperature incubation of $\alpha + \beta1$, as with $\beta1 + \beta2$ at 105ºC, $\beta1$ was not visible on the native gel, even though it was stable enough to remain visible under the same conditions in the absence of α. Whether it was destabilized by negative interactions with α or is simply not seen because it was partially bound to the α rings remains to be determined. The $\beta1$ was initially expected to exhibit activity, since it contains the characteristic active-site "TTT" seen in all forms of the active proteasome β subunits. It was found that another archaeon, *A. pernix* K1, encodes both inactive and active forms of the β subunit. Specifically, the gene encoding APE0507 does not contain the active-site "TTT" region (Kawarabayasi *et al.*, 1999). In the archaeon *H. volcanii*, which contains two α subunits and one β subunit, the roles of the two α subunits appeared to be different, though still unclear, with separate proteasome structures being assembled with varying stoichiometric ratios of subunits per structure (Kaczowka & Maupin-Furlow, 2003; Wilson *et al.*, 1999). Furthermore, in many eukaryotic forms of the proteasome, which contain 7 α-type and 7 β-type subunits, with yeast forms encoding 7 β-type (Heinemeyer *et al.*, 1994) and higher eukaryotic forms encoding as many as 10 β-type genes (Baumeister *et al.*, 1997;

Coux *et al*., 1996; Fu *et al.*, 1998), each of the subunits plays a specific and distinct role within the structure. These roles include presence or absence of proteolytic activity, as well as different types of proteolytic activity. For example, Groll et al. (1997) found that the yeast proteasome β1, β2, and β5 subunits catalyzed caspase-, trypsin-, and chymotrypsin-like activities, respectively. This was similar to other studies suggesting that β-type subunits create pairs of active and inactive structures to catalyze the different peptidase activities (Chen & Hochstrasser, 1996; Dick *et al*., 1998; Enenkel *et al*., 1994; Gueckel *et al*., 1998; Heinemeyer *et al*., 1991; Hilt *et al.*, 1993). Furthermore, identical active sites within a proteasome may exhibit different types of activity with varying kinetic characteristics, which appeared to be controlled by subunit interactions (Arribas & Castano, 1990; Djaballah & Rivett, 1992; Maupin-Furlow *et al.*, 2001; Orlowski *et al*., 1991). From the relatively low activity by the *P. furiosus* β1 seen in this study and the clear evidence showing the importance of various roles played by different versions of the subunits, it is clear that the β1 subunit is not a primary catalytic center in the *P. furiosus* proteasome, while β2 is able to exhibit full activity under favorable conditions.

The P. furiosus β2 is the main catalytic center during non-stress conditions in the cell, while β1 may play more of a stabilizing role at other times, with different ratios of each β incorporated into the structure based on the environmental conditions. When the mixtures of $\alpha + \beta1 + \beta2$ were incubated at 90ºC, 98ºC, and 105ºC, the gradual disappearance of isolated β1 (β1 not incorporated into the full proteasome) with increase of assembly temperature was clear from both the native gel, where the β1 was not visible at the 105ºC

case (Fig. 6.4b), and the DSC results, where the melting curve for β1 was prominent in the 90ºC case and undetectable in the 105ºC case (Fig. 6.2b). When the three subunits were assembled at 98ºC (unpublished data), an intermediate amount of β1, compared to the same mixtures assembled at 90ºC and 105ºC, was seen on a native gel, as well. In other words, the lower the assembly temperature, the more β1 remained unincorporated from the full proteasome. These results lead to two possibilities, the first being that β1 may have destabilized at the higher temperature and simply fallen out of solution, and the second being that increasingly more of β1 may have been incorporated into the proteasome when the assembly temperature was increased from 90ºC to 98ºC and, finally, to 105ºC. When examining the gel and DSC data alone, it is difficult to determine which of the possibilities is more likely. However, there are several pieces of evidence that indicate important differences between the various assemblies. One difference is the specific activities on VKM-MCA of all of the tested proteasomes. The version lacking β1 ($\alpha + \beta 2$) had increased specific activity as assembly temperature was increased, while the version containing β1 acted in an opposite fashion (Table 6.4). An even stronger piece of evidence is the altered stability, as seen with the high-temperature (115ºC) incubation studies, of the $\alpha + \beta 1 + \beta 2$ structure after assembly at 90ºC versus 105ºC (Fig. 6.5a). The stability of the 90ºC $\alpha + \beta 1 + \beta 2$ is not only different from the 105ºC $\alpha + \beta 1 + \beta 2$, but nearly the same as both structures that lack β1. From the figure and calculated decay constants (Fig. 6.5a), it is clear that the $\alpha + \beta 1 + \beta 2$ structure assembled at 105ºC was far more stable than the same mixture assembled at 90ºC and the other mixtures lacking β1. The most likely conclusion that can be drawn from this is that more β1 is incorporated into the higher temperature assembly and that

it plays a stabilizing role. Therefore, β1 is, in fact, not a near copy of β2, but plays a distinctly different role that may include stabilization at stress conditions. The composition of the proteasome, particularly the ratio of β1:β2 may be specifically tailored to the environment of the enzyme, with more β1 incorporated into the structure when greater stability is necessary for survival of the proteasome and the *P. furiosus* cell.

The P. furiosus PAN is unstable in an isolated state and is unable to stabilize the proteasome. It was found that, during heat shock of *P. furiosus* cells, the level of PAN increased by 7- to 15-fold within the first 5-10 min, suggesting that it may be an important player in the stress response pathway. However, in an isolated state, PAN was relatively unstable, exhibiting a melting temperature of 94.2ºC. With this knowledge, it seems likely that the PAN may not exist as an isolated structure for long periods in the cell, but may instead be continuously associated with nucleotides, substrate proteins that must be targeted for hydrolysis by the proteasome, or other stabilizing factors including the proteasome itself. Interestingly, the presence of a stable 26S proteasome structure has not been determined with confidence in the archaea. Wilson et al. (2000) reported the presence of the *M. jannaschii* 20S proteasome singly or doubly capped by PAN as seen in transmission electron micrographs. However, the concentrations of such complexes were extremely low and, as a result, it was suggested that the interaction between the proteasome and PAN is transitory or that this interaction required further stabilizing factors that remained unknown. With this evidence, it may be possible that PAN is up-regulated at such high levels during heat shock simply because of its instability; those particles that are not able to associate with a

stabilizing factor immediately after expression cannot survive in the cell during a stress event.

There was no apparent affect by PAN to further stabilize the proteasomes, either when it was added together with the proteasome subunits during assembly or during high temperature (115ºC) incubation of native and pre-assembled proteasomes. In addition, PAN had no affect on the ability of the proteasome to hydrolyze small peptides, particularly the MCA-linked flourogenic peptides tested in this study. These results verify the accepted beliefs about the *M. jannaschii* PAN, including the presence of a different type of α subunit "gate" compared to the eukaryotic proteasomes, which allows entry of small peptide substrates even in the absence of the PAN ATPase (Navon & Goldberg, 2001; Zwickl *et al.*, 1999a). These results also show that, although PAN has been found to associate with the proteasome, both through direct viewing on transmission electron micrographs (Wilson *et al.*, 2000) and through the indirect observation via the stimulation of polypeptide hydrolysis (Benaroudj & Goldberg, 2000; Wilson *et al.*, 2000; Zwickl *et al.*, 1999a), PAN does not appear to have an affect on the assembly efficiency or stability of the proteasome.

The varying reactions of the P. furiosus proteasomes to the presence of PAN further support the hypothesis that the proteasome exists in different forms. Akopian et al. (1997) and Kisselev et al. (1998) found that the *T. acidophilum* form of the proteasome degraded the polypeptide substrate β-casein in a processive manner, quite different from the majority of proteolytic enzymes. Specifically, it was found that each polypeptide substrate (including insulin-like growth factor (Kisselev *et al.*, 1998), lactalbumin (Kisselev *et al.*, 1998), β-casein (Akopian *et al.*, 1997; Kisselev *et al.*, 1998), and alkaline phosphatase (Kisselev *et al.*,

1998)) was hydrolyzed to oligopeptides before being released by the proteasome, with a new particle attacked for hydrolysis. Instead of just making one cut and then releasing the polypeptide, the proteasome consistently made a series of cuts in the polypeptide before release. Once the full polypeptide substrate molecules were depleted, the proteasome was also able to further degrade the small peptides. However, it was hypothesized that the primary role of the archaeal proteasome is most likely that of polypeptide degradation, with the need for other functioning peptidases to complete the cycle of protein degradation to single residues (Kisselev *et al.*, 1998). This hypothesis was derived, not only from the experimentation done with the *T. acidophilum* proteasome, which directly proved its processive nature (Akopian *et al.*, 1997; Kisselev *et al.*, 1998), but also from the extensive information about the eukaryotic form. The eukaryotic proteasome, along with the ubiquitin system (the "tagging" system for selective degradation of proteins), has been reviewed multiple times and all reviews stress the role of the proteasome as a critical protease in the bulk turnover of proteins and targeted hydrolysis of proteins that enter into the ubiquitin-mediated pathway (Baumeister *et al.*, 1997; Baumeister *et al.*, 1998; Coux *et al.*, 1996; Lupas *et al*., 1995; Maupin-Furlow *et al.*, 2001; Maupin-Furlow *et al*., 2000; Tanaka, 1998; Voges *et al.*, 1999; Zwickl *et al*., 1999b). Therefore, it was assumed that the primary activity of the proteasome from *P. furiosus* was in the polypeptide degradation, rather than the peptidase activity.

The archaeal *M. jannaschii* PAN has been well characterized, with all studies in agreement that it is able to stimulate polypeptide degradation (but not small peptide degradation) by the archaeal proteasome (Benaroudj & Goldberg, 2000; Benaroudj *et al.*, 2003; Navon & Goldberg, 2001; Wilson *et al.*, 2000; Zwickl *et al.*, 1999a). In the study of

the *P. furiosus* native proteasome with PAN, there is agreement with the findings of the *M. jannaschii* versions. Specifically, PAN stimulates the degradation of insulin B by the native proteasome at two different assay temperatures. However, during the investigation of the NHS version of the proteasome, the opposite effect was observed, with PAN dramatically inhibiting the proteasome activity. More specifically, the ATPase activity of the PAN exhibited this negative effect, since the presence of PAN alone did not have any affect. These effects are illustrated in Fig. 6.8 and Table 6.5a. From this result, several possibilities arise, one of which is the chance that the NHS form already has PAN and/or a similar enzyme associated with it and is in a form similar to the 26S, so that the addition of more PAN brings the level of control enzyme so high that it becomes inhibitory. It was found that the optimum ratio of PAN to proteasome was 4:1 in studies of the *M. jannaschii* version, with a leveling of activity occurring as this ratio became larger (Zwickl *et al.*, 1999a). The fact that the levels of PAN dramatically increase in the initial minutes of the heat shock event in the *P. furiosus* cells and fall again as the high temperature period reaches 1 hr, coupled with the finding that PAN, alone, has such a low melting point (<95ºC) that it cannot possibly remain in an isolated state during a heat shock event, leads to the conclusion that PAN must associate with another stabilizing factor immediately after it is expressed in the cell (particularly during a stress response). Because of the proven interactions that are known to occur between the PAN and proteasome (Wilson *et al.*, 2000), one of the most likely choices for the stabilizing factor is the proteasome, which, by itself, is able to remain stable well above 105ºC. However, in this study of the NHS proteasome, it was also found that the addition of ATP alone had no affect on the activity. This argument can also be disputed by the possibility that there is already ATP associated with the 26S structures of

proteasome + PAN in the NHS sample. It has not yet been determined whether or not this NHS form is an archaeal version of the full 26S eukaryotic form. If it is, this would be the first reported finding of a stable archaeal 26S proteasome.

Along with differences between the native and NHS proteasomes, the recombinant forms also exhibited variations in activity upon exposure to PAN ATPase activity. All of the 90ºC and 98ºC assemblies, in addition to exhibiting similar degradation patterns to the native proteasome (though there were peak variations among the assemblies), were also stimulated by PAN in the same way (Table 6.5b). However, the $\alpha + \beta 2$ and $\alpha + \beta 1 + \beta 2$ mixtures assembled at 105ºC, were noticeably inhibited by the ATPase activity of PAN, as was seen with the NHS form. Given this observation, coupled with the individual peak patterns of degradation, it is clear that the high temperature assemblies are not the same as the lower temperature ones (PAN does not affect them in the same way), but are also not the same as the NHS form (degradation patterns are different). Instead, they may be an intermediate form between the native and NHS proteasomes. The most obvious choice, given the data on its assembly characteristics, is that this form must have a higher percentage of β1 compared to the lower temperature cases (recall that the 90ºC assembly clearly showed that little of β1 had been incorporated into the proteasome structure, yet was much less stable than the 105ºC assembly). It is possible that one of the important roles of β1 is stabilization of the structure during stress response and that, in carrying out this role, β1 is incorporated into the proteasome at a higher level during times of higher stress. The resulting alteration in β1:β2 ratio per proteasome molecule likely has an effect on the biochemical characteristics of the enzyme, such as its association with PAN. Though there are most likely several separate phenomena occurring. One effect may be the presence of more β1 at times of high stress to

stabilize the structure – the need for stabilization in this period is priority over basic catalysis, which is the apparent role of β2, an unstable, but very active subunit under favorable conditions. Another, perhaps separate occurrence may be that the α "gate", which is normally in a semi-closed state (allowing small peptides, but not full proteins with secondary and tertiary structure to enter the catalytic sites), remains open during stress response so that proteins are fed through the catalytic region at a faster pace, without the control of the ATPase activity, but perhaps under control of a different associating enzyme (unfoldase or anti-chaperone). From the clearly negative effect of PAN on the high temperature recombinant assemblies and the NHS proteasome, coupled with the positive affect on the native and other recombinant proteasomes, it must be concluded that PAN does not associate with the "stress-response" form of the proteasome, but acts as a controlling enzyme under normal conditions in the cell. The levels of PAN do increase dramatically at the initial stages of stress response, meaning that PAN could very likely play a role during this period. However, a possible sequence of events could be that PAN is initially highly expressed during the first minutes of the stress event. This PAN may associate immediately with proteasome forms already present, which were previously assembled during normal conditions. These proteasomes would be the native forms, containing the lower percentage of β1 and the α rings in a conformation similar to the "closed gate" state. As the period of stress continues, the larger amounts of the NHS proteasome are assembled, which are more stable structures containing a higher percentage of β1 and the α rings in an "open gate" state. As a higher percentage of NHS proteasome is created, less of the PAN is expressed, as seen from microarray data, where after longer periods of stress response, the PAN is up-regulated at a much lower level compared to the initial stages of stress. Therefore, by the end of stress

response, the cell may contain a mixture of 26S proteasomes, which could be the native, low-β1 form, combined with PAN, and the NHS form, which could be a high-β1 form with an open gate for activity that is not under control of the PAN ATPase activity, but likely under control of a different enzyme, since it exhibited a different degradation pattern from both the native and the recombinant proteasomes assembled even at the higher temperatures.

VI. CONCLUSIONS

As with the other archaeal proteasomes characterized to date (Maupin-Furlow *et al.*, 1998; Seemuller *et al.*, 1996; Wilson *et al.*, 2000; Zwickl *et al.*, 1994), the native and recombinant *P. furiosus* α subunit appears to assemble into larger (α_7 or higher) structures that act as matrices for assembly of the proteasome. The *P. furiosus* proteasome differs from the other characterized archaeal proteasomes (*T. acidophilum, M. thermophila, M. jannaschii*) in that it contains more than one copy of the β subunit. However, its assembly pattern may still be more like these archaeal forms rather than the bacterial form *R. erythropolis*, which contains 2 copies each of α and β, but is unable to create full α_7 ring structures spontaneously. This bacterial proteasome was found to require the presence of the β subunits in order for any organized assembly of the α_7 rings to occur (Zuhl *et al.*, 1997a). Since the *P. furiosus* α subunits appeared in stable form in both the native proteasome sample (as found from sequencing) and was quite stable when expressed (as seen with SDS-PAGE and DSC), it most likely is able to spontaneously form stable α_7 rings without the presence of the β subunits. This finding can be coupled with the observation that the β subunits of *P. furiosus* do not process their own pro-sequences or appear to form large assemblies until combined with α. It can be deduced, then, that the spontaneously formed α_7 rings most likely act as the scaffolding for β processing and assembly.

In all of the archaeal proteasomes characterized to date, there is only one known form of each subunit, as determined from sequence information. However, the *P. furiosus* proteasome, along with several other of the archaeal proteasomes listed in Table 6.1, contains

a second β subunit. As compared to some of the eukaryotic subunits, which can be 90% identical (Fu *et al.*, 1999), the two β subunits of *P. furiosus* are relatively different, with only 48% of the amino acids identical. Furthermore, when the two β subunits were directly compared in terms of stability (SDS-PAGE, DSC scans, activity decay experiments including long incubation at 115ºC) and activity (MCA-linked peptides and polypeptides), there were clear differences. The β2 was much less stable, but much more active than β1 on both peptides and polypeptides. However, full proteasome structures were more stable when they included β1. Overall, it can be concluded that the β1 subunit is not an active copy of the β2 subunit, but instead plays a distinctly different role in the proteasome structure.

Once it was determined that the two β subunits of *P. furiosus* were not redundant copies of each other, the individual roles of each subunit still remained unknown. The $\alpha + \beta 2$ proteasome recombinant assembly appeared to exhibit similar levels of activity compared to the native form, which was found to contain all three subunits. Therefore, the presence or absence of β1 in the structure did not appear to affect the catalytic ability of the proteasome. However, in structures where β1 was not included (the $\alpha + \beta 2$ assemblies) or less of β1 appeared to be incorporated (the $\alpha + \beta 1 + \beta 2$ assembled at 90ºC), the stability of the structure was much lower than that of the proteasomes that contained a higher amount of β1. Therefore, it appears that the *P. furiosus* β2 is the main catalytic center during non-stress conditions in the cell, while β1 may play more of a stabilizing role at other times, with different ratios of each β incorporated into the structure based on the environmental conditions.

The archaeal *M. jannaschii* PAN has been characterized in several separate studies and was found to stimulate the proteasome on polypeptides (but not small peptides) in the presence of ATP and CTP (Benaroudj & Goldberg, 2000; Wilson *et al.*, 2000; Zwickl *et al.*, 1999a). There are no additional characterization studies published on other archaeal PAN structures. In this study, the extent of interaction between the *P. furiosus* PAN and proteasome was examined. Specifically, the effects of PAN on assembly of the individual proteasome subunits and the stability of the structure at high temperature (115ºC) were tested. The *P. furiosus* PAN was found to be not only unstable in an isolated state (as seen in DSC scans), but was also unable to stabilize the proteasome during and after assembly. Overall, it did not appear to interact with the proteasome during its assembly pathway or in conditions outside of polypeptide degradation (particularly with stability and small peptide hydrolysis).

Although the *P. furiosus* PAN did not appear to interact with the proteasome during assembly and small peptide studies, it did have an effect on the ability of the proteasome to degrade polypeptides. Furthermore, the varying reactions of the *P. furiosus* proteasomes to the presence of PAN further support the hypothesis that the proteasome exists in different forms. In particular, under normal conditions, the proteasome is likely in a form containing primarily β2, with the α subunits in the "closed gate" state, so that degradation of polypeptides is strictly regulated by PAN (and perhaps other energy-dependent proteins). During stress events, a high level of PAN may be expressed to accelerate the degradation activity of the proteasomes already existing in the cell (the form with a high level of β2). At the same time, the NHS proteasome forms may be increasing in population over the period of the stress event. This NHS form may contain a higher number of β1, stabilizing (but not

catalytically active) subunits. As well, it may be in more of an "open-gate" conformation, with the α N-termini no longer blocking off the catalytic sites in the same manner. In this case, regulation by PAN is not necessary, although some other type of regulation (by a separate, unknown enzyme) may be occurring. This was primarily assumed from the observation that the ATPase activity of PAN had a detrimental effect on the proteolytic ability of the NHS proteasome and the high-temperature recombinant proteasomes (those assembled at 105ºC), particularly with degradation of polypeptides.

VII. REFERENCES

Akopian, T. N., Kisselev, A. F. & Goldberg, A. L. (1997). Processive degradation of proteins and other catalytic properties of the proteasome from Thermoplasma acidophilum. *J Biol Chem* **272**(3), 1791-8.

Arendt, C. S. & Hochstrasser, M. (1999). Eukaryotic 20S proteasome catalytic subunit propeptides prevent active site inactivation by N-terminal acetylation and promote particle assembly. *Embo J* **18**(13), 3575-85.

Arribas, J. & Castano, J. G. (1990). Kinetic studies of the differential effect of detergents on the peptidase activities of the multicatalytic proteinase from rat liver. *J Biol Chem* **265**(23), 13969-73.

Baker, D., Shiau, A. K. & Agard, D. A. (1993). The role of pro regions in protein folding. *Curr Opin Cell Biol* **5**(6), 966-70.

Bauer, M. W., Bauer, S. H. & Kelly, R. M. (1997). Purification and characterization of a proteasome from the hyperthermophilic archaeon Pyrococcus furiosus. *Appl Environ Microbiol* **63**(3), 1160-1164.

Baumeister, W., Cejka, Z., Kania, M. & Seemuller, E. (1997). The proteasome: a macromolecular assembly designed to confine proteolysis to a nanocompartment. *Biol Chem* **378**(3-4), 121-30.

Baumeister, W., Walz, J., Zuhl, F. & Seemuller, E. (1998). The proteasome: paradigm of a self-compartmentalizing protease. *Cell* **92**(3), 367-80.

Benaroudj, N. & Goldberg, A. L. (2000). PAN, the proteasome-activating nucleotidase from

archaebacteria, is a protein-unfolding molecular chaperone. *Nat Cell Biol* **2**(11), 833-9.

Benaroudj, N., Zwickl, P., Seemuller, E., Baumeister, W. & Goldberg, A. L. (2003). ATP hydrolysis by the proteasome regulatory complex PAN serves multiple functions in protein degradation. *Mol Cell* **11**(1), 69-78.

Bornemann, S., Ramjee, M. K., Balasubramanian, S., Abell, C., Coggins, J. R., Lowe, D. J. & Thorneley, R. N. (1995). Escherichia coli chorismate synthase catalyzes the conversion of (6S)-6- fluoro-5-enolpyruvylshikimate-3-phosphate to 6-fluorochorismate. Implications for the enzyme mechanism and the antimicrobial action of (6S)-6-fluoroshikimate. *J Biol Chem* **270**(39), 22811-5.

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**, 248-54.

Bult, C. J., White, O., Olsen, G. J., Zhou, L., Fleischmann, R. D., Sutton, G. G., Blake, J. A., FitzGerald, L. M., Clayton, R. A., Gocayne, J. D., Kerlavage, A. R., Dougherty, B. A., Tomb, J. F., Adams, M. D., Reich, C. I., Overbeek, R., Kirkness, E. F., Weinstock, K. G., Merrick, J. M., Glodek, A., Scott, J. L., Geoghagen, N. S. & Venter, J. C. (1996). Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii. *Science* **273**(5278), 1058-73.

Chen, P. & Hochstrasser, M. (1996). Autocatalytic subunit processing couples active site formation in the 20S proteasome to completion of assembly. *Cell* **86**(6), 961-72.

Chhabra, S. R., Shockley, K. R., Conners, S. B., Scott, K. L., Wolfinger, R. D. & Kelly, R.

M. (2003). Carbohydrate-induced differential gene expression patterns in the hyperthermophilic bacterium Thermotoga maritima. *J Biol Chem* **278**(9), 7540-52.

Ciechanover, A. (1994). The ubiquitin-proteasome proteolytic pathway. *Cell* **79**(1), 13-21.

Clayton, R. A., White, O., Ketchum, K. A. & Venter, J. C. (1997). The first genome from the third domain of life. *Nature* **387**(6632), 459-62.

Copeland, R. A. (1996). *Enzymes: A practical Introduction to Structure, Mechanism, and Data Analysis*, VCH Publishers, Inc., New York, NY.

Coux, O., Tanaka, K. & Goldberg, A. L. (1996). Structure and functions of the 20S and 26S proteasomes. *Annu Rev Biochem* **65**, 801-47.

Creamer, L. K., Richardson, T. & Parry, D. A. (1981). Secondary structure of bovine alpha s1- and beta-casein in solution. *Arch Biochem Biophys* **211**(2), 689-96.

Dahlmann, B., Kopp, F., Kuehn, L., Niedel, B., Pfeifer, G., Hegerl, R. & Baumeister, W. (1989). The multicatalytic proteinase (prosome) is ubiquitous from eukaryotes to archaebacteria. *FEBS Lett* **251**(1-2), 125-31.

Dahlmann, B., Kuehn, L., Grziwa, A., Zwickl, P. & Baumeister, W. (1992). Biochemical properties of the proteasome from Thermoplasma acidophilum. *Eur J Biochem* **208**(3), 789-97.

De Mot, R., Nagy, I., Walz, J. & Baumeister, W. (1999). Proteasomes and other self-compartmentalizing proteases in prokaryotes. *Trends Microbiol* **7**(2), 88-92.

Dib, R., Chobert, J. M., Dalgalarrondo, M., Barbier, G. & Haertle, T. (1998). Purification, molecular properties and specificity of a thermoactive and thermostable proteinase from Pyrococcus abyssi, strain st 549, hyperthermophilic archaea from deep-sea hydrothermal ecosystem. *FEBS Lett* **431**(2), 279-84.

Dick, T. P., Nussbaum, A. K., Deeg, M., Heinemeyer, W., Groll, M., Schirle, M., Keilholz, W., Stevanovic, S., Wolf, D. H., Huber, R., Rammensee, H. G. & Schild, H. (1998). Contribution of proteasomal beta-subunits to the cleavage of peptide substrates analyzed with yeast mutants. *J Biol Chem* **273**(40), 25637-46.

Ditzel, L., Huber, R., Mann, K., Heinemeyer, W., Wolf, D. H. & Groll, M. (1998). Conformational constraints for protein self-cleavage in the proteasome. *J Mol Biol* **279**(5), 1187-91.

Djaballah, H. & Rivett, A. J. (1992). Peptidylglutamyl-peptide hydrolase activity of the multicatalytic proteinase complex: evidence for a new high-affinity site, analysis of cooperative kinetics, and the effect of manganese ions. *Biochemistry* **31**(16), 4133-41.

Durner, J. & Boger, P. (1995). Ubiquitin in the prokaryote Anabaena variabilis. *J Biol Chem* **270**(8), 3720-5.

Enenkel, C., Lehmann, H., Kipper, J., Guckel, R., Hilt, W. & Wolf, D. H. (1994). PRE3, highly homologous to the human major histocompatibility complex- linked LMP2 (RING12) gene, codes for a yeast proteasome subunit necessary for the peptidylglutamyl-peptide hydrolyzing activity. *FEBS Lett* **341**(2-3), 193-6.

Fenteany, G., Standaert, R. F., Lane, W. S., Choi, S., Corey, E. J. & Schreiber, S. L. (1995). Inhibition of proteasome activities and subunit-specific amino-terminal threonine modification by lactacystin. *Science* **268**(5211), 726-31.

Fiala, G. & Stetter, K. O. (1986). Pyrococcus furiosus sp. nov. represents a novel genus of marine heterotrophic archaebacteria growing optimally at 100-degrees C. *Arch Microbiol* **145**(1), 56-61.

Fitz-Gibbon, S. T., Ladner, H., Kim, U. J., Stetter, K. O., Simon, M. I. & Miller, J. H. (2002).

Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum aerophilum. *Proc Natl Acad Sci U S A* **99**(2), 984-9.

Fu, H., Doelling, J. H., Arendt, C. S., Hochstrasser, M. & Vierstra, R. D. (1998). Molecular organization of the 20S proteasome gene family from Arabidopsis thaliana. *Genetics* **149**(2), 677-92.

Fu, H., Girod, P. A., Doelling, J. H., van Nocker, S., Hochstrasser, M., Finley, D. & Vierstra, R. D. (1999). Structure and functional analysis of the 26S proteasome subunits from plants. *Mol Biol Rep* **26**(1-2), 137-46.

Gerards, W. L., de Jong, W. W., Bloemendal, H. & Boelens, W. (1998a). The human proteasomal subunit HsC8 induces ring formation of other alpha-type subunits. *J Mol Biol* **275**(1), 113-21.

Gerards, W. L., de Jong, W. W., Boelens, W. & Bloemendal, H. (1998b). Structure and assembly of the 20S proteasome. *Cell Mol Life Sci* **54**(3), 253-62.

Gerards, W. L., Enzlin, J., Haner, M., Hendriks, I. L., Aebi, U., Bloemendal, H. & Boelens, W. (1997). The human alpha-type proteasomal subunit HsC8 forms a double ringlike structure, but does not assemble into proteasome-like particles with the beta-type subunits HsDelta or HsBPROS26. *J Biol Chem* **272**(15), 10080-6.

Goldberg, A. L., Akopian, T. N., Kisselev, A. F. & Lee, D. H. (1997). Protein degradation by the proteasome and dissection of its in vivo importance with synthetic inhibitors. *Mol Biol Rep* **24**(1-2), 69-75.

Goldberg, A. L. & Rock, K. (2002). Not just research tools--proteasome inhibitors offer therapeutic promise. *Nat Med* **8**(4), 338-40.

Gottesman, S., Wickner, S. & Maurizi, M. R. (1997). Protein quality control: triage by

chaperones and proteases. *Genes Dev* **11**(7), 815-23.

Groll, M., Bajorek, M., Kohler, A., Moroder, L., Rubin, D. M., Huber, R., Glickman, M. H. & Finley, D. (2000). A gated channel into the proteasome core particle. *Nat Struct Biol* **7**(11), 1062-7.

Groll, M., Ditzel, L., Lowe, J., Stock, D., Bochtler, M., Bartunik, H. D. & Huber, R. (1997). Structure of 20S proteasome from yeast at 2.4 A resolution. *Nature* **386**(6624), 463-71.

Grziwa, A., Baumeister, W., Dahlmann, B. & Kopp, F. (1991). Localization of subunits in proteasomes from Thermoplasma acidophilum by immunoelectron microscopy. *FEBS Lett* **290**(1-2), 186-90.

Grziwa, A., Maack, S., Puhler, G., Wiegand, G., Baumeister, W. & Jaenicke, R. (1994). Dissociation and reconstitution of the Thermoplasma proteasome. *Eur J Biochem* **223**(3), 1061-7.

Gueckel, R., Enenkel, C., Wolf, D. H. & Hilt, W. (1998). Mutations in the yeast proteasome beta-type subunit Pre3 uncover position-dependent effects on proteasomal peptidase activity and in vivo function. *J Biol Chem* **273**(31), 19443-52.

Hegerl, R., Pfeifer, G., Puhler, G., Dahlmann, B. & Baumeister, W. (1991). The three-dimensional structure of proteasomes from Thermoplasma acidophilum as determined by electron microscopy using random conical tilting. *FEBS Lett* **283**(1), 117-21.

Heinemeyer, W., Kleinschmidt, J. A., Saidowsky, J., Escher, C. & Wolf, D. H. (1991). Proteinase yscE, the yeast proteasome/multicatalytic-multifunctional proteinase: mutants unravel its function in stress induced proteolysis and uncover its necessity for cell survival. *Embo J* **10**(3), 555-62.

Heinemeyer, W., Trondle, N., Albrecht, G. & Wolf, D. H. (1994). PRE5 and PRE6, the last missing genes encoding 20S proteasome subunits from yeast? Indication for a set of 14 different subunits in the eukaryotic proteasome core. *Biochemistry* **33**(40), 12229-37.

Hershko, A. (1996). Lessons from the discovery of the ubiquitin system. *Trends Biochem Sci* **21**(11), 445-9.

Hershko, A. & Ciechanover, A. (1998). The ubiquitin system. *Annu Rev Biochem* **67**, 425-79.

Hilt, W., Enenkel, C., Gruhler, A., Singer, T. & Wolf, D. H. (1993). The PRE4 gene codes for a subunit of the yeast proteasome necessary for peptidylglutamyl-peptide-hydrolyzing activity. Mutations link the proteasome to stress- and ubiquitin-dependent proteolysis. *J Biol Chem* **268**(5), 3479-86.

Hobbie, J. E., Daley, R. J. & Jasper, S. (1977). Use of nuclepore filters for counting bacteria by fluorescence microscopy. *Appl Environ Microbiol* **33**(5), 1225-8.

Hochstrasser, M. (1996). Ubiquitin-dependent protein degradation. *Annu Rev Genet* **30**, 405-39.

Ishiura, S., Nishikawa, T., Tsukahara, T., Momoi, T., Ito, H., Suzuki, K. & Sugita, H. (1990). Distribution of Alzheimer's disease amyloid A4-generating enzymes in rat brain tissue. *Neurosci Lett* **115**(2-3), 329-34.

Kaczowka, S. J. & Maupin-Furlow, J. A. (2003). Subunit topology of two 20S proteasomes from Haloferax volcanii. *J Bacteriol* **185**(1), 165-74.

Kannan, Y., Koga, Y., Inoue, Y., Haruki, M., Takagi, M., Imanaka, T., Morikawa, M. & Kanaya, S. (2001). Active subtilisin-like protease from a hyperthermophilic archaeon in a form with a putative prosequence. *Appl Environ Microbiol* **67**(6), 2445-52.

Kawarabayasi, Y., Hino, Y., Horikawa, H., Jin-no, K., Takahashi, M., Sekine, M., Baba, S., Ankai, A., Kosugi, H., Hosoyama, A., Fukui, S., Nagai, Y., Nishijima, K., Otsuka, R., Nakazawa, H., Takamiya, M., Kato, Y., Yoshizawa, T., Tanaka, T., Kudoh, Y., Yamazaki, J., Kushida, N., Oguchi, A., Aoki, K., Masuda, S., Yanagii, M., Nishimura, M., Yamagishi, A., Oshima, T. & Kikuchi, H. (2001). Complete genome sequence of an aerobic thermoacidophilic crenarchaeon, Sulfolobus tokodaii strain7. *DNA Res* **8**(4), 123-40.

Kawarabayasi, Y., Hino, Y., Horikawa, H., Yamazaki, S., Haikawa, Y., Jin-no, K., Takahashi, M., Sekine, M., Baba, S., Ankai, A., Kosugi, H., Hosoyama, A., Fukui, S., Nagai, Y., Nishijima, K., Nakazawa, H., Takamiya, M., Masuda, S., Funahashi, T., Tanaka, T., Kudoh, Y., Yamazaki, J., Kushida, N., Oguchi, A., Kikuchi, H. & et al. (1999). Complete genome sequence of an aerobic hyper-thermophilic crenarchaeon, Aeropyrum pernix K1. *DNA Res* **6**(2), 83-101, 145-52.

Kawarabayasi, Y., Sawada, M., Horikawa, H., Haikawa, Y., Hino, Y., Yamamoto, S., Sekine, M., Baba, S., Kosugi, H., Hosoyama, A., Nagai, Y., Sakai, M., Ogura, K., Otsuka, R., Nakazawa, H., Takamiya, M., Ohfuku, Y., Funahashi, T., Tanaka, T., Kudoh, Y., Yamazaki, J., Kushida, N., Oguchi, A., Aoki, K. & Kikuchi, H. (1998). Complete sequence and gene organization of the genome of a hyper- thermophilic archaebacterium, Pyrococcus horikoshii OT3. *DNA Res* **5**(2), 55-76.

Kisselev, A. F., Akopian, T. N. & Goldberg, A. L. (1998). Range of sizes of peptide products generated during degradation of different proteins by archaeal proteasomes. *J Biol Chem* **273**(4), 1982-9.

Kisselev, A. F. & Goldberg, A. L. (2001). Proteasome inhibitors: from research tools to drug

candidates. *Chem Biol* **8**(8), 739-58.

Kisselev, A. F., Songyang, Z. & Goldberg, A. L. (2000). Why does threonine, and not serine, function as the active site nucleophile in proteasomes? *J Biol Chem* **275**(20), 14831-7.

Kodama, T., Fukui, K. & Kometani, K. (1986). The initial phosphate burst in ATP hydrolysis by myosin and subfragment- 1 as studied by a modified malachite green method for determination of inorganic phosphate. *J Biochem (Tokyo)* **99**(5), 1465-72.

Kohler, A., Cascio, P., Leggett, D. S., Woo, K. M., Goldberg, A. L. & Finley, D. (2001). The axial channel of the proteasome core particle is gated by the Rpt2 ATPase and controls both substrate entry and product release. *Mol Cell* **7**(6), 1143-52.

Lanzetta, P. A., Alvarez, L. J., Reinach, P. S. & Candia, O. A. (1979). An improved assay for nanomole amounts of inorganic phosphate. *Anal Biochem* **100**(1), 95-7.

Larsen, C. N. & Finley, D. (1997). Protein translocation channels in the proteasome and other proteases. *Cell* **91**(4), 431-4.

Loe, D. W., Glover, J. R., Head, S. & Sharom, F. J. (1989). Solubilization, characterization, and detergent interactions of lymphocyte 5'-nucleotidase. *Biochem Cell Biol* **67**(4-5), 214-23.

Lowe, J., Stock, D., Jap, B., Zwickl, P., Baumeister, W. & Huber, R. (1995). Crystal structure of the 20S proteasome from the archaeon T. acidophilum at 3.4 A resolution. *Science* **268**(5210), 533-9.

Lupas, A., Flanagan, J. M., Tamura, T. & Baumeister, W. (1997a). Self-compartmentalizing proteases. *Trends Biochem Sci* **22**(10), 399-404.

Lupas, A., Kania, M. & Baumeister, W. (1998). Archaean proteasome. In *Handbook of*

Proteolytic Enzymes (Barrett, A. J., Rawlings, N. D. & Woessner, J. F., eds.), pp. 486-89. Academic Press, San Diego, CA.

Lupas, A., Zuhl, F., Tamura, T., Wolf, S., Nagy, I., De Mot, R. & Baumeister, W. (1997b). Eubacterial proteasomes. *Mol Biol Rep* **24**(1-2), 125-31.

Lupas, A., Zwickl, P., Wenzel, T., Seemuller, E. & Baumeister, W. (1995). Structure and function of the 20S proteasome and of its regulatory complexes. *Cold Spring Harb Symp Quant Biol* **60**, 515-24.

Maupin-Furlow, J. A., Aldrich, H. C. & Ferry, J. G. (1998). Biochemical characterization of the 20S proteasome from the methanoarchaeon Methanosarcina thermophila. *J Bacteriol* **180**(6), 1480-7.

Maupin-Furlow, J. A. & Ferry, J. G. (1995). A proteasome from the methanogenic archaeon Methanosarcina thermophila. *J Biol Chem* **270**(48), 28617-22.

Maupin-Furlow, J. A., Kaczowka, S. J., Ou, M. S. & Wilson, H. L. (2001). Archaeal proteasomes: proteolytic nanocompartments of the cell. *Adv Appl Microbiol* **50**, 279-338.

Maupin-Furlow, J. A., Wilson, H. L., Kaczowka, S. J. & Ou, M. S. (2000). Proteasomes in the archaea: from structure to function. *Front Biosci* **5**, D837-65.

Mayr, J., Seemuller, E., Muller, S. A., Engel, A. & Baumeister, W. (1998). Late events in the assembly of 20S proteasomes. *J Struct Biol* **124**(2-3), 179-88.

Navon, A. & Goldberg, A. L. (2001). Proteins are unfolded on the surface of the ATPase ring before transport into the proteasome. *Mol Cell* **8**(6), 1339-49.

Ng, W. V., Kennedy, S. P., Mahairas, G. G., Berquist, B., Pan, M., Shukla, H. D., Lasky, S.

R., Baliga, N. S., Thorsson, V., Sbrogna, J., Swartzell, S., Weir, D., Hall, J., Dahl, T. A., Welti, R., Goo, Y. A., Leithauser, B., Keller, K., Cruz, R., Danson, M. J., Hough, D. W., Maddocks, D. G., Jablonski, P. E., Krebs, M. P., Angevine, C. M., Dale, H., Isenbarger, T. A., Peck, R. F., Pohlschroder, M., Spudich, J. L., Jung, K. W., Alam, M., Freitas, T., Hou, S., Daniels, C. J., Dennis, P. P., Omer, A. D., Ebhardt, H., Lowe, T. M., Liang, P., Riley, M., Hood, L. & DasSarma, S. (2000). Genome sequence of Halobacterium species NRC-1. *Proc Natl Acad Sci U S A* **97**(22), 12176-81.

Orlowski, M., Cardozo, C., Hidalgo, M. C. & Michaud, C. (1991). Regulation of the peptidylglutamyl-peptide hydrolyzing activity of the pituitary multicatalytic proteinase complex. *Biochemistry* **30**(24), 5999-6005.

Pamnani, V., Tamura, T., Lupas, A., Peters, J., Cejka, Z., Ashraf, W. & Baumeister, W. (1997). Cloning, sequencing and expression of VAT, a CDC48/p97 ATPase homologue from the archaeon Thermoplasma acidophilum. *FEBS Lett* **404**(2-3), 263-8.

Puhler, G., Weinkauf, S., Bachmann, L., Muller, S., Engel, A., Hegerl, R. & Baumeister, W. (1992). Subunit stoichiometry and three-dimensional arrangement in proteasomes from Thermoplasma acidophilum. *Embo J* **11**(4), 1607-16.

Ramos, P. C., Hockendorff, J., Johnson, E. S., Varshavsky, A. & Dohmen, R. J. (1998). Ump1p is required for proper maturation of the 20S proteasome and becomes its substrate upon completion of the assembly. *Cell* **92**(4), 489-99.

Rao, M. B., Tanksale, A. M., Ghatge, M. S. & Deshpande, V. V. (1998). Molecular and biotechnological aspects of microbial proteases. *Microbiol Mol Biol Rev* **62**(3), 597-635.

Robb, F. T., Maeder, D. L., Brown, J. R., DiRuggiero, J., Stump, M. D., Yeh, R. K., Weiss, R. B. & Dunn, D. M. (2001). Genomic sequence of hyperthermophile, Pyrococcus furiosus: implications for physiology and enzymology. *Methods Enzymol* **330**, 134-57.

Ruepp, A., Eckerskorn, C., Bogyo, M. & Baumeister, W. (1998). Proteasome function is dispensable under normal but not under heat shock conditions in Thermoplasma acidophilum. *FEBS Lett* **425**(1), 87-90.

Ruepp, A., Graml, W., Santos-Martinez, M. L., Koretke, K. K., Volker, C., Mewes, H. W., Frishman, D., Stocker, S., Lupas, A. N. & Baumeister, W. (2000). The genome sequence of the thermoacidophilic scavenger Thermoplasma acidophilum. *Nature* **407**(6803), 508-13.

Schirmer, E. C., Queitsch, C., Kowal, A. S., Parsell, D. A. & Lindquist, S. (1998). The ATPase activity of Hsp104, effects of environmental conditions and mutations. *J Biol Chem* **273**(25), 15546-52.

Schmidt, M., Zantopf, D., Kraft, R., Kostka, S., Preissner, R. & Kloetzel, P. M. (1999). Sequence information within proteasomal prosequences mediates efficient integration of beta-subunits into the 20 S proteasome complex. *J Mol Biol* **288**(1), 117-28.

Schmidtke, G., Kraft, R., Kostka, S., Henklein, P., Frommel, C., Lowe, J., Huber, R., Kloetzel, P. M. & Schmidt, M. (1996). Analysis of mammalian 20S proteasome biogenesis: the maturation of beta- subunits is an ordered two-step mechanism involving autocatalysis. *Embo J* **15**(24), 6887-98.

Seemuller, E., Lupas, A. & Baumeister, W. (1996). Autocatalytic processing of the 20S proteasome. *Nature* **382**(6590), 468-71.

Seemuller, E., Lupas, A., Stock, D., Lowe, J., Huber, R. & Baumeister, W. (1995). Proteasome from Thermoplasma acidophilum: a threonine protease. *Science* **268**(5210), 579-82.

She, Q., Singh, R. K., Confalonieri, F., Zivanovic, Y., Allard, G., Awayez, M. J., Chan-Weiher, C. C., Clausen, I. G., Curtis, B. A., De Moors, A., Erauso, G., Fletcher, C., Gordon, P. M., Heikamp-de Jong, I., Jeffries, A. C., Kozera, C. J., Medina, N., Peng, X., Thi-Ngoc, H. P., Redder, P., Schenk, M. E., Theriault, C., Tolstrup, N., Charlebois, R. L., Doolittle, W. F., Duguet, M., Gaasterland, T., Garrett, R. A., Ragan, M. A., Sensen, C. W. & Van der Oost, J. (2001). The complete genome of the crenarchaeon Sulfolobus solfataricus P2. *Proc Natl Acad Sci U S A* **98**(14), 7835-40.

Shinde, U. & Inouye, M. (1994). The structural and functional organization of intramolecular chaperones: the N-terminal propeptides which mediate protein folding. *J Biochem (Tokyo)* **115**(4), 629-36.

Shockley, K. R., Ward, D. E., Chhabra, S. R., Conners, S. B., Montero, C. I. & Kelly, R. M. (2003). Heat Shock Response by the Hyperthermophilic Archaeon Pyrococcus furiosus. *Appl Environ Microbiol* **69**(4), 2365-71.

Stock, D., Ditzel, L., Baumeister, W., Huber, R. & Lowe, J. (1995). Catalytic mechanism of the 20S proteasome of Thermoplasma acidophilum revealed by X-ray crystallography. *Cold Spring Harb Symp Quant Biol* **60**, 525-32.

Tamura, T., Nagy, I., Lupas, A., Lottspeich, F., Cejka, Z., Schoofs, G., Tanaka, K., De Mot, R. & Baumeister, W. (1995). The first characterization of a eubacterial proteasome: the 20S complex of Rhodococcus. *Curr Biol* **5**(7), 766-74.

Tanaka, K. (1998). Proteasomes: structure and biology. *J Biochem (Tokyo)* **123**(2), 195-204.

Tanaka, K., Yoshimura, T., Kumatori, A., Ichihara, A., Ikai, A., Nishigai, M., Kameyama, K. & Takagi, T. (1988). Proteasomes (multi-protease complexes) as 20 S ring-shaped particles in a variety of eukaryotic cells. *J Biol Chem* **263**(31), 16209-17.

Verhagen, M. F., Menon, A. L., Schut, G. J. & Adams, M. W. (2001). Pyrococcus furiosus: large-scale cultivation and enzyme purification. *Methods Enzymol* **330**, 25-30.

Voges, D., Zwickl, P. & Baumeister, W. (1999). The 26S proteasome: a molecular machine designed for controlled proteolysis. *Annu Rev Biochem* **68**, 1015-68.

Ward, D. E., Shockley, K. R., Chang, L. S., Levy, R. D., Michel, J. K., Conners, S. B. & Kelly, R. M. (2002). Proteolysis in hyperthermophilic microorganisms. *Archaea* **1**, 63-74.

Wenzel, T. & Baumeister, W. (1993). Thermoplasma acidophilum proteasomes degrade partially unfolded and ubiquitin-associated proteins. *FEBS Lett* **326**(1-3), 215-8.

Whitby, F. G., Masters, E. I., Kramer, L., Knowlton, J. R., Yao, Y., Wang, C. C. & Hill, C. P. (2000). Structural basis for the activation of 20S proteasomes by 11S regulators. *Nature* **408**(6808), 115-20.

Wilson, H. L., Aldrich, H. C. & Maupin-Furlow, J. (1999). Halophilic 20S proteasomes of the archaeon Haloferax volcanii: purification, characterization, and gene sequence analysis. *J Bacteriol* **181**(18), 5814-24.

Wilson, H. L., Ou, M. S., Aldrich, H. C. & Maupin-Furlow, J. (2000). Biochemical and physical properties of the Methanococcus jannaschii 20S proteasome and PAN, a homolog of the ATPase (Rpt) subunits of the eucaryal 26S proteasome. *J Bacteriol* **182**(6), 1680-92.

Witt, E., Zantopf, D., Schmidt, M., Kraft, R., Kloetzel, P. M. & Kruger, E. (2000).

Characterisation of the newly identified human Ump1 homologue POMP and analysis of LMP7(beta 5i) incorporation into 20 S proteasomes. *J Mol Biol* **301**(1), 1-9.

Wolf, S., Lottspeich, F. & Baumeister, W. (1993). Ubiquitin found in the archaebacterium Thermoplasma acidophilum. *FEBS Lett* **326**(1-3), 42-4.

Yao, Y., Toth, C. R., Huang, L., Wong, M. L., Dias, P., Burlingame, A. L., Coffino, P. & Wang, C. C. (1999). alpha5 subunit in Trypanosoma brucei proteasome can self-assemble to form a cylinder of four stacked heptamer rings. *Biochem J* **344 Pt 2**, 349-58.

Zhao, D., Elimban, V. & Dhalla, N. S. (1991). Characterization of the purified rat heart plasma membrane Ca2+/Mg2+ ATPase. *Mol Cell Biochem* **107**(2), 151-60.

Zuhl, F., Seemuller, E., Golbik, R. & Baumeister, W. (1997a). Dissecting the assembly pathway of the 20S proteasome. *FEBS Lett* **418**(1-2), 189-94.

Zuhl, F., Tamura, T., Dolenc, I., Cejka, Z., Nagy, I., De Mot, R. & Baumeister, W. (1997b). Subunit topology of the Rhodococcus proteasome. *FEBS Lett* **400**(1), 83-90.

Zwickl, P., Baumeister, W. & Steven, A. (2000). Dis-assembly lines: the proteasome and related ATPase-assisted proteases. *Curr Opin Struct Biol* **10**(2), 242-50.

Zwickl, P., Grziwa, A., Puhler, G., Dahlmann, B., Lottspeich, F. & Baumeister, W. (1992a). Primary structure of the Thermoplasma proteasome and its implications for the structure, function, and evolution of the multicatalytic proteinase. *Biochemistry* **31**(4), 964-72.

Zwickl, P., Kleinz, J. & Baumeister, W. (1994). Critical elements in proteasome assembly. *Nat Struct Biol* **1**(11), 765-70.

Zwickl, P., Lottspeich, F. & Baumeister, W. (1992b). Expression of functional

Thermoplasma acidophilum proteasomes in Escherichia coli. *FEBS Lett* **312**(2-3), 157-60.

Zwickl, P., Ng, D., Woo, K. M., Klenk, H. P. & Goldberg, A. L. (1999a). An archaebacterial ATPase, homologous to ATPases in the eukaryotic 26 S proteasome, activates protein breakdown by 20 S proteasomes. *J Biol Chem* **274**(37), 26008-14.

Zwickl, P., Voges, D. & Baumeister, W. (1999b). The proteasome: a macromolecular assembly designed for controlled proteolysis. *Philos Trans R Soc Lond B Biol Sci* **354**(1389), 1501-11.

Table 6.1: The reported hypothetical and characterized archaeal proteasome components, including the proposed energy-dependent regulatory component PAN. As indicated with each organism name in the first column, the genome sequences for several of the archaeal species are not yet complete. The next three columns contain the α-type, β-type and PAN-like components, along with each gene name and number, and the similarity and identity to the corresponding components in *P. furiosus*.

Organism name (genome availability)[a]	**α subunits (gene name, gene number, identity/similarity to Pfα)**[b]	**β subunits (gene name, gene number, identity/similarity to Pfβ1; identity/similarity to Pfβ2)**[c]	**PAN-like (gene name, gene number, similarity/identity to PfPAN)**[d]
Aeropyrum pernix (F)	hypothetical α (APE1449) 47 / 68	hypothetical β (APE0521) 46 / 65; 43 / 65 hypothetical β (APE0507) 37 / 60; 41 / 59	H 26S (APE2012) 58 / 75
Archaeoglobus fulgidus (F)	psmA (AF0490) 57 / 77	psmB (AF0481) 45 / 68; 47 / 67	26S 4 (AF1976) 64 / 80
Ferroplasma acidarmanus (U)	HP (Faci1260) 48 / 70	HP (Faci0876) 46 / 67; 39 / 64	---
Halobacterium sp. NRC-1 (F)	psmB/PSMA (VNG0166G) 46 / 64	psmA (VNG0880G) 39 / 61; 39 / 60	prrIV1 / PAN1 (VNG 2000G) 50 / 71 prrIV2 / PAN2 (VNG 0510G) 52 / 71
Haloferax volcanii (U)	psmA1 (T48678) 50 / 69 psmA2 (T48679) 44 / 62	psmB1 (T48677) 37 / 61; 42 / 65	---
Methanococcus / Methanocaldococcus jannaschii (F)	psmA (MJ0591) 58 / 79	psmB (MJ1237) 48 / 69; 46 / 64	PR AAA-ATPase (MJ1176) 59 / 77

Table 6.1 ***(continued)***

Methanopyrus kandleri AV19 (F)	HslV_1 (MK0385) 61 / 77	HslV_2 (MK1228) 44 / 68; 47 / 68		RPT1 / ATPd 26S (MK0878) 59 / 77	
Methanosarcina acetivorans (F)	psmA (MA1779) 61 / 78	psmB (MA3873) 42 / 69; 44 / 64		PAN (MA4268) 58 / 79	
Methanosarcina barkeri (U)	HP (Meth1878) 61 / 78	HP (Meth3136) 45 / 71; 45 / 65		HP1 (Meth 3002) 59 / 79	HP2 (Meth 2182) 57 / 75
Methanosarcina mazei Goel (F)	α (MM2620) 61 / 77	β (MM0694) 42 / 68; 46 / 66		PAN (MM 1006) 58 / 79	26S RPT2/S (MM 0798) 50 / 71
Methanosarcina thermophila (U)	psmA (MTU30483) 62 / 78	psmB (MTU22157) 42 / 67; 44 / 65		---	
Methanothermobacter thermautotrophicus (F)	α (MTH686) 61 / 76	β (MTH1202) 42 / 68; 44 / 66		ATPd 26S 4 (MTH728) 62 / 79	
Pyrobaculum aerophilum (F)	α (PAE2215) 49 / 68	β (PAE3595) 39 / 61; 41 / 62	β (PAE0807) 39 / 60; 41 / 57	No PAN; but contains CDC / VAT-like (Zwickl *et al.*, 1999b)	
Pyrococcus abyssi (F)	psmA (PAB0417) 91 / 97	psmB (PAB1867) 75 / 88; 47 / 66	psmB (PAB2199) 47 / 69; 89 / 96	26S 4 (PAB2233) 93 / 97	
Pyrococcus furiosus (F)	psmA (PF1571) 100 / 100	psmB-1 (PF1404) 100 / 100; 48 / 69	psmB-2 (PF0159) 48 / 69; 100 / 100	PAN (PF0115) 100 / 100	
Pyrococcus horikoshii (F)	hypothetical α (PH1553) 91 / 97	hypothetical β (PH1402) 81 / 89; 47 /67	hypothetical β (PH0245) 47 / 68; 89 / 97	H 26S (PH0201) 93 / 97	

Table 6.1 ***(continued)***

Sulfolobus acidocaldarius (U)	---	---		H 26S 4 (AAO73475) 54 / 74
Sulfolobus solfataricus (F)	PSMA (SSO0738) 46 / 67	PSMB (SSO0766) 44 / 64; 46 / 66	20S subunit (SSO0278) 42 / 61; 37 / 57	AAAf ATPase (SSO0271) 57 / 76
Sulfolobus tokodaii (F)	psmA (ST0446) 47 / 66	hypothetical β (ST0477) 45 / 65; 48 / 68	hypothetical β (ST0324) 41 / 63; 39 / 59	PAN (ST0330) 56 / 73
Thermoplasma acidophilum (F)	PSMA (TA1288) 47 / 70	PSMB (TA0612) 42 / 62; 40 / 60		No PAN; but contains CDC / VAT-like (Ruepp *et al.*, 2000)
Thermoplasma volcanium (F)	α (TVN0304) 47 / 69	β (TVN0663) 42 / 62; 40 / 60		No PAN; but contains CDC / VAT-like (Ruepp *et al.*, 2000)
Thermococcus sp. KS-1 (U)	PSMA (T43887) 81 / 92	---		---

[a] F = complete genome sequence online; U = incomplete genome sequence online. Web address: http://www.ncbi.nlm.nih.gov/.

[b] Pfα = *P. furiosus* α subunit amino acid sequence; HP = hypothetical protein

[c] Pfβ1 = *P. furiosus* β1 subunit amino acid sequence; Pfβ2 = *P. furiosus* β2 subunit amino acid sequence; HP = hypothetical protein

[d] PfPAN = *P. furiosus* proteasome-activating nucleotidase (PAN) subunit amino acid sequence; H = hypothetical; ATPd = ATP-dependent; HP = hypothetical protein

Table 6.2: Comparison of *P. furiosus* differential gene expression before and after a temperature shift from 90ºC to 105ºC. Numbers represent fold-changes occurring at four time points after the *P. furiosus* culture temperature was increased. Column '0-B' represents fold-change upon temperature increase (zero time point), column '5-B' is 5 min after temperature change, column '10-B' is 10 min after temperature change, and column '60-B' represents fold-change from a separate experiment, the results of which are detailed in Shockley et al. (2003).

PF ID[a]	Function[b]	0-B	5-B	10-B	60-B[c]
PF1974	Thermosome	2.3	5.5	1.3	>4.0, >4.2
PF1719	PfpI	3.3	1.8	0.7	1.8, 1.1
PF1883	SmHSP	12.8	11.1	10.8	>7.7, >6.9
PF1882	VAT1	8.2	6.1	6.5	6.9, 3.7
PF0963	VAT2	3.1	2.1	3.7	4.6, 2.7
PF1404	β1	1.4	1.3	1.8	2.0, 2.0
PF0159	β2	1.2	0.9	1.0	1.9, 1.4
PF0115	PAN	6.1	15.3	7.4	1.2, -1.1
PF1571	α	-1.3	-2.4	-2.6	-3.9, -2.2

[a] PF ID = *P. furiosus* gene identifier
[b] PfpI = *P. furiosus* protease I; SmHSP = small heat-shock protein; VAT = valosine-containing protein-like ATPase; β = β subunit of 20S proteasome; PAN = proteasome-activating nucleotidase; α = α subunit of 20S proteasome
[c] Data from Shockely et al. (2003).

Table 6.3: Relative activities of recombinant *P. furiosus* proteasome forms on MCA-linked peptides. Expressed and purified 20S proteasome subunits from *P. furiosus* were combined in a 1:1 molar ratio and heated at either 90ºC or 105ºC for 1 hr. Activities on 16 different types of MCA-linked substrates were tested in duplicate on microtiter plates at 95ºC for 15 minutes. Flourescence readings were averaged, compared to blank values (substrate with no enzyme), and normalized to the fluorescence of the highest activity (all three subunits assembled at 90ºC tested on VKM-MCA). No activity was detected toward any of the substrates by individual subunits at either assembly temperature; nor was activity detected by the 1:1 molar ratio of β1 and β2. Extremely low or no activity was detected by all mixtures on the following MCA-linked substrates (type of activity is in parenthesis following each subustrate name): Suc-AAPF (C), Suc-LY (C), Suc-IIW (C), Bz-FVR (T = trypsin-like), Z-FR (T), Z-GGR (T), Boc-FSR (T), Boc-VLK (T), Ac-YVAD (PGPH), Ac-VEHD (PGPH), and Suc-IA (SNAAP = small neutral amino acid peptide hydrolyzing). Ac = N-acetyl; Suc = succinyl; Bz = benzoyl; Boc = *tert*-butyloxycarbonyl; Z = N-carbobenzyloxy.

Activity[a]	**Substrate**	**Normalized fluorescence of subunit mixtures:**[b]			
		α+β2 90	**α+β2 105**	**α+β1+β2 90**	**α+β1+β2 105**
C	**AAF-MCA**	10	5	10	5
	LLVY-MCA	20	10	20	10
PGPH	**LLE-MCA**	50	30	40	30
20S	**LLL-MCA**	50	50	60	50
Amy	**VKM-MCA**	100	80	100	100

[a] Type of activity: C = chymotrypsin-like; PGPH = peptidyl-glutamyl peptide hydrolyzing; 20S = 20S proteasome peptidase; Amy = amyloid A4-generating enzyme

[b] Values reported in rounded percentage compared to VKM-MCA activity by α+β1+β2; 90 = 90ºC assembly temperature; 105 = 105ºC assembly temperature

Table 6.4: Specific activity of recombinant *P. furiosus* proteasomes on VKM-MCA. Values are based on total protein per sample, which may include both full proteasome and partially assembled or unassembled subunits.

Mixture	Presence of PAN / ATP during assay[a]	Assembly temperature	Specific activity[b] (nmol / min / mg enzyme)
α+β2	-PAN / ATP	90ºC	266.4 ± 6.5
		98ºC	98.6 ± 1.6
		105ºC	140.3 ± 2.0
	+PAN / ATP	98ºC	123.2 ± 1.9
α+β1+β2	-PAN / ATP	90ºC	190.7 ± 6.6
		98ºC	164.2 ± 3.4
		105ºC	248.5 ± 7.4
	+PAN / ATP	98ºC	177.6 ± 9.9

[a] 1 mM ATP and a 4:1 (PAN:proteasome) molar ratio PAN added to substrate and buffer before preheating at 95ºC for 3 min

[b] Assays contained 5 μM VKM-MCA and 0.002 mg enzyme per ml with a reaction temperature of 95ºC. Substrate was preheated for 3 min before addition of preheated enzyme over a reaction time of 45 sec in a total volume of 250 μl. Reactions were quenched on ice and fluorescence of 100 μl was recorded (2 readings per reaction). Three reactions were run per sample and average fluorescence values were corrected for thermal degradation of substrate before activity values were calculated.

Table 6.5a: Relative effects of PAN + ATP on insulin B hydrolysis by the native and native heat-shocked proteasomes. Note that the degradation pattern (individual hydrolysis products created) of each proteasome was not affected by any of the given changes in assay condition and remained unchanged throughout all assays. All assays were carried out using 30 µg insulin B, approximately 1-1.2 µg proteasome, and listed amounts of ATP and PAN in a total volume of 200 µl. Reactions were incubated at listed temperatures for 1.5-3 hr, quenched on ice and 1% TFA was added. Products were passed through a 100,000 Da MWCO filter before being separated in an acetonitrile gradient through a C_{18} Nucleosil column.

Assay conditions			**Relative extent of hydrolysis normalized to 90ºC (no PAN / ATP)**[c]	
Assay T[a] **(ºC)**	**ATP (mM)**	**Molar ratio 20S:PAN**[b]	**Native**	**Native heat-shocked**
90	--	--	+ (1.0)	+++++ (1.00)
90	1	--	+ (1.0)	+++++ (1.00)
90	--	1:4	+ (1.0)	+++++ (1.00)
90	1	1:4	++(1.1)	++++ (0.8)
80	--	--	+++ (1.6)	++++ (0.8)
80	1	--	+++ (1.6)	++++ (0.8)
80	--	1:4	+++ (1.6)	++++ (0.8)
80	1	1:1	++++ (1.7)	+++ (0.6)
80	1	1:4	++++ (1.7)	++ (0.6)
80	1	1:10	++++ (1.7)	+ (0.5)

[a] T = temperature

[b] 20S = 20S proteasome; Molar ratio was calculated assuming a 20S proteasome molecular weight of 650 kDa and a PAN molecular weight of 550 kDa and that all protein mass was in the assembled 20S or PAN structure.

[c] Qualitative trends are indicated with '+' symbols and shading, and are based on simultaneous decrease in substrate insulin B and increase in hydrolysis products in chromatograms resulting from RP-HPLC separation of degradation products. Quantitative comparisons are in parenthesis and represent the percent hydrolysis of insulin B per hour of assay incubation time (% hydrolysis / hr). Values were determined by comparing areas under the insulin B peaks. Values were then normalized to the 90ºC (no PAN or ATP added) data point within in each column.

Table 6.5b: Relative effects of PAN + ATP on insulin B hydrolysis by the recombinant proteasomes. Note that the degradation pattern (individual hydrolysis products created) of each proteasome was not affected by any of the given changes in assay condition and remained unchanged throughout all assays. All assays were carried out using 30 µg insulin B, approximately 1-1.2 µg proteasome, and listed amounts of ATP and PAN in a total volume of 200 µl. Reactions were incubated at listed temperatures for 1.5-3 hr, quenched on ice and 1% TFA was added. Products were passed through a 100,000 Da MWCO filter before being separated in an acetonitrile gradient through a C_{18} Nucleosil column.

Assay conditions			**Relative extent of hydrolysis normalized to 90ºC (no PAN / ATP)[c]**					
Assay T[a] (ºC)	**ATP (mM)**	**Molar ratio 20S:PAN[b]**	**α+β2 90**	**α+β2 98**	**α+β2 105**	**α+β1+β2 90**	**α+β1+β2 98**	**α+β1+β2 105**
90	--	--	++ (1.0)	+ (1.0)	++ (1.0)	+ (1.0)	++ (1.0)	++ (1.0)
90	1	--	++ (1.0)	+ (1.0)	++ (1.0)	+ (1.0)	++ (1.0)	++ (1.0)
90	--	1:4	++ (1.0)	+ (1.0)	++ (1.0)	+ (1.0)	++ (1.0)	++ (1.0)
90	1	1:4	+++ (1.1)	++ (1.2)	+ (0.9)	+ (1.0)	+++ (1.1)	+ (0.9)
80	--	--	+ (0.9)	+ (1.0)	NT[d]	+ (1.0)	+ (0.8)	NT
80	1	--	+ (0.9)	+ (1.0)	NT	+ (1.0)	+ (0.8)	NT
80	--	1:4	+ (0.9)	+ (1.0)	NT	+ (1.0)	+ (0.8)	NT
80	1	1:1	+++ (1.1)	+++ (1.4)	NT	+++ (1.2)	+++ (1.1)	NT
80	1	1:4	+++ (1.1)	+++ (1.4)	NT	+++ (1.2)	+++ (1.1)	NT
80	1	1:10	+++ (1.1)	+++ (1.4)	NT	+++ (1.2)	+++ (1.1)	NT

[a] T = temperature

[b] 20S = 20S proteasome; Molar ratio was calculated assuming a 20S proteasome molecular weight of 650 kDa and a PAN molecular weight of 550 kDa and that all protein mass was in the assembled 20S or PAN structure.

[c] Qualitative trends are indicated with '+' symbols and shading, and are based on simultaneous decrease in substrate insulin B and increase in hydrolysis products in chromatograms resulting from RP-HPLC separation of degradation products. Quantitative comparisons are in parenthesis and represent the percent hydrolysis of insulin B per hour of assay incubation time (% hydrolysis / hr). Values were determined by comparing areas under the insulin B peaks. Values were then normalized to the 90ºC (no PAN or ATP added) data point within in each column.

[d] NT = not tested

Table 6.6: Activity of recombinant *P. furiosus* PAN on nucleotides and the stimulation of nucleotidase activity by the presence of the proteasome substrate β-casein (1000-fold molar excess over PAN).

Nucleotide	Hydrolysis rate[a] (nmol / hr / µg)	Relative rate[c] (%)	Hydrolysis rate with β-casein[d] (nmol / hr / µg)	Hydrolysis rate fold-change[e]
ATP	62 ± 3	100	196 ± 3	3
CTP	55 ± 4	88	126 ± 4	2
GTP	9 ± 2	15	35 ± 2	4
UTP	4 ± 1	7	16 ± 1	4
ADP	UD[b]	0	UD	0
AMP-PNP	UD	0	UD	0

[a] Determined by release of organic phosphate (P_i) from nucleotide substrate in 50 mM Tris pH 8.9 + 100 mM NaCl + 10 mM $MgCl_2$ + 1 mM DTT. Nucleotide concentration was 1 mM, with 0.75 µg PAN used in a 50-µl assay volume. Mixtures were heated at 90ºC for 5 min, quenched on ice, and P_i was detected using the modified molybdate/malachite green method (Lanzetta, 1979). Values reported are averages of two separate assays, with each run in triplicate.

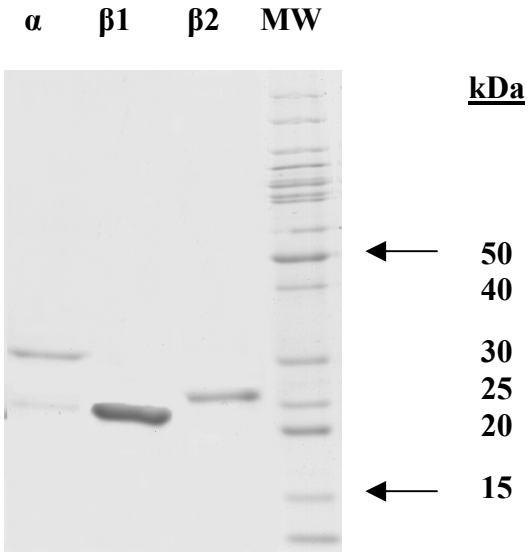
[b] UD = undetectable activity

[c] Relative rate is compared to the hydrolysis rate of PAN on ATP.

[d] The usual assay was run to detect release of P_i, except 1000-fold molar excess of β-casein over PAN was added immediately before addition of the enzyme.

[e] Relative stimulation of PAN nucleotidase activity by the presence of β-casein. Values represent the ratio of the hydrolysis rate with β-casein in the assay against the hydrolysis rate without β-casein.

a)



b)

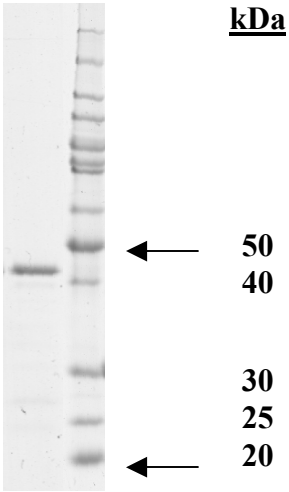
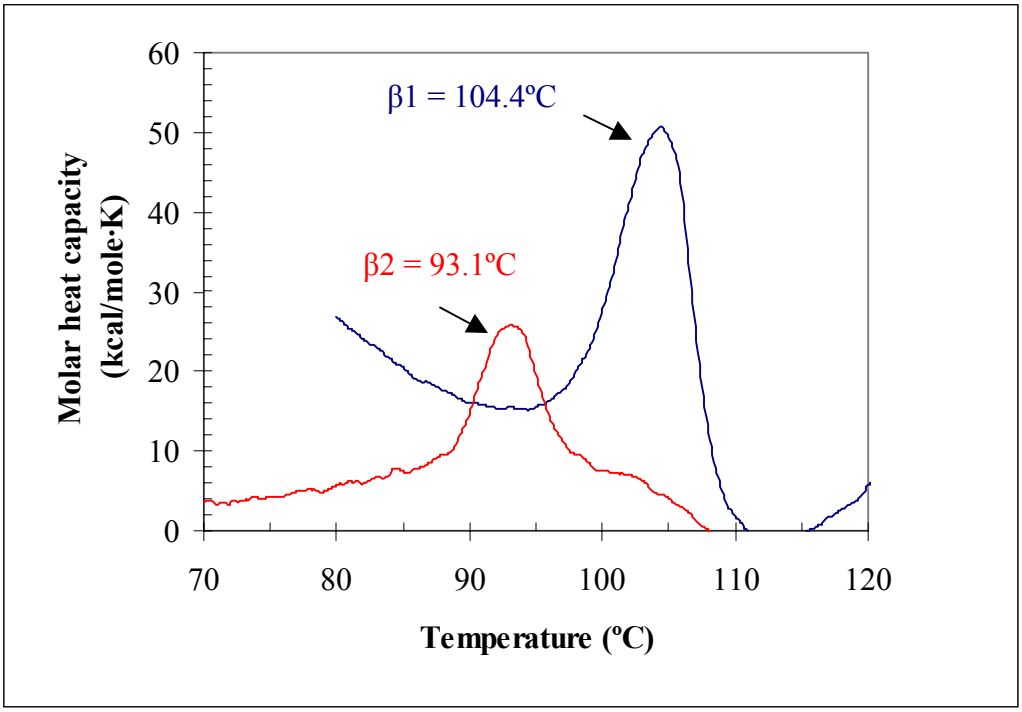
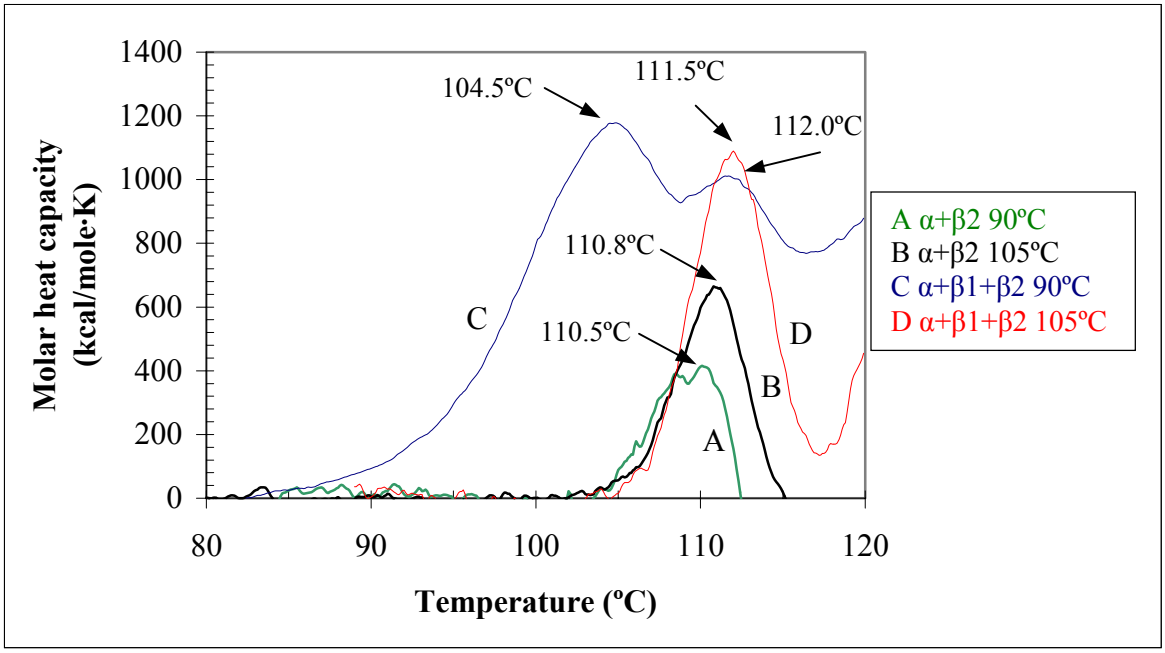


Figure 6.1: a) A 12.5% SDS-PAGE with the expressed and purified *P. furiosus* proteasome subunits in lanes 1-3 and a protein ladder in lane 4, with corresponding molecular weights (kDa) listed to the right of the gel. The α subunit ran at approximately 30 kDa (hypothetical molecular weight of 20.009 kDa), the β1 ran at 23 kDa (hypothetical molecular weight of 22.002 kDa), and the β2 ran at 26 kDa (hypothetical molecular weight of 21.649 kDa).
b) A 12.5% SDS-PAGE gel showing the *P. furiosus* proteasome-activating nucleotidase (PAN) expressed in *E. coli* and purified. PAN is in the left-hand lane and a protein ladder is in the right-hand lane, with corresponding molecular weights (kDa) to the right of the gel. The estimated molecular weight of PAN, based on sequence data, is 44.805 kDa and the observed molecular weight is 43 kDa.

a)


β1 = 104.4ºC
β2 = 93.1ºC
Molar heat capacity (kcal/mole·K)
Temperature (ºC)
0
10
20
30
40
50
60
70
80
90
100
110
120

b)


104.5ºC
111.5ºC
112.0ºC
110.8ºC
110.5ºC
C
D
B
A
A α+β2 90ºC
B α+β2 105ºC
C α+β1+β2 90ºC
D α+β1+β2 105ºC
Molar heat capacity (kcal/mole·K)
Temperature (ºC)
0
200
400
600
800
1000
1200
1400
80
90
100
110
120

c)

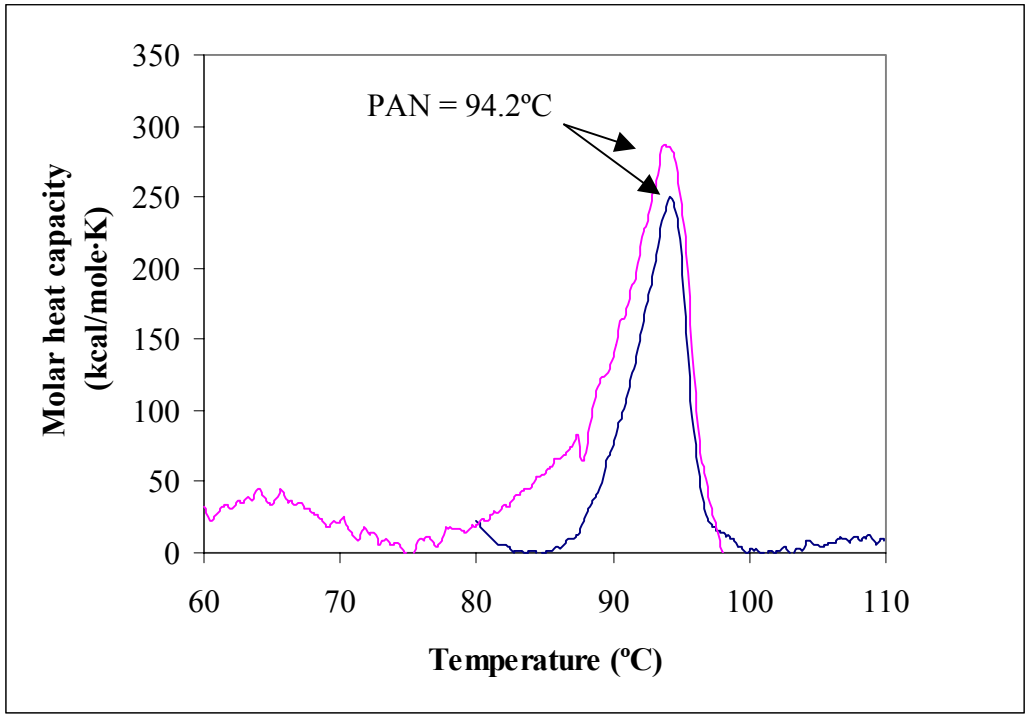


Figure 6.2: Differential scanning calorimetry curves showing melting points of *P. furiosus* recombinant proteasome subunits (part a), subunit combinations (part b), and recombinant PAN (part c). The α subunit did not exhibit a melting peak, whereas the two β subunits (a) melted at very different temperatures. All subunit mixtures are shown on a single plot (b), with each subunit mixture and corresponding melting curve assigned a letter in the following order: A = α+β2 assembled at 90ºC, B = α+β2 assembled at 105ºC, C = α+β1+β2 assembled at 90ºC, and D = α+β1+β2 assembled at 105ºC. Each subunit mixture (b) exhibited a peak in molar heat capacity within the range 110.5-112ºC, with the two forms containing the β1 subunit at the upper end of the range. The α+β1+β2 form assembled at 90ºC also exhibited a transition at 104.5ºC, which was the determined melting temperature of the recombinant β1 subunit isolated in solution. The recombinant PAN (c) was analyzed in two separate scans, with both melting peaks occurring at 94.2ºC.

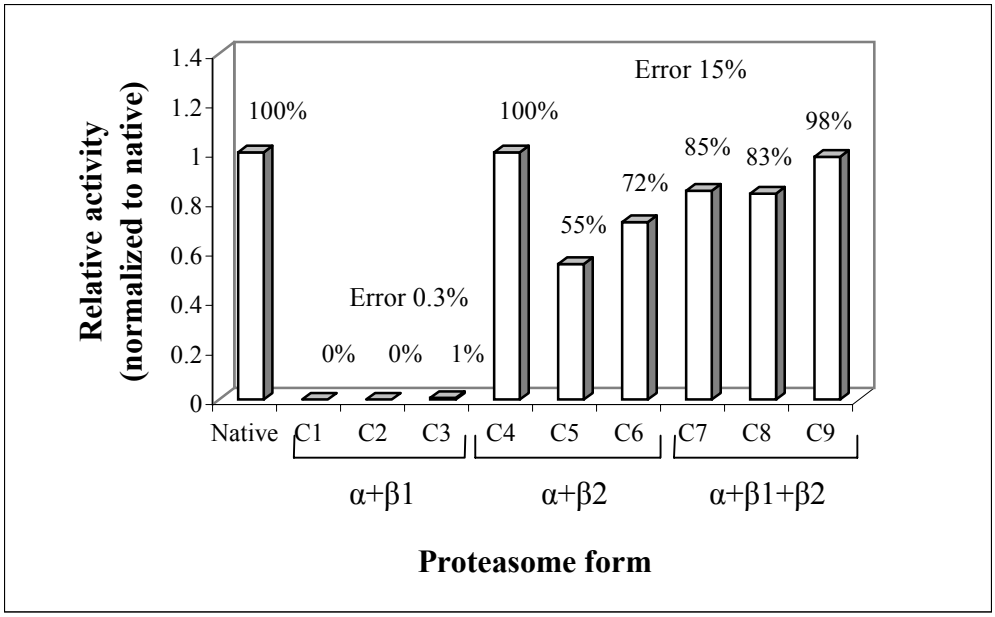
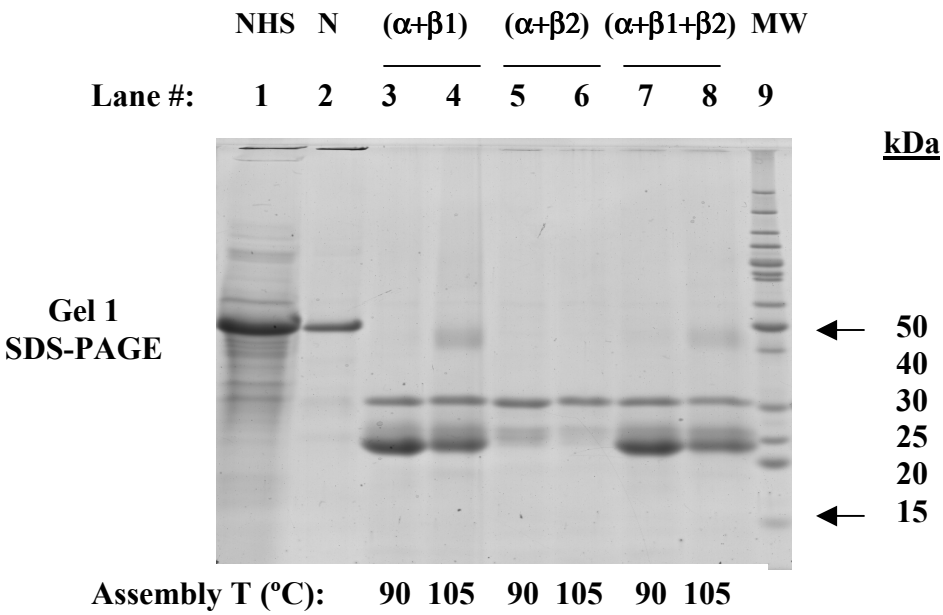


Figure 6.3: The relative activity of the native and recombinant proteasome forms on the peptide substrate VKM-MCA. Endpoint assays were used, with mixtures containing 5 μM VKM-MCA and 0.3 μg protein in microtiter plates. Incubation was at 95ºC for 15 min. C1, C4, C7 = 90ºC assembly temperature; C2, C5, C8 = 98ºC assembly temperature; and C3, C6, C9 = 105ºC assembly temperature.

a)


NHS
N
(α+β1)
(α+β2)
(α+β1+β2)
MW
Lane #: 1 2 3 4 5 6 7 8 9
kDa
Gel 1
SDS-PAGE
50
40
30
25
20
15
Assembly T (ºC): 90 105 90 105 90 105

b)

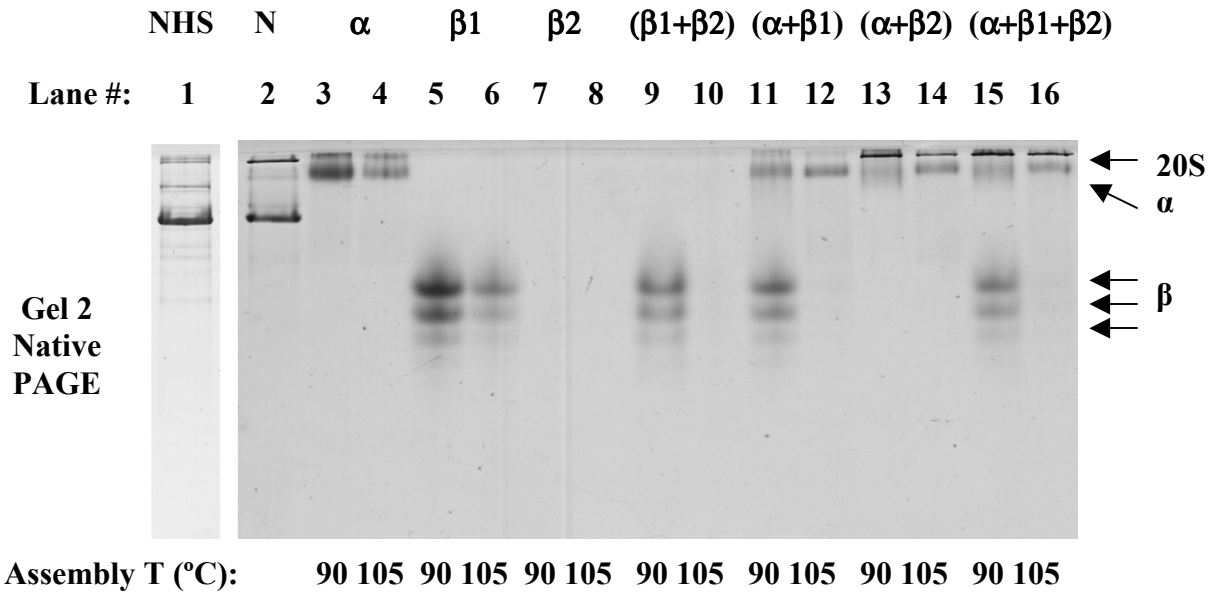
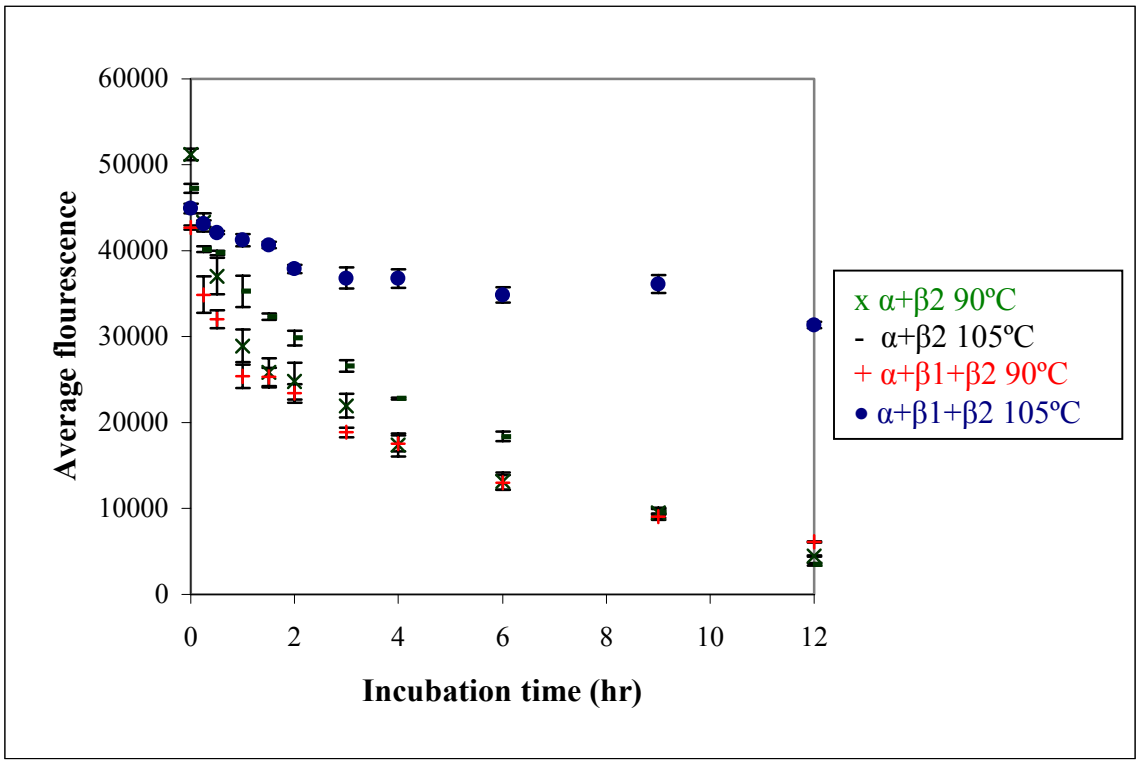

NHS
N
α
β1
β2
(β1+β2)
(α+β1)
(α+β2)
(α+β1+β2)
Lane #: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
Gel 2
Native
PAGE
20S
α
β
Assembly T (ºC): 90 105 90 105 90 105 90 105 90 105 90 105 90 105

Figure 6.4: a) **Gel 1:** Combinations of expressed and purified *P. furiosus* 20S proteasome subunits on 12.5% SDS-PAGE. The most active combinations (α+β1, α+β2, and α+β1+β2) are shown when assembled (1 hr incubation) at 90ºC versus 105ºC. Lane 1 / 2: native heat-shocked (NHS) / native 20S proteasome (N) purified from biomass; lane 3: α+β1 90ºC; lane 4: α+β1 105ºC; lane 5: α+β2 90ºC; lane 6: α+β2 105ºC; lane 7: α+β1+β2 90ºC; lane 8: α+β1+β2 105ºC; lane 9: protein ladder with corresponding molecular weights listed at right.
b) **Gel 2:** Recombinant *P. furiosus* 20S proteasomes on 10% native PAGE. Subunits were combined as listed and incubated at 90ºC or 105ºC for 1 hr. Lane 1: native 20S proteasome purified from heat-shocked *P. furiosus* biomass (NHS); lane 2: native 20S proteasome purified from *P. furiosus* biomass (N); lane 3: α 90ºC; lane 4: α 105ºC; lane 5: β1 90ºC; lane 6: β1 105ºC; lane 7: β2 90ºC; lane 8: β2 105ºC; lane 9: β1+β2 90ºC; lane 10: β1+β2 105ºC; lane 11: α+β1 90ºC; lane 12: α+β1 105ºC; lane13: α+β2 90ºC; lane 14: α+β2 105ºC; lane 15: α+β1+β2 90ºC; lane 16: α+β1+β2 105ºC.

a)



Proteasome form	Decay constant (k_{obs})
α+β2 90ºC	0.18 ± 0.01
α+β2 105ºC	0.19 ± 0.01
α+β1+β2 90ºC	0.15 ± 0.01
α+β1+β2 105ºC	0.025 ± 0.004

b)

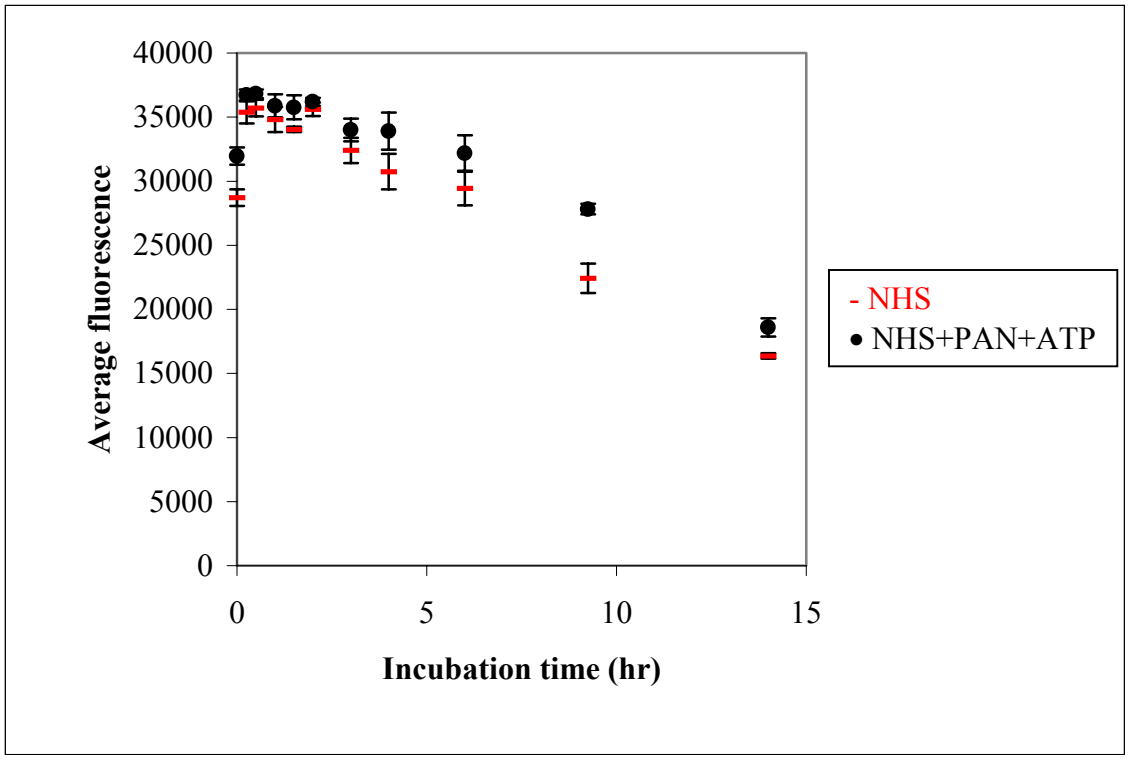
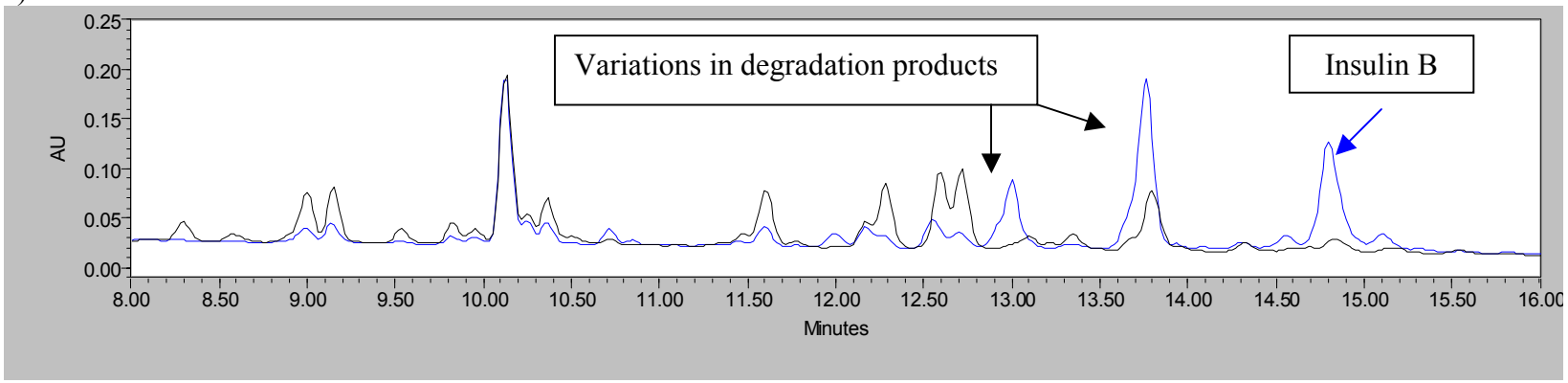
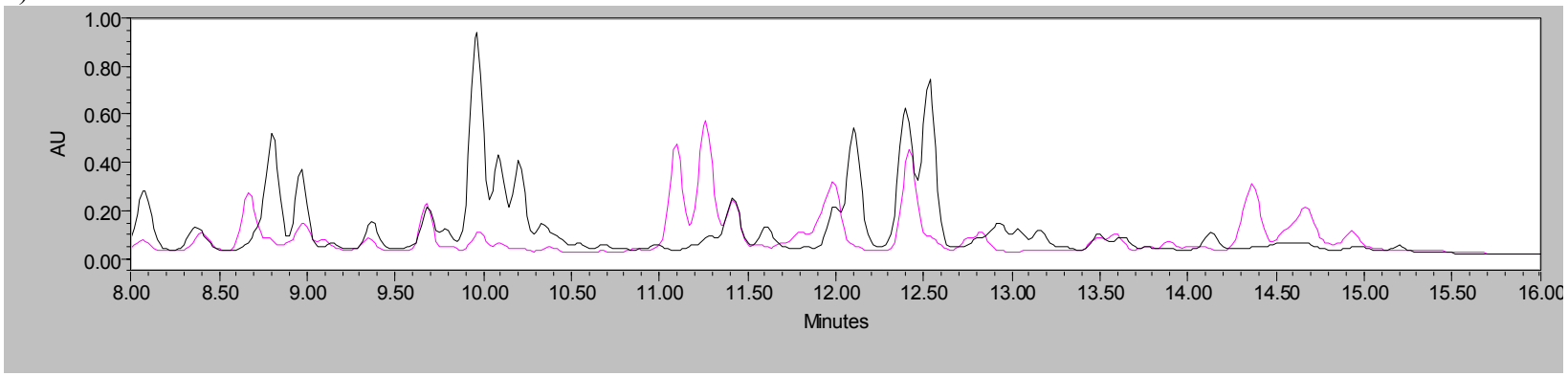


Figure 6.5: Average fluorescence of *P. furiosus* proteasomes versus incubation time at 115ºC. Error bars are shown on plots.
a) Stability of recombinant proteasomes. Samples were adjusted to a final concentration of 0.15 mg/ml and heated in an oil bath at 115ºC for up to 12 hr. Aliquots were removed at the selected time points and stored on ice until the end of the 115ºC incubation. VKM-MCA endpoint activities for all points were determined using a fixed volume of each aliquot in triplicate in microtiter plates and at an assay incubation temperature and time of 95ºC and 15 min, respectively.
b) Stability of the native heat-shocked proteasome (NHS). Experimental procedure was identical to that outlined in part a) for the recombinant proteasome stability study, with total incubation time at 115ºC increased to 14 hr. The addition of a 4:1 molar ratio of PAN over the proteasome did not have an affect on stability.

a)


Variations in degradation products
Insulin B
AU
0.25
0.20
0.15
0.10
0.05
0.00
8.00
8.50
9.00
9.50
10.00
10.50
11.00
11.50
12.00
12.50
13.00
13.50
14.00
14.50
15.00
15.50
16.00
Minutes

b)


AU
1.00
0.80
0.60
0.40
0.20
0.00
8.00
8.50
9.00
9.50
10.00
10.50
11.00
11.50
12.00
12.50
13.00
13.50
14.00
14.50
15.00
15.50
16.00
Minutes

c)

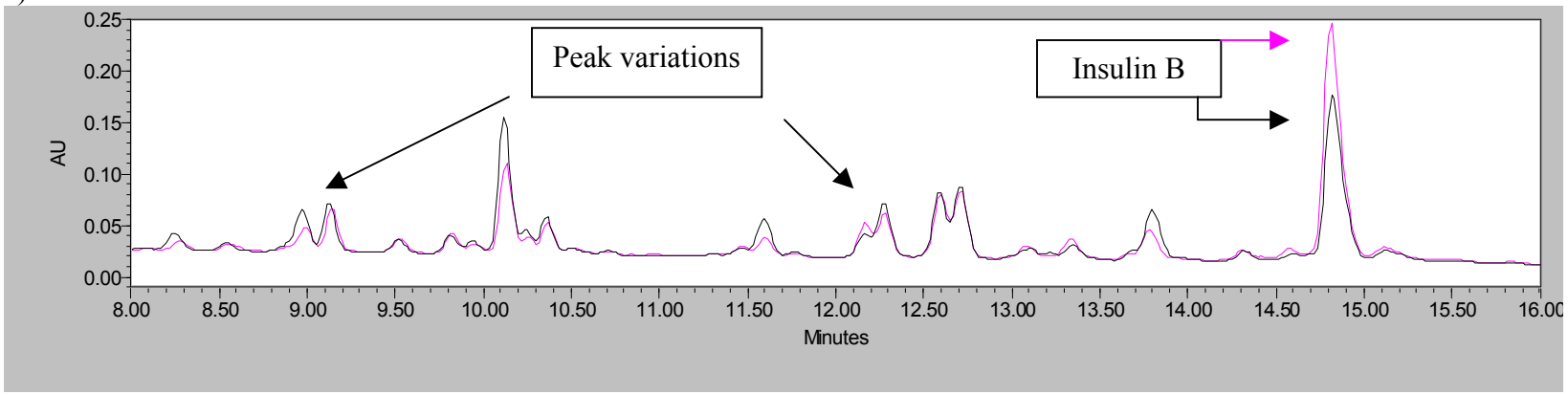
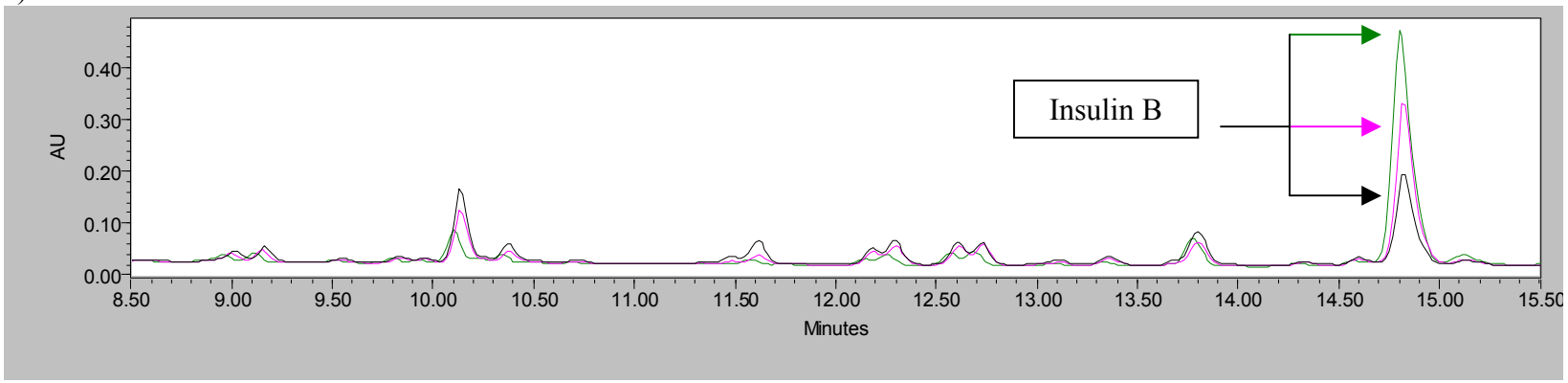


Figure 6.6: Chromatograms from HPLC separation of proteasome degradation products on insulin B. a) Native heat-shocked proteasome (blue) overlay with α+β1+β2 assembled at 105ºC. The assay temperature was 90ºC and incubation time was 2 hr. b) Native heat-shocked proteasome (pink) overlay with α+β1+β2 assembled at 90ºC. The assay temperature was 80ºC and the incubation time was 4 hr. c) α+β2 assembled at 90ºC (pink) overlay with α+β1+β2 assembled at 90ºC. The assay temperature was 90ºC and incubation time was 2 hr.

a)



b)

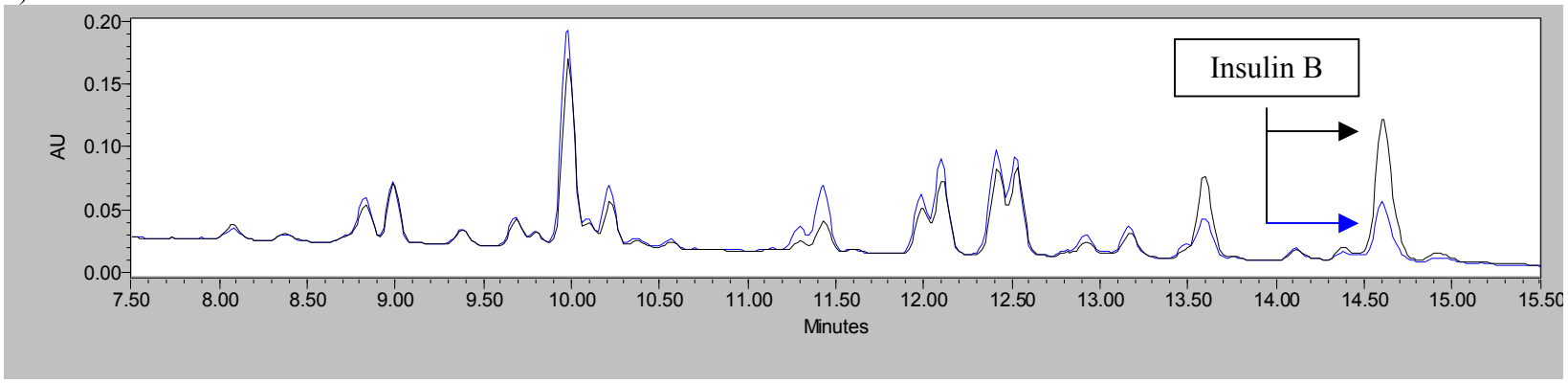
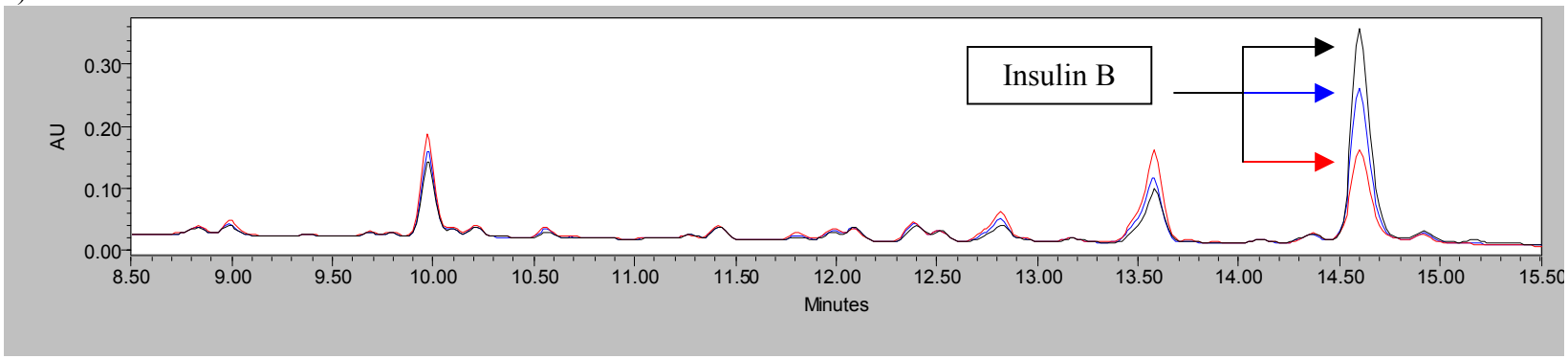


Figure 6.7: Chromatograms from HPLC separation of native proteasome degradation products on insulin B showing effects of temperature and PAN/ATP. a) Native proteasome at 90ºC (green) is least active, followed by the more favorable temperature of 80ºC (pink), and the most active conditions 80ºC with PAN (1:4 molar ratio with proteasome) and ATP (black). Assay time was 2 hr. b) Native proteasome at 80ºC without PAN/ATP (black) and with PAN (1:1 molar ratio with proteasome) and ATP. Assay time was 3 hr.

a)



b)

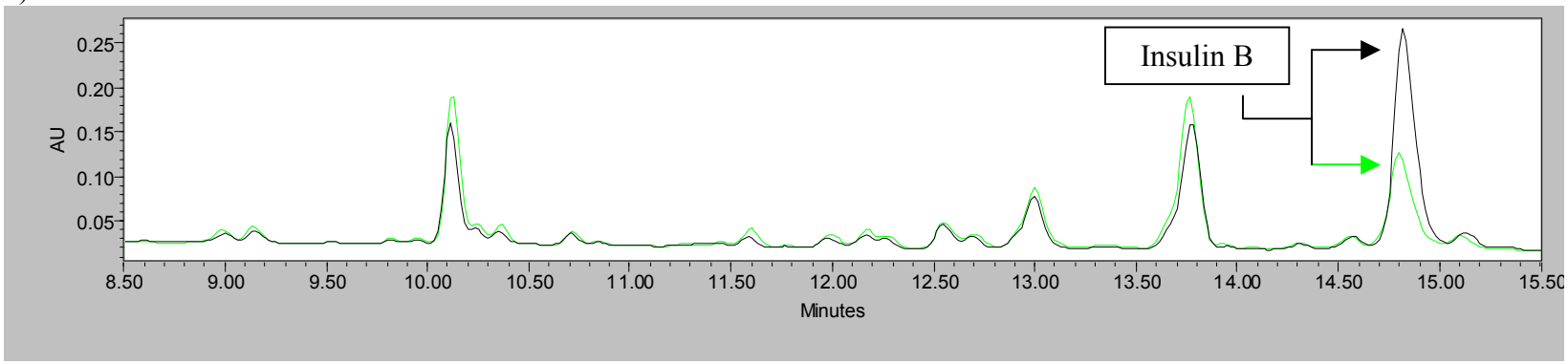
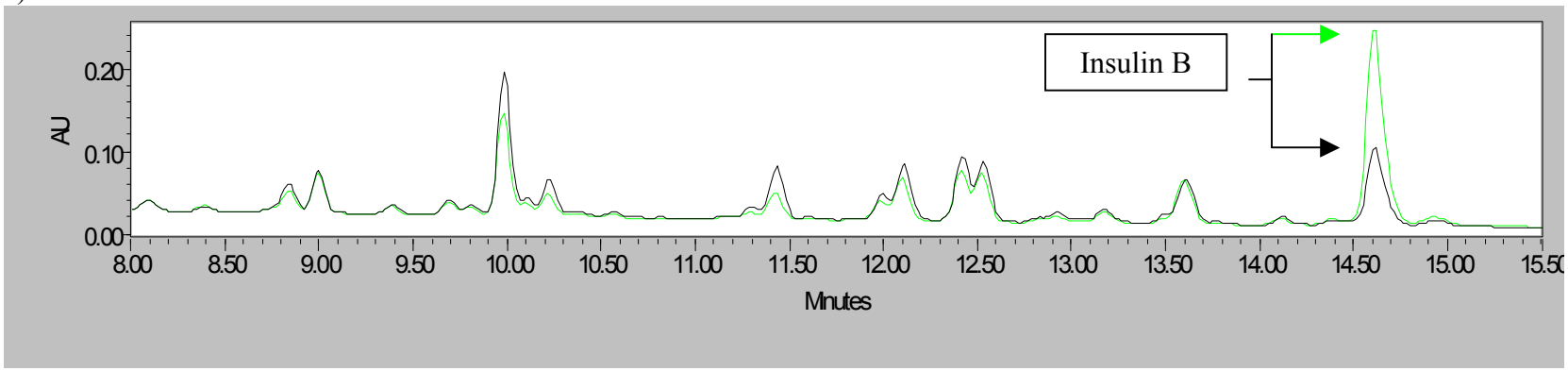
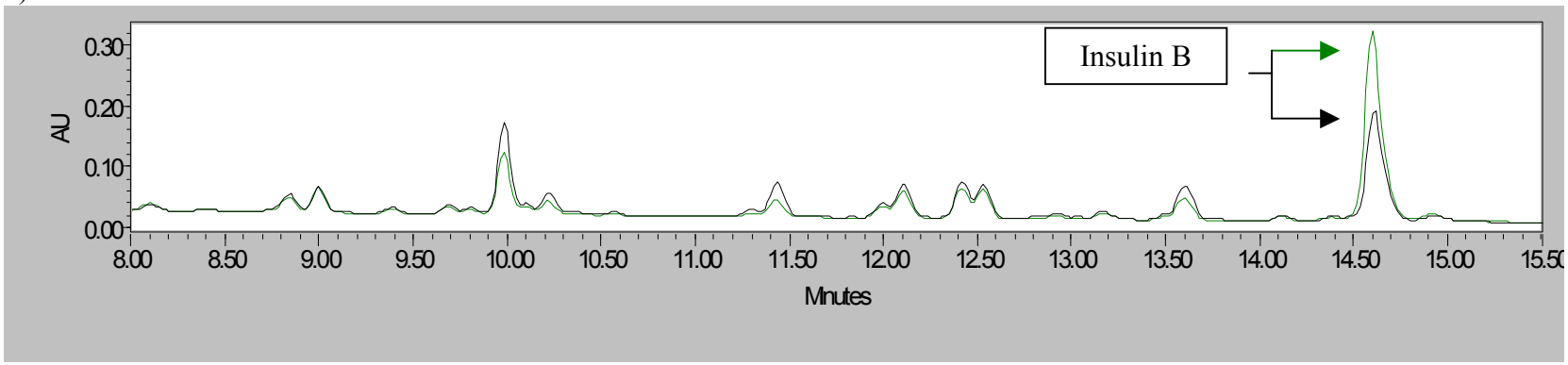


Figure 6.8: Chromatograms from HPLC separation of native heat-shocked proteasome (NHS) degradation products on insulin B showing the inhibitory effects of PAN/ATP at 80ºC and 90ºC. a) NHS at 80ºC with PAN (1:10 molar ratio with proteasome) and ATP (black) is least active, followed by the more favorable PAN ratio of 1:1 (blue), and the most active condition at 80ºC with the absence of PAN and ATP (red). Assay time was 3 hr. b) NHS at 90ºC without PAN/ATP (green) and with PAN (1:4 molar ratio with proteasome) and ATP. Assay time was 2 hr.

a)


Insulin B
AU
0.20
0.10
0.00
8.00
8.50
9.00
9.50
10.00
10.50
11.00
11.50
12.00
12.50
13.00
13.50
14.00
14.50
15.00
Minutes

b)


Insulin B
AU
0.30
0.20
0.10
0.00
8.00
8.50
9.00
9.50
10.00
10.50
11.00
11.50
12.00
12.50
13.00
13.50
14.00
14.50
15.00
Minutes

c)

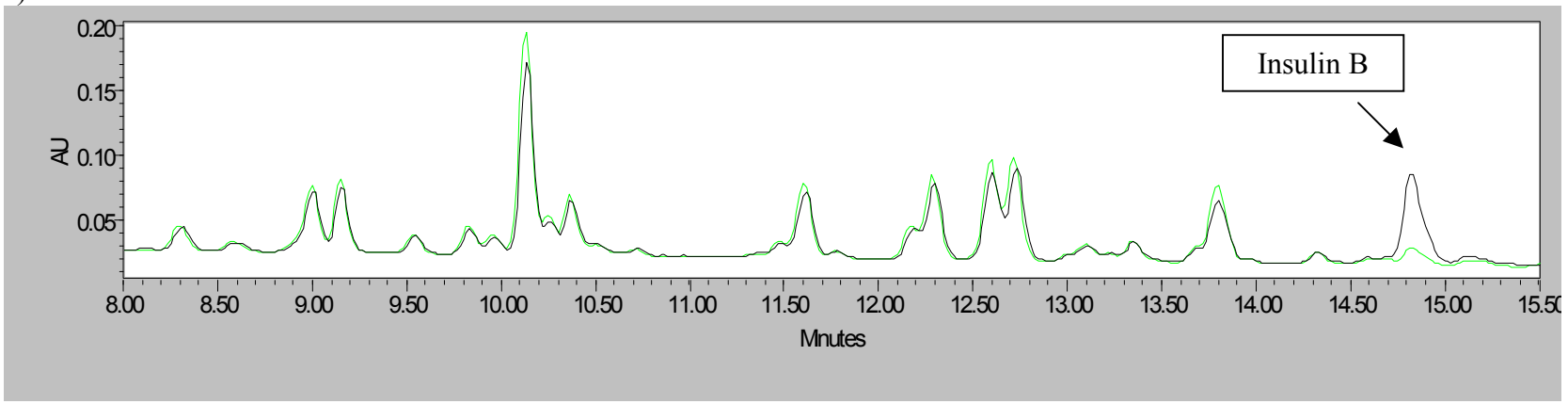


Figure 6.9: Chromatograms from HPLC separation of recombinant proteasome degradation products on insulin B showing effects of PAN/ATP. a) The effect of PAN/ATP on the activity at 80ºC of the α+β1+β2 proteasome assembled at 90ºC. The effects on the α+β2 structure assembled at 90ºC are identical. The activity is stimulated by PAN (molar ratio with proteasome does not matter) and ATP (black). Assay time was 3 hr. b) The effect of PAN/ATP on the activity at 80ºC of the α+β1+β2 proteasome assembled at 98ºC. The activity is stimulated by PAN (molar ratio with proteasome does not matter) and ATP (black). Assay time was 3 hr. c) The effect of PAN/ATP on the activity at 90ºC of the α+β1+β2 proteasome assembled at 105ºC. The activity is inhibited by PAN (1:4 molar ratio with proteasome) and ATP (black). Assay time was 2 hr.

CHAPTER 7:

Concluding remarks

and

future work

Two multisubunit, self-compartmentalizing proteases from the hyperthermophilic archaeon *Pyrococcus furiosus* were investigated in this series of studies. The first was the homo-mulitimeric cysteine protease, PfpI (*Pyrococcus furiosus* protease I), and the second was the hetero-multimeric threonine protease containing three individual subunits, the 20S proteasome. The regulatory homo-multimeric ATPase, named proteasome-activating nucleotidase (PAN), was also studied in combination with the proteasome. From the observations outlined in this dissertation, the most likely physiological role of these two proteases may involve their combined activities to achieve full degradation of protein substrates. Specifically, the proteasome was found to break down larger polypeptides, while PfpI was only able to degrade small oligopeptides. Within the cell, a protein substrate, which may be a protein that has been misfolded or has completed its function within the cell, must be recycled so its amino acid residues may be used to rebuild new proteins. In the case of *P. furiosus*, the likely sequence of events, based on evidence described here, may be initialized by the action of the proteasome and other energy-dependent proteinases degrading the protein substrates into smaller peptide pieces. Next, peptidases such as PfpI may break these peptide pieces down into individual residues or di- and tri-peptides. However, to fully prove this theory, studies focused on the activities of PfpI and the proteasome together must be carried out.

As outlined in chapters 3 and 4, much has been accomplished in elucidating the activity and basic structure of PfpI. However, the remaining difficulty lies in expression of the recombinant form. Several expression attempts were made in this study, as outlined in Appendix A. In all cases, a stable assembly of subunits was created but never exhibited activity. Since the DNA inserts were sequenced and always remained correct, it is likely that

the correct PfpI subunit was synthesized. The absence of activity was in the formation of the active site during folding and assembly. It would be useful to obtain active forms of this so that DNA-based investigations can be done, including mutational studies of the putative active-site region. In the future, expression of PfpI must make use of chaperones to aid in functional active site formation. An initial attempt might involve expression of the same PfpI from the pET system as outlined in Appendix A. However, immediately after sonication, a crude extract of *P. furiosus* intracellular protein could be added to the *Escherichia coli* cell extract containing the expressed PfpI. The *P. furiosus* extract may likely contain a native PfpI chaperone (assuming one exists), which could facilitate the folding of the active sites within the expressed PfpI. If this does not work, the reason could be that the expressed PfpI is already misfolded at the active site. In this case, the next step could be to unfold this protein using a common denaturant such as guanidine hydrochloride (6 M final concentration) and then add the *P. furiosus* cell extract to the expressed and unfolded PfpI while simultaneously diluting out (refolding) the PfpI. As a final attempt, PfpI could be cloned and expressed again using a different expression system that makes use of a general chaperone during expression in *E. coli*.

Even if synthesis of a PfpI active form is not achieved, there are several critical studies that should be carried out. The most important studies include: a careful look at the possible ATP-dependence of PfpI, its activity on polypeptides in the presence of other ATPases such as PAN and VAT (PfpI did not show activity on longer polypeptides in these studies), and its activity in the presence of the proteasome. In particular, the synergy of protein degradation, both with short unfolded proteins (insulin and casein) and with longer

folded proteins (*P. furiosus* proteins), between the proteasome and PfpI may yield useful information.

The observations of this study offered clues into the roles of the three subunits from the 20S proteasome. Specifically, the α subunit may be a "gate" into the active sites and a matrix for assembly of the structure, including processing and assembly of the two β subunits. The β1 may be a non-catalytic stabilizing factor, and β2 may be the catalytic center. Furthermore, the varying interactions of PAN with the different proteasomes proved that the proteasome exists in different forms depending on the cellular environment. However, to fully prove these theories beyond a doubt, several additional studies would be useful. First, the different assemblies of α + β1 + β2 (low- versus high-temperature) should be combined in relatively high concentration, purified (to separate the individual subunit structures from the assembled proteasome), and the compositions of the remaining proteasomes determined. This may be able to show the absence of β1 in the low-temperature assembly as was theorized in Chapter 6. If possible, the crystal structures of each proteasome form should be found. However, an extremely high concentration of each proteasome form, and subsequently each subunit will have to be achieved.

Beyond their physical compositions, the activities of the various proteasome forms on polypeptide substrates other than insulin should be investigated. It would be particularly useful to look at substrates in a folded form, such as small proteins from *P. furiosus*. In theory, the proteasomes should not have the ability to break down these types of substrates without the presence of ATPases. But it is also possible that the simple presence of the ATPases would not be enough to impart proteolytic activity and that a form of substrate recognition (homologous to the eukaryotic ubiquitin) is necessary for cleavage of such

substrates. However, this is still yet to be determined. Finally, the other possible associating ATPase, VAT, should be either purified or recombinantly expressed to determine the presence or absence of its interaction with the proteasome. This may be particularly interesting in combination with the native-heat shocked proteasome, which exhibited decreased activity in the presence of PAN ATPase activity.

APPENDIX A:

Cloning and expression of protease I from *Pyrococcus furiosus* and *Pyrococcus horikoshii*

Cloning and expression were attempted on two forms of protease I, including those from *Pyrococcus furiosus* (PfpI) and *Pyrococcus horikoshii* (PhpI). In both cases, the genes encoding each PfpI form were cloned into the pET-21b(+) vector. The gene encoding PfpI was also cloned into the intein-based vector pCYB1 (IMPACT –CN system or Intein Mediated Purification with an Affinity Chitin-Binding Tag; New England Biolabs, Beverly, MA). After cloning, inserts were sequenced and found to be correct. Following the cloning, expression was attempted under various conditions. Initially, the common expression run was attempted, with induction by 0.8 mM IPTG and a 3-4 hr period at 37ºC and 250 rpm before harvesting of cells. Next, the induction period was altered so that the temperature after induction was 18ºC or 28ºC and the length was varied within the range to 4-8 hr; varying amounts of IPTG were used, ranging from 0.2 mM – 1 mM. In addition, cells were sonicated after a –20ºC freeze/thaw cycle initially, with this step later omitted in the possible case that the freezing process had a negative affect on the expressed protein. Furthermore, additives such as EDTA and DTT were added to harvested cells before release of the expressed protease I from the cells during sonication. These additives were then kept in the buffer at all times. After all of these expression runs, stable forms of PfpI and PhpI were present, appearing at several sizes on SDS-PAGE, but never exhibited activity on azocasein, gelatin-based zymogram, or the ideal MCA-linked substrate N-Suc-AAF-MCA (previous studies reported native PfpI proteolytically active on all of these substrates). One representative gel of an expressed PhpI, after expression and after heat treatment is illustrated in Figure A.1.

Several attempts were made to stimulate the activity of the expressed PfpI and PhpI. These included:

1) Various heat-treatment based purification steps to remove any possible inhibiting factors;
2) guanidine hydrochloride refolding procedures (addition of 6M GuHCl);
3) addition of EDTA, DTT, and the combination during activity assays, since protease I was most recently determined to be a cysteine protease;
4) addition of divalent metals, individually and in combination during assays (Coolbear *et al.*, Biochem. J. 1992 Oct 15;287 (Pt 2):367-74). These metals included: Ca^{2+}, Co^{2+}, Mg^{2+}, Sr^{2+}, Zn^{2+}, and Mn^{2+};
5) affinity-based purification of the expressed PfpI, during which no binding occurred.

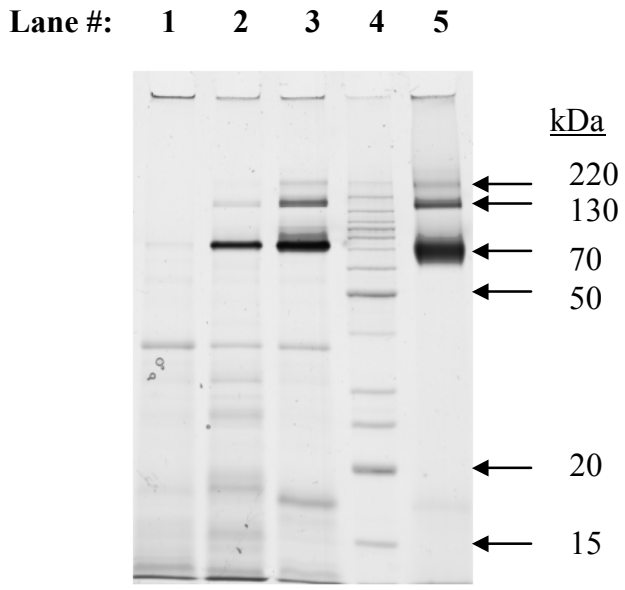


Figure A.1: Expression of recombinant PhpI in *E. coli* on 12% SDS-PAGE.
Expression was carried out after transforming the appropriate vectors into *E. coli* strain BL21(DE3) and growing to optical density (595 nm) of 0.6-0.75. IPTG was added to a final concentration of 0.8 mM and cells (60 ml per sample) were grown for 3.5 hours at 28ºC. Cell pellets were resuspended in 50 mM sodium phosphate buffer pH 7 with 5 mM DTT, sonicated, and supernatant from sonication was heated at 98ºC for 30 min. After spinning, the supernatant from heat treatment was applied to the polyacrylamide gels shown.

Lane 1: Expression product of pET-21b(+) vector without *phpI* gene insert (negative control);
Lane 2: Expression product of pET-21b(+) vector with *phpI* gene insert, but no IPTG added at induction;
Lane 3: Expression product of same vector as in lane 2 with 0.8 mM IPTG added at induction;
Lane 4: Benchmark protein ladder including sizes (from top) 220, 160, 120, 100, 90, 80, 70, 60, 50 (dark band), 40, 30, 25, 20 (dark band), and 15 kDa;
Lane 5: Same sample as in lane 3 with additional 24 hours of heating at 98ºC and spin-down at 14,000 rpm for 30 min (supernatant fraction shown). Note that the same heat treatment was done with the negative control sample in lane 1 and the accompanying SDS-PAGE results showed a blank (protein-free) lane.