BIOCHEMICAL AND PHYSIOLOGICAL CHARACTERIZATION OF CASPASE ACTIVITY IN

HALOARCHAEA

by

MANSHA SETH PASRICHA

A dissertation submitted to the
Graduate School-New Brunswick
Rutgers, The State University of New Jersey
in partial fulfillment of the requirements
For the degree of
Doctor of Philosophy
Graduate Program in Microbiology and Molecular Genetics
Written under the direction of
Kay D. Bidle
Approved by

New Brunswick, New Jersey

JANUARY, 2015
ABSTRACT OF THE DISSERTATION

Biochemical and physiological characterization of caspase activity in Haloarchaea

By Mansha Seth Pasricha

Dissertation Director:
Kay D. Bidle

Caspases, cysteine aspartate-specific proteases, are key initiators and executioners of programmed cell death across a wide array of life. Archaea had been absent from the caspase inheritance discussion due to a notable lack of gene homologues with diagnostic domain signatures. Nonetheless, extremely high, basal caspase-like catalytic activity linked to the cellular stress response was recently demonstrated in the model haloarchaeon, Haloferax volcanii, and shown to be widespread among diverse phyla of archaeal extremophiles. In this dissertation, the catalytic specificity of observed caspase activity in H. volcanii was rigorously tested using hydrolytic assays with a diverse suite of canonical, fluorogenic protease substrates and inhibitors, with model serine and cysteine proteases serving as controls. It was demonstrated that H. volcanii possesses an extremely high level and highly specific caspase-like activity in exponentially growing cells that most closely resembles caspase-4. It is the dominant cellular proteolytic activity, is preferentially inhibited by a pan-caspase inhibitor, and has no cross-reactivity with other known protease families.
Biochemical purification and in situ trapping with biotinylated fmk- and AOMK-based inhibitors, combined with genome-enabled proteomics and structural alignments were collectively used to identify the protein(s) that are associated with this caspase activity. These analyses identified a diverse suite of cellular proteins including thermosomes, proteasomes, a cell division protein, an ATPase (recently identified as an activator of proteasomal degradation), a putative nuclease, a putative aminopeptidase, elongation factor αEF-2, and an ornithine cyclodeaminase as key proteins associated with caspase activity. These findings biochemically connected caspase activity in *H. volcanii* to specific stress-related protein complexes, including those involved in the unfolded protein response (UPR).

A subset of these candidate proteins were targeted for gene knockouts to empirically test their relationship to caspase activity, to assess their physiological roles, and link to UPR through incubations with canavanine. We show that loss of this activity or reduction from a critical threshold level has important consequences to organismal fitness, placing caspase activity in a novel cellular context. Given the deep archaeal roots of eukaryotes, we posit that it evolved as part of a cellular protein-quality control system, ensuring proper production, folding, and degradation.
DEDICATION

This thesis is dedicated

To my parents, Sudesh and Neeta Seth, who sowed the seeds for today to be a reality,

To my children, Mitalee and Viraaj, for being magical, and to help me appreciate and revel at my own biology,

To, most importantly, my husband, Mohit, for his strength, perseverance and unflinching commitment to make my vision his own.
ACKNOWLEDGEMENT

It has taken a village to rear this dissertation and I cherish this opportunity to thank everyone who is so tightly interwoven in this fabric I created.

First of all, I thank Kay and Kelly Bidle, for being partners in crime in every which way during the course of this PhD. Kay has been very encouraging and supportive of the professional transition I represented in the lab. I thank him for his confidence in me, and the independence he gave me to think about my science, and guide and inspire me to tread forward. I thank Kelly for her letting me rear her science “baby” and in standing neck to neck with Kay in her time, patience, and willingness to help my research efforts, despite being physically separated from Rutgers.

I have been extremely lucky to have a diverse, brilliant, and a humble committee. I thank each one of you, Costa, Lee, and Vik, for the advice and feedback I have received, not only in committee meetings, but offline as well. I greatly regard your willingness to let me walk in to your office and discuss any issues I have faced. Additionally I thank Costa for making me a part of his lab family, both scientifically and socially.

I have met some fantastic people by way of Kay. Being in Kay’s lab has been one of the most wonderful experiences of my life. I have made friends who I have shared many an emotion with. I thank Kim for being such a good friend, for her prudence and scientific advice; Liti for the spectrum of life and conversations she brings to the lab; Frank for making sure I had a seamless experimental life and for the hilarious conversations we have shared; Brittany, Christien, Chris Brown, for being such good friends and all others I have missed. I thank Jeremy and Brandon for giving me the opportunity to be their mentor. Suzanne you are thought of at this time.
Paul Falkowski and his lab group have been an integral part of my research life having practically lived half my research life in Paul’s lab. I am so grateful to Paul for encouraging me and listening to my science ideas, and sharing many thoughtful as well as funny conversations with me. I thank Kevin for helping my research efforts despite being Kay’s student, and especially looking out for me when the lab group was out for a 6 week cruise. I also am grateful to Piotr Nawrot and Charlotte Fuller for keeping my equipment in working condition, especially Piotr who was committed even when he was on vacation.

I am lucky to have spent the last 3 and half years at IMCS (now DMCS). I have met some very diverse people and learnt so much science in our oceans. I thank the GPO and DMCS for accepting my participation in the oceanographic world even though I came from microbiology and molecular genetics. I especially thank Oscar Schofield for making me a part of his research group to come to Antarctica. It is such a phenomenal opportunity for me!!

From my past PhD life, I thank Ruth Steward, but for her, I would have not met and loved my current life. I thank my ex-committee member, Kim McKim, and the chair of my previous committee, Mark Gartenberg, for the extremely encouraging and wonderful words they had for me. I thank Kim specifically for helping me find the strength and motivation to move forward. I also thank Andrew Vershon, my department chair, for keeping his faith in me and helping me in my transition from my previous lab to Kay’s lab.

As I transitioned to Kay’s lab, I am grateful for some wonderful people who encouraged me and without their support, this dissertation would not be possible; Dean Barbara Bender, and my friends, Drs. Sangya and Sanjay Varma, Dena Siedel, Morit Shavelsky, and my dear friend Ceren Aralpe.
I have an army of friends and family who have supported me and my family, with their love and help: Ceren, Jeana, Morit, Greg, Chuck, Donato and Patri, Shemaila and Javed, the Khosla’s, the Rocky Hill family, and the Higate Manor family! I am blessed to have an extremely loving family who have supported me every step of the way. I am grateful to my father for treating me like a person and not his princess, and encouraging me to assess and have an opinion ever since I was very little. Both him and my mother are my source of strength. My jump and establishment across the seven seas was possible because of my older sister and her husband, Nidhi and Subodh. Joined by an extremely loving older brother, Anurag and his family, back in India who has been more of a father in all these years and looked out for my social responsibilities in India. I have my husband’s family to thank, particularly his late mother, Shama Pasricha, who made a place for my intellectual status in an extremely conservative family. I additionally thank my husband’s uncle and his wife, Parveen and Prabha Pasricha, and our ex-domestic help, Vaneeta; it is truly their support on the India side of my life that made it possible for me to be in this part of the world.
# TABLE OF CONTENTS

Title page ........................................................................................................................................... i
Abstract ............................................................................................................................................... ii
Dedication ............................................................................................................................................. iv
Acknowledgement ............................................................................................................................ v
Table of Contents .............................................................................................................................. viii
List of Tables ....................................................................................................................................... x
List of Figures ..................................................................................................................................... xi

1. Introduction Title ............................................................................................................................ 1
   1.1. Introduction ............................................................................................................................... 2
   1.2. Tables and Figures .................................................................................................................... 17
   1.3 References .................................................................................................................................. 27

2. Specificity of archaeal caspase activity in the extreme halophile *Haloferax volcanii* .......... 33
   2.1 Published Paper ......................................................................................................................... 34
   2.2 Acknowledgement ..................................................................................................................... 45

3. Isolation and Identification of Proteins Associated with Caspase Activity in
   *Haloferax volcanii* .......................................................................................................................... 46
   3.1. Abstract ....................................................................................................................................... 47
   3.2. Introduction ................................................................................................................................. 48
   3.3. Materials and Methods ............................................................................................................ 50
   3.4. Results and Discussion ............................................................................................................ 60
   3.5. Conclusions ............................................................................................................................... 76
   3.6. Acknowledgement ..................................................................................................................... 78
   3.7. Tables and Figures ..................................................................................................................... 79
   3.8. References ................................................................................................................................... 103
4. Phenotypic characterization of \textit{H.volcanii} mutants for caspase activity and growth under unfolded protein stress ........................................................................................................ 106

4.1. Abstract ........................................................................................................................................................................... 107

4.2. Introduction ........................................................................................................................................................................ 108

4.3. Materials and Methods ......................................................................................................................................................... 110

4.4. Results and Discussion ......................................................................................................................................................... 113

4.5. Future Directions ................................................................................................................................................................. 123

4.6. Acknowledgement ............................................................................................................................................................... 127

4.7. Tables and Figures ................................................................................................................................................................. 128

4.8. References ........................................................................................................................................................................... 136
# LIST OF TABLES

1. **1.1. Analysis of *H. volcanii* proteins containing IETD tetrapeptide caspase cleavage sites** ....17

3. **3.1. Rank order of a subset of *H. volcanii* proteins recovered in a biotinylated fmk-based inhibitor trapping and identified by LC MS/MS analysis based on the sum total of spectral counts of unique peptides identified** .................................................................79

3.2. **Rank order of a subset of *H. volcanii* proteins recovered in a biotinylated AOMK-based inhibitor trapping and identified by LC MS/MS analysis based on the sum total of spectral counts of unique peptides identified** .................................................................80

3.3. **Skyline quantification analysis of *H. volcanii* proteins recovered in biotinylated AOMK-based inhibitor trapping** .................................................................................................................................................................81

3.4. **Rank order of the top twenty-five structural alignments for predicted tertiary structures of *H. volcanii* proteins to the crystal structure of caspase-6 inhibitor-bound active site (VEID-CHO; 3P4U) via TopMatch and Cytoscape analysis** .................................................................84

4. **4.1. Primers used in the creation of mutant strains** .................................................................................................................................128

4.2. **List of strains and plasmids created and used in this work** ..................................................................................................................129
# LIST OF FIGURES

1. **1.1.** A depiction of caspase activation and recognition of caspase-specific substrates and their mode of detection by spectrofluorometry .............................................................. 18

2. **1.2.** Crystal structure of human caspase-8 ............................................................................. 19

3. **1.3.** Phylogenetic analysis of caspase-family of proteases ..................................................... 20

4. **1.4.** High caspase-specific activity in H. volcanii ..................................................................... 21

5. **1.5.** *H. volcanii* has a high caspase-like activity at different salinities and phases of growth .22

6. **1.6.** Effect of caspase inhibition on cell growth and in vivo caspase-specific activity during incubation in low (1.5 M), optimal (2.1 M), and high (3.5 M) salt .............................................. 23

7. **1.7.** Caspase activity and protein expression in representatives from archaeal and bacterial domains .............................................................................................................. 24

8. **1.8.** Original characterization of the model haloarchaeon, *Haloferax volcanii* .................... 25

9. **1.9.** *H. volcanii* has a tractable genetic system ................................................................. 26

2. **2.1.** Substrate specificity of *H. volcanii* cell extracts and various purified proteases ............. 36

2. **2.2.** Caspase-specific activities in cell extracts from *H. volcanii* grown in low (1.5 M), optimal (2.1 M), and high (3.5 M) NaCl concentrations and harvested at mid-exponential phase ........................................................................................................ 37

3. **2.3.** Selected inhibition of caspase activity in *H. volcanii* cell extracts ............................... 38

2. **2.4.** Efficacy of specific fluoromethylketone inhibitors at abolishing caspase-specific activity ............................................................................................................................... 39

2. **2.5.** Caspase catalytic activity in *H. volcanii* is stimulated by magnesium ......................... 40

S1. **S1.** Panel of proteolytic activity in partially purified protein fractions ................................. 43

S2. **S2.** Protease inhibitor efficacy on model protease substrate .............................................. 44

3. **3.1.** Caspase specific activity during initial purification with ammonium sulfate .................. 85

3. **3.2.** Visualization of total protein profiles during initial AmSO₄ purification ......................... 86

3. **3.3.** Visualization of proteins fractionated by size exclusion chromatography from *H. volcanii* cleared cell extract .............................................................................................. 87
3.4. Test of efficacy of fmk-based inhibitors on caspase-specific activity in *H. volcanii* ..........88
3.5. Efficacy of fmk-based inhibitors on human caspase-8 activity ......................................89
3.6. Visualization of b-VAD-fmk-based inhibitor trapping of human caspase-8 .........................90
3.7. Efficacy of biotinylated-fmk-based inhibitors on caspase-specific activity in pre-cleared
*H. volcanii* AmSO₄ cell extracts prior to inhibitor trapping with streptavidin beads ..........91
volcanii* cell extracts ...........................................................................................................92
3.9. Visualization of b-FA-fmk-based inhibitor trapping of non-caspase-associated proteins in
*H. volcanii* cell extracts ...................................................................................................93
3.11. Efficacy of AOMK-based inhibitors on the activity of human caspase-8 .........................95
3.12. Efficacy of AOMK-based inhibitors on caspase activity of *H. volcanii* cell extracts .......96
3.13. Efficacy of b-VAD-AOMK on caspase activity of *H. volcanii* cell extracts ......................97
3.14. Visualization of AOMK-based inhibitor trapping of proteins from *H. volcanii* cell extracts
...........................................................................................................................................98
3.15. Efficacy of b-FA-fmk inhibitor on the hydrolysis of LLBY-AMC activity in the *H. volcanii
80% and 100% AmSO₄ salt precipitated samples ....................................................................99
3.16. Visualization of structural alignments of select *H. volcanii* candidate proteins with the
crystal structures of human caspases ....................................................................................100
3.17. Structural alignment comparison of inhibitor-trapped *H. volcanii* proteins with the
known modeled, active site tertiary structure of caspases ....................................................101
3.18. Gel visualization of caspase-active proteins by incubation with a fluorescently labeled
AOMK-based inhibitor ............................................................................................................102

4 4.1. Examination of caspase activity in all mutant strains ......................................................130
4.2. Growth of H26 in Hv minimal and canavanine stress ......................................................131
4.3. Impact growth conditions on caspase and chymotrypsin activity in *H. volcanii* ............132
4.4. Impact of various growth conditions in MSP1 .....................................................................133
4.5. Impact of various growth conditions on thermosome mutant strains ..............................134
4.6. Impact of various growth conditions on proteasome mutants ..........................................135
CHAPTER 1

INTRODUCTION
Section 1.1  

**INTRODUCTION**

Cysteine aspartate proteases (caspases)

Cysteine aspartate proteases, or caspases, are a conserved family of highly specific and refined cysteine proteases that cleave target proteins at specific tetrapeptide motifs (e.g., IETD, LEVD, VEID, WEHD) after the aspartate residue in the P1 position (Figure 1.1). They are among the most specific proteases in biology with the different substrate specificities of different caspases resting in their recognition of different amino acids in the P2-P4 position. Caspases are key initiators and executioners of programmed cell death (PCD), an irreversible form of cell death (1, 2), both in unicellular microbes and higher eukaryotes, whereby their activity cleaves a variety of cellular proteins that post-translationally activate the PCD molecular machinery, disassemble the cell, and irreversibly commit it to die (3-7). Caspases are produced as zymogens and undergo proteolytic cleavage for activation; their high specificity is consistent with the fact that PCD and apoptosis is not accompanied by indiscrete protein digestion but site-specific cleavage.

PCD was first discovered in multicellular organisms that use genetically controlled mechanisms of cell death for development, morphogenesis, and defense (1, 8, 9). Manipulations and repression of apoptosis is associated with human pathologies like cancer, neurodegenerative diseases, and autoimmune disorders (3), so its activity is essential to proper cell and organismal function. Once thought to be a proprietary feature of metazoans, PCD, and its core molecular machinery, is now recognized as a ubiquitous cellular trait throughout
nature, having been found in a diverse range of prokaryotes and unicellular eukaryotes (10-13). The discovery of the PCD machinery in unicellular organisms is particularly intriguing as it likely represents an initial ancestral core of executioners that led to the emergence of the cell death machinery (11, 12).

The first caspase, Interleukin 1β converting enzyme (ICE) was discovered and cloned from human blood monocytes (14, 15). Their role in PCD and apoptosis was first elucidated in the nematode, Caenorhabditis elegans as the products of two genes, ced-3 and ced-4, were implicated in initiation and execution of cell death (16). Since that time, approximately 14 distinct ‘classic’ caspases have been found in metazoan cells, all possessing a conserved domain structure with a histidine- and cysteine-containing catalytic dyad, with the cysteine residue as the catalytic nucleophile. Activated caspases are heterodimers containing two subunits, a large 17-20 kDa subunit and a small 10-12 kDa subunit comprising a domain. Each caspase is referred to as a “homodimer of a heterodimer”. Each domain comprises twisted and mostly parallel β-sheets flanked by two layers of α-helices. The heterodimer comprises the two domains that are formed by the alignment of the β6 strand at the C-terminal end in an antiparallel manner (Figure 1.2) (17-19). Two heterodimers come together to form a functional caspase and hence the term “homodimer of a heterodimer”.

There has been an intense interest in the origin of eukaryotic apoptotic systems as part of an emerging and ongoing discussion on the evolutionary origin(s) and cellular roles of PCD and its varied molecular machinery. Research in the last decade or so has provided evidence of the ubiquity of PCD, catalyzed by caspases, or caspase-orthologous proteins-metacaspases, paracaspases, phytaspases, and saspases-in metazoans, plants, protozoa, fungi, and a few
bacteria (3, 20-28) (Figure 1.3). Saspases and phytaspases are functional homologs of caspases found in plants, with specificity for caspase tetrapeptide substrates and participation in PCD, but no structural homology to caspases (20, 26, 27). Paracaspases, found in slime molds and animals, have structural homology to caspases, are not known to be involved in PCD and do not recognize tetrapeptide caspase substrates (29). The discovery of metacaspases, in morphologically diverse organisms such as plants, fungi, unicellular protozoa and a few bacterial species (30) suggests that they may represent an initial, ancestral core of executioners that led to the emergence of the cell death machinery. Two types of metacaspases have been defined in plants and fungi (Type I and II); *Arabidopsis* harbors nine genes for these metacaspases. Together with the yeast metacaspase (Mca1) gene, these proteases are caspase orthologs with an arginine/lysine endopeptidase activity (31). While the activity and cellular roles of metacaspases still remain an open question, they are part of an emerging discussion about the origins of eukaryotic apoptotic machinery. They are particularly abundant in the genomes of bacteria that employ cell differentiation processes in their complex life cycles (e.g., Actinomycetes, Cyanobacteria and α-Proteobacteria) and their presence in α-Proteobacteria genomes points to a potential bacterial origin, with their subsequent establishment in eukaryotes occurring after a mitochondrial endosymbiotic event and transfer to the nuclear genome (32, 33). The presence of these genes in cyanobacteria, the progenitors of the chloroplast, argue for an alternative endosymbiotic descent in eukaryotic phytoplankton and plant lineages. Indeed, metacaspases (along with other putative PCD-related proteins) are widespread among prokaryotic and eukaryotic phytoplankton genomes (34). Moreover, morphological and biochemical characteristics that are consistent with caspase-mediated PCD
have been found in widely diverse evolutionary lineages of prokaryotic and eukaryotic phytoplankton, including cyanobacteria, chlorophytes, coccolithophores, diatoms, and dinoflagellates (35-40), indicating that PCD is widespread among ‘lower’ taxonomic groups. However, a strong phylogenetic divergence suggests essential evolutionary and ecological roles for PCD in aquatic microbes but the processes that selected for its retention over a microbial three billion year evolutionary history remain unknown (34, 41, 42).

**Archaea and caspases**

Archaea comprise the second major prokaryotic domain of life (43, 44). Since their distinct and separate classification from bacteria, the ubiquity and diversity of the Archaea has been shown in the metabolically diverse extremophiles, those microbes that have adapted to and thrive in adverse ecological niches, normally uninhabitable by most living things (e.g., halophiles, hyperthermophiles), as well as in non-extreme (e.g., non-cultivated soil and oceanic species) representatives [reviewed in (45)]. From an evolutionary perspective, the study of the Archaea is of great relevance, as their origin and phylogenetic placement compared with eukaryotes and bacteria remains controversial and unresolved (46). Phenotypically, several fundamental molecular features of the Archaea share a great deal in common with eukaryotes (e.g., replication, transcription, and translation machinery) (47-52) though their size, gene organization (e.g., operons), and circular chromosomes align more closely with bacteria (11, 12, 25, 53). Unique ether-linked lipid membrane composition and, for some, the capability to perform methanogenesis, remain exclusive hallmarks of Archaea.

Given the deep archaeal roots of eukaryotes (54) it was of considerable interest to investigate whether caspase-like activity can be found in Archaea (55). Archaea have historically
been absent from the discussion on the evolutionary origin of PCD likely due to a lack of recognizable genetic homologues to key apoptotic genes, which do have representation in a few bacterial genomes (33). Bidle et al. (53) searched the annotated Haloferax volcanii 4.01 Mb genome sequence (http://archaea.ucsc.edu/cgi-bin/hgGateway?db=haloVolc1) for homologues to caspases (Homo sapiens caspase 8, XM_054989), paracaspases (Caenorhabditis elegans, AAG38591) and metacaspases (Saccharomyces cerevisiae, NP_014840; the cyanobacterium Trichodesmium erythraeum, ZP_00074337; the α-proteobacterium Rhodobacter sphaeroides, ZP_00005822) all of which contain a conserved caspase domain structure (COG4249). Using the BLAT analysis server provided by the H. volcanii genome website, there was no detectable sequences among the 4,074 annotated H. volcanii proteins with similarity to this representative subset of proteins in the caspase superfamily. Likewise, a survey of the annotated Pfam domains yielded no proteins containing the PF00656 Peptidase_C14 pfam domain, which is characteristic of the caspase family of proteases. Similarly, a tBLASTn search of the annotated genomes of closely related halophilic euryarchaea, Halogeometricum borinquense DSM 11551 and Halorubrum lacusprofundi ATCC 49239 (http://www.jgi.doe.gov/genome-projects/), yielded no sequence homologues with conserved caspase domain structure. These observations corroborate previous bioinformatics analyses (33) and suggest that, at the level of bioinformatics and sequence similarity, these proteins are absent from haloarchaeal genomes. The incidence and roles of caspase-like proteins are now beginning to be recognized in this third domain of life. Bidle et al. (56) first demonstrated the presence of a extremely high basal caspase-8 like activity in the model haloarchaeon, Haloferax volcanii, despite a lack of sequence homologs to the conserved caspase domain structure. Instead, H. volcanii caspase
activity was measured using its diagnostic catalytic activity and hydrolysis of a canonical, caspase 8-specific substrate (IETD-AFC). Notably, this catalytic activity was found tuned and adapted to the optimal environmental conditions for *H. volcanii* growth (53) at high salt (1.5M NaCl) and temperature (42°C) (Figure 1.4). This activity was more than an order of magnitude higher in actively growing cells, and dwarfs these rates in other unicellular microbes, putting them in a unique catalytic and physiological context.

To better define a potential cellular role for the observed caspase activity, Bidle et al. (56) tested *H. volcanii* cell extracts from cultures grown in differing salinities and stages of growth. High basal caspase specific activity was detected in exponentially growing cells under optimal growth conditions, as well as under different environmental conditions or stages of growth, suggesting that it may be essential for growth and may serve a housekeeping function. Activity was also slightly elevated under suboptimal low and high salt stress (Figure 1.5), indicative that it may also play a role cellular salt stress responses. Successful *in vivo* inhibition of cellular caspase activity by the caspase specific inhibitor, z-VAD-fmk did not affect growth under optimal conditions, but noticeably inhibited growth under suboptimal low (1.5 M) and high (3.5 M) salt conditions (53) (Figure 1.6). Together with the elevated activities observed at these salt levels, these results strongly suggested that caspase activity was an important component of the cellular salt stress response.

Given the very high, basal caspase specific activity, as well as its induction under salt stress and death, Bidle et al. (53) further tested whether expressed proteins in *H. volcanii* extracts displayed immunoreactivity to an affinity purified, polyclonal antibody raised against recombinant human caspase-8. Indeed, western analyses detected several distinct proteins in
*H. volcanii* that cross-reacted with the caspase-8 antibody, most notably \(\sim 37\) and \(\sim 100\) kilodalton (kDa) proteins that appeared to be differentially expressed at different stages of growth and under different environmental conditions. The immunoreactive \(\sim 37\) kDa protein was expressed only in mid-exponentially growing cells and appeared to be more abundant with increasing salt concentrations. In contrast, an immunoreactive \(\sim 100\) kDa protein was expressed in both growing and dying cells. It was especially prevalent in dying cells under optimal salinity. In contrast, it was largely absent under high salt conditions (53).

To test for a potential link with caspase activity results, Bidle et al also tested whether treatment of optimally growing *H. volcanii* cells with z-VAD-FMK influenced the expression of immunoreactive proteins on Western blots. This treatment abolished *in vivo* caspase-like activity after 24 h. Western analyses of inhibited cultures revealed a similar dramatic reduction in both the \(\sim 37\) kDa and \(\sim 100\) kDa proteins normally found in cultures expressing caspase activity. While the \(\sim 100\) kDa proteins were prominently expressed over a four day period in the presence of DMSO, they were largely absent in the presence of z-VAD-FMK. Likewise, the 37 kDa protein virtually disappeared after Day 1. While this data show similar responses for the observed caspase activity and the expression of caspase-8 immunoreactive proteins (i.e., both are significantly reduced), it is unknown whether they represent the same proteins and, hence, whether they can be justifiably linked. Nonetheless, they likely play a role in a similar pathway such that caspase-like protein expression was strongly repressed by the presence the broad-spectrum caspase inhibitor. Inhibition of caspase-like activity may have a negative affect on an upstream regulator of protein expression and/or a positive affect on a downstream regulator of protein turnover.
To determine if other diverse Archaea possessed similar caspase activity and immunohybridizing proteins to human caspase-8, a number of species from diverse archaeal phyla within the Euryarchaeota and Crenarchaeota were examined. Among the organisms studied were two different genera of haloarchaea, *Halorubrum* and *Haloarcula*, as well as a methanogen (*Methanosarcina acetivorans*), a hyperthermophile (*Pyrococcus furiosus*), and an acidophile (*Sulfolobus solfataricus*). A number of diverse bacteria were also examined, including two bacterial halophiles (*Salicola marasensis* and *Alkalibacillus* sp.), *Escherichia coli* and *Bacillus cereus*. All organisms were tested for caspase activity via IETD-AFC hydrolysis, and protein expression via Western blots (Figure 1.7). No detectable activity or expression of immunoreactive proteins was found in the bacteria, but all Archaea showed either high levels of basal caspase activity and/or immunoreactivity to human caspase-8 antisera in exponentially growing cells. Hydrolysis rates were at a similar order of magnitude to the salt-adapted activity in haloarchaea. While *P. furiosus* and *S. solfataricus* showed little to no immunoreactivity, they possessed very high, basal caspase specific activity when assayed near their optimal growth temperatures (98°C and 65°C, respectively), indicative of temperature-adapted enzymes. It was notable that no measurable activity was detected for the strict anaerobe *M. acetivorans*, probably due to the aerobic assay conditions (Z. Kellman and K. Sowers, personal communication). Nonetheless, it is noteworthy that *M. acetivorans* showed relatively strong expression of caspase-8 immunoreactive proteins, paralleling that seen for the haloarchaea. These results indicated that caspase-like activity and expression might be a widespread and characteristic trait of the archaeal domain. Unfortunately, the genetic signatures and identity of the responsible proteins are still lacking. This is interesting given Archaea have whole families of
uncharacterized peptidases; Euryarcheota alone have 49 known peptidase species and 68 unassigned sets (57).

**Putative caspase substrates in the H. volcanii proteome**

The canonical, target tetrapeptide cleavage recognition sequence for caspase-8 is IETD (Ile-Glu-Thr-Asp). Given our initial observations of very high, basal caspase-8-like specific activity, both under optimal and salt stress conditions, the annotated H. volcanii proteome was searched for proteins that contain this amino acid sequence and could serve as potential proteolytic targets of activated caspases. Eighteen potential target proteins were identified from an in silico analysis of the 4,063 total annotated proteins as having the site-specific, IETD signature sequence (Table 1.1), including one protein encoded on the 635.7 kb megaplasmid (53). Based on functional analysis, they play diverse roles in cell physiology like central energy metabolism, cell membrane biosynthesis, chaperones, signal transduction, nucleases, and transcription, with several having intriguing, putative roles in osmotic stress, allosteric regulation, and cell communication. For example, one protein (HVO_0600) contains an ACT domain, which is a structural motif and one of several intracellular small molecule-binding domains that function to control metabolism, solute transport, allosteric regulation, and signal transduction (58). A putative secreted protein (HVO_1184) shows strong homology to the ‘late embryogenesis abundant proteins’ (IPR004238) that, although of unknown function, are expressed at different stages of late embryogenesis in higher plant seed embryos and under conditions of dehydration stress (59). Furthermore, a ‘conditioned medium-induced protein 4’ (HVO_0219) has even been linked with evidence of quorum sensing in H. volcanii (M. Mevarech, personal communication). Added to this repertoire are an additional 53 putative
substrate proteins with LEVD and VEID signature motifs [unpublished data]; together with the much higher caspase-4 like and caspase-6 like activities in our recent experiments [(60); also chapter 2 of this dissertation]. Given the diverse array of known cellular caspase substrates (61), it is not surprising that potential target proteins in \textit{H. volcanii} represent diverse aspects of cell physiology. Cleavage and post-translational, proteolytic processing of this subset of proteins by caspase-like activity appears to be a critical component of a regulated yet multifaceted, salt stress response.

\textbf{Halofex volcanii as a model system}

\textit{Halofex volcanii} is an ideal model system for these studies. It is a moderate halophilic archaeon that grows at NaCl concentrations ranging from 1.5- 3.5 M (62). It was first discovered when Benjamin Elazari-Volcani isolated the first aquatic “orange-pigment producer”, a gram-negative microorganism that was 1.6 x 1.6 \textmu m, with remarkably high salt tolerance of 28-29% (63) (Fig 1.8 A). Volcani described many bacteria that grew in the Dead Sea and one of the most interesting microorganisms (64, 65) was eponymously named \textit{Halobacterium volcanii} [later classified and renamed \textit{H. volcanii}; (65)]. It was of interest due to both its unusual pleomorphism and relatively “low” salt requirement, as compared with other haloarchaea.

The Dead Sea is a fascinating ecosystem to study microflora and their ecophysiology, as it has the highest salt concentration of all natural lakes and high concentrations of divalent cations (e.g. \textit{Mg}^{2+} and \textit{Ca}^{2+}), which can be inhibitory to most living cells. \textit{H. volcanii} was collected from sediments at 1m depth of water in the northern end of Dead Sea (62). It required about 2.6 M NaCl (15\% w/v) for growth, with optimal NaCl being 1.7 M at a temperature of 30\(^\circ\)C, and 2.5 M at 40\(^\circ\)C (Figure 1.8 B)(62, 66). Importantly, \textit{H. volcanii} also has a
requirement for high levels of Mg, likely reflective of the abundant MgCl₂ concentrations found in the Dead Sea, which constitute 50.8% of the anhydrous chlorides on a weight percentage basis, compared with 30.4% for NaCl (67). This is remarkably different from the composition of ocean seawater which is ~97% NaCl.

*H. volcanii* is considered a model haloarcheon for several reasons. **First**, it is ecologically relevant having been isolated from the Dead Sea environment and found in various solar saltern environments (68). **Second**, its ease of growth in the laboratory in both batch liquid culture and on solid media under optimal (2.1 M) and suboptimal (1.5 M and 3 M) salinities (48, 69-73) makes it tractable to study osmoregulation and salt stress responses. Rodriguez-Valera et al. (66) demonstrated that these chemo-organotrophs could use defined inorganic media with a single carbon source such as glucose and pyruvate. **Third**, its genome has been completed (2.9 Mbp chromosome and 4 megaplasmds ranging in size from 635.7, 437.9, 85.1, to 6.4 kb; [http://archaea.ucsc.edu/](http://archaea.ucsc.edu/)) (74-77). It encodes 4,063 proteins, 226 of which are hypothetical proteins and have no BLASTP matches to proteins of known function. **Fourth**, it has a well-established genetic system and availability of selectable markers (e.g., *pyrE*, orotate phosphoribosyltransferase) so that targeted gene knockouts can be routinely generated using plasmid based transformation systems (76, 78-80) (Figure 1.9). This is in contrast to another halophile, *Halobacterium salinarum*, which has an unstable genome due to frequent transpositions (81).

**Co-evolution of PCD biochemical machinery with key metabolic pathways**

The ecological role(s) of PCD-related genes (such as caspase-like proteases) in unicellular organisms and the evolutionary drivers selecting for their retention remain unknown. Their
retention in diverse microbes with ancient evolutionary histories strongly supports a physiological benefit even though their preservation seems to impart a negative selection pressure. Bidle and Falkowski discuss several hypotheses that have been put forth as potential evolutionary drivers in unicellular organisms, revolving around the beneficial balance between community and individual (34) for example, as an altruistic adaptation in heterotrophic bacteria and yeast, designed to benefit a population (82-85), by eliminating damaged cells from a population, and hence ensuring sufficient nutrient supply of organic compounds to surviving cells. This mechanism has also been proposed to eliminate individual cells damaged by exposure to bactericidal agents like antibiotics (84). At the same time, maintenance of PCD in unicellular organisms may actually convey a benefit to the individual in the form of increased inclusive fitness.

According to the Kin selection model (86), “a behavior can increase in frequency if it favors the fitness of genes, regardless of whether the genes are in an individual or in closely related kin. Since the genetic material within an individual is equal to the sum of shared alleles within related individuals, behavior that benefits the kin of the individual may enhance its inclusive fitness (the sum of the reproductive success of the individual and its relatives), essentially ensuring genetic longevity”.

This strategy would clearly be effective for clonal populations of microorganisms that reproduce asexually, such as Archaea, and is relevant for environmental microbes grown in culture. The ecological relevance of this model is unclear in mixed natural populations, as release of cellular nutrients into the water would seem just as likely to benefit competing species as closely related or clonal individuals. Nonetheless, it is well accepted now quorum sensing is the mode of communication in bacteria to exhibit social behavior, often living in large, complex and organized communities (87). It is currently unknown whether Archaea ‘communicate’ in a similar manner.
On the other hand, our data documenting extremely high caspase activity and expression in exponentially growing *H. volcanii* cells with an essential role in adaptation to salt stress strongly support a hypothesis that the PCD biochemical machinery co-evolved with other metabolic pathways—one where retention and expression served housekeeping and/or regulatory functions. Our findings provide a role for caspase-like proteases in normal cell function, suggesting the potential biological functions of caspase-like proteases beyond apoptosis and cell death (88). It was recently shown that *Aspergillus fumigatus* metacaspases CasA and CasB facilitated growth under physiological stress (89), rather than elicited cell death. Similar observations have been seen in the diatom *T. pseudonana*, where distinct metacaspases are expressed differently in response to Fe availability raising questions about their importance in the ecological success of diatoms (39). Similarly yeast metacaspase, Mca1, was recently suggested to be involved in preventing protein aggregates in *Saccharomyces cerevisiae*, and facilitates the removal of unfolded proteins in a protein quality control pathway of the cell (90). These questions ultimately have critical implications for the roles of these enzymes in influencing the fate of cells in natural environments.

**Goals of this thesis**

It is now evident that caspase-like proteins are indeed represented in all domains of life and a widespread characteristic of unicellular organisms. The extremely high basal caspase activity in haloarchaea together with dramatic physiological impairment when inhibited under salt stress is noteworthy given the lack of caspase genetic homologues in archaeal genomes. The identity of proteins, which encode for caspase-like protease activity and do not share conserved caspase domain structures at the sequence level, is currently unknown in Archaea
and will be the focus of future studies. It is intriguing to speculate that they represent a unique evolutionary lineage of these enzymes. Given the deep archaeal roots of eukaryotes (55), it is possible that these archaeal enzymes have played a unique and heretofore unappreciated role in their establishment and evolution in higher eukaryotes.

The overarching hypothesis of this thesis is that caspase-like proteins may have emerged in early evolving microbes like Archaea, for essential housekeeping functions by participating in the cellular stress responses, such as activation of the unfolded protein response (UPR). I further posit that it served as a foundation for stress and PCD pathways in higher organisms. Proposed work employs a suite of biochemistry, molecular genetics, and physiological approaches to address this hypothesis by 1) characterizing the catalytic activity and specificity of the extremely high caspase activity in H. volcanii, 2) isolating and identifying the protein(s) associated with the extremely high caspase-like activity, and 3) elucidating its cellular function(s) and possible participation in UPR.

Chapter 2, which has already been published (60), establishes the strict specificity of the caspase activity and firmly roots caspase activity as the predominant proteolytic activity in H volcanii cells. Through a detailed analysis of substrate specificity to four model proteases, (trypsin, cathepsin, papain, and recombinant human caspase-8), caspase activity in H. volcanii was reported to be highly specific and preferential to the canonical caspase-4 substrate (LEVD-AMC). This specificity was further verified by challenging cell extracts with a panel of protease inhibitors that target a diverse array of proteases including serine proteases, cysteine proteases, metalloproteases, and caspases. Results demonstrate that z-VADfmk, a pan caspase inhibitor, specifically inhibits H. volcanii caspase activity.
Armed with this diagnostic catalytic specificity, Chapter 3 uses standard protein purification techniques coupled with genome-enabled proteomics to isolate and identify candidate proteins associated with the high, basal caspase activity in *H. volcanii* cell extracts. It specifically employs *in situ* trapping using biotinylated fmk- and AOMK-based inhibitors combined with genome-enabled proteomics and structural alignments to identify candidate proteins associated. Using *in situ* trapping and ‘structure mining’ approaches, this work identified a suite of diverse proteins that have a similar active site fold like caspases, including thermosomes, chaperonins, proteasomes, cell division protein 48 (Cdc48), an ATPase, recently identified as an activator of proteasomal degradation, and more prominently, a putative nuclease and putative aminopeptidase.

Genes encoding these proteins served as promising candidates for knock out experiments in Chapter 4 in order to more directly assess a link with caspase activity and provide a more resolved cellular context. A collection of knockout mutants were either obtained or generated *de novo* for candidate genes that either had some putative role in cell stress pathways and/or were among the top structural homologues in the active site fold. Significant losses in caspase activity were observed for cdc48, proteasome, and thermosome mutants. Furthermore, impaired growth in the presence of canavanine, a chemical inducer of protein folding stress, indicated a role for archaeal caspase-activity in cellular responses to proteotoxicity. Taken together, this work strongly suggests that the caspase-like activity functions as part of UPR as part of complex interactions with proteasome and thermosome function.
Section 1.2 TABLES AND FIGURES

Table 1.1. Analysis of *H. volcanii* proteins containing IETD tetrapeptide caspase cleavage sites.

The table presents likely substrate proteins for caspase-like proteases in the *H. volcanii* proteome. *H. volcanii* has an additional 53 proteins in the genome with LEVD (caspase-4), and VEID (caspase-6) tetrapeptide motifs (unpublished data). Adapted from Bidle et al. (53).

<table>
<thead>
<tr>
<th>Protein ID</th>
<th>Category</th>
<th>Putative function (organism; accession number; e-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HVO_0707</td>
<td>Central metabolism</td>
<td>Rieske iron sulfur protein-like protein 1 (<em>Natrachloris magadii</em>, ZP_03693093; 9E-34)</td>
</tr>
<tr>
<td>HVO_1657</td>
<td></td>
<td>Phosphoribosylamine/glycine ligase (<em>Halorubrum lacusprofundi</em>, YP_0025665957; 0.0)</td>
</tr>
<tr>
<td>HVO_1711</td>
<td></td>
<td>Glycosyl hydrolase, glucamylase (<em>Halogeometricum borinquense</em>, ZP_04000277; 0.0)</td>
</tr>
<tr>
<td>HVO_2109</td>
<td></td>
<td>Pentulose/hexulose kinase (<em>Thermobaculum terrenum</em>, ZP_03858929; 9e-32)</td>
</tr>
<tr>
<td>HVO_2244</td>
<td></td>
<td>2-Keto-4-pentenoate hydratase (<em>Halogeometricum borinquense</em>, ZP_03999158; 6e-99)</td>
</tr>
<tr>
<td>HVO_2340</td>
<td></td>
<td>Protoporphyrogen oxidase (<em>Halogeometricum borinquense</em>, ZP_03999228; 0.0)</td>
</tr>
<tr>
<td>HVO_2808</td>
<td></td>
<td>Succinate dehydrogenase subunit A (<em>Halogeometricum borinquense</em>, ZP_03999753; 0.0)</td>
</tr>
<tr>
<td>HVO_0702</td>
<td>Cell membrane</td>
<td>Uncharacterized membrane protein (<em>Halogeometricum borinquense</em>, ZP_03998039; 0.0)</td>
</tr>
<tr>
<td>HVO_1412</td>
<td></td>
<td>Diphosphomevalonate decarboxylase (<em>Halogeometricum borinquense</em>, ZP_03997342; 1e-149)</td>
</tr>
<tr>
<td>HVO_0459</td>
<td>Nuclease</td>
<td>Endoribonuclease L-PSP (<em>Halogeometricum borinquense</em>, ZP_03998257; 3e-54)</td>
</tr>
<tr>
<td>HVO_A0069†</td>
<td></td>
<td>Hypothetical protein OES289R, predicted nuclease, RecB family (<em>Halobacterium salinarum</em>, YP_001690727; 9e-12)</td>
</tr>
<tr>
<td>HVO_0651</td>
<td>Chaperone</td>
<td>Prefoldin, beta subunit, archaeal (<em>Halogeometricum borinquense</em>, ZP_03998084; 9e-39)</td>
</tr>
<tr>
<td>HVO_1233</td>
<td></td>
<td>Cold-shock DNA-binding domain protein (<em>Halorubrum lacusprofundi</em>, YP_002566634; 3E-23)</td>
</tr>
<tr>
<td>HVO_1184</td>
<td>Unknown function</td>
<td>Conserved secreted protein (<em>Halogeometricum borinquense</em>, ZP_03997584; 5e-96)</td>
</tr>
<tr>
<td>HVO_2673</td>
<td></td>
<td>Hypothetical protein rmAC3092 (<em>Haloarcula marismortui</em>, YP_137512; 9e-36)</td>
</tr>
<tr>
<td>HVO_0875</td>
<td>Transcription</td>
<td>RNA polymerase beta subunit (<em>Halorubrum lacusprofundi</em>, YP_002564783; 0.0)</td>
</tr>
<tr>
<td>HVO_0600</td>
<td>Signal transduction</td>
<td>Amino acid-binding ACT domain protein (<em>Haloarcula marismortui</em>, YP._136929; 4e-64)</td>
</tr>
<tr>
<td>HVO_0219</td>
<td>Cell communication</td>
<td>Conditioned medium-induced protein 4 (<em>Halofex volcanii</em>, AAL35836; 1e-103)</td>
</tr>
</tbody>
</table>

†. Based on genome annotation (http://archaea.ucsc.edu/cgi-bin/hgGateway?db=haloVolc1) and best BLASTP hit (http://blast.ncbi.nlm.nih.gov/).

b. Protein encoded on the 635.7 kb megaplasmid.
**Figure 1.1.** A depiction of caspase activation and recognition of caspase-specific substrates and their mode of detection by spectrofluorometry. After protease activation, the activated caspases recognize their respective substrates, which are covalently linked to the fluorogenic dye, 7-amino-4- methyl coumarin (AMC). Upon cleavage by the respective caspase, the free fluorescent dye can be detected using a plate reader with a 380 nm Excitation and 460 nm Emission filter. Adapted from Takara Clontech (http://www.takara.co.kr).
**Figure 1.2. Crystal structure of human caspase-8.** This figure (pdb 1QTN) exemplifies the fundamental caspase fold (a) followed by a simplified topological model of caspase structure (b). Both figures indicate two domains. Each domain comprises the heterodimer of the large and small subunit constituting. Together the two domains represent a twisted 12-stranded β-sheet sandwiched by α-helices. Adapted from Fuentes-Prior and Salvesen (17).
**Figure 1.3.** Phylogenetic analysis of caspase-family of proteases. Caspase family tree comprises homologs and orthologs in the eukaryotic and bacterial domain, but there is no representation of caspases, metacaspases, or paracaspases in Archaea. Metacaspases from both the "red" (*T. pseudonana* and *E. huxleyi*; colored red) and "green" (*C. reinhardtii*; colored green) eukaryotic, phytoplankton lineages tightly cluster within a group of metacaspases including unicellular protists such as fungi, and trypanosomes, as well as with higher plants. Metacaspase-like proteases in cyanobacteria (blue) display considerable diversity and form several clusters independent of eukaryotic lineages. Purple non-sulfur photosynthetic bacteria are colored purple. Abbreviations: Csp, caspase; PC, paracaspase; MC, metacaspase; HP, hypothetical protein; WD, WD-40 containing protein; *Hs*, *Homo sapiens*; *Dr*, *Danio rerio*; XI, *Xenopus laevis*; Dm, *Drosophila melanogaster*; Ce, *Caenorhabditis elegans*; Hv, *Hydra vulgaris*; Dd, *Dictyostelium discoideum*; Rp, *Rhodopseudomonas palustris*; Bj, *Bradyrhizobium japonicum*; MI, *Mesorhizobium loti*; Av, *Anabaena variabilis*; Np, *Nostoc punctiforme*; Te, *Trichodesmium erythraeum*; Gv, *Gloeobacter violaceus*; Sar, Sargasso Sea; At, *Arabidopsis thaliana*; Os, *Oryza sativa*; Cr, *Chlamydomonas reinhardtii*; Tp, *Thalassiosira pseudonana*; Eh, *Emiliania huxleyi*; An, *Aspergillus nidulans*; Sc, *Saccharomyces cerevisiae*; Sp, *Schizosaccharomyces pombe*; Tb, *Trypanosoma brucei*; Ca, *Chlorochromatium aggregatum*; Rs, *Rhodobacter sphaeroides*; Rr, *Rhodospirillum rubrum*; Syn, *Synechocystis* sp.; The, *Thermosynechococcus elongatus*. Scale bar indicates base pair substitutions per nucleotide position. The general tree topology is similar to a previous analysis of the caspase family (33), which did not include phytoplankton representatives. Adapted from Bidle and Falkowski (91).
**Figure 1.4. High caspase specific activity in *H. volcanii*.** (A) NaCl and (B) temperature optimization of caspase specific activity in *H. volcanii* cell extracts (white bars) compared to purified recombinant human caspase 8 (black bars). Caspase activity is highest at 1.5 M NaCl and at optimal temperature of 42°C. Adapted from Bidle et al. (53).
**Figure 1.5.** *H. volcanii* has a high caspase-like activity at different salinities and phases of growth. Caspase specific activity in *H. volcanii* cells grown in low (1.5 M), optimal (2.1 M), and high (3.5 M) NaCl concentrations and harvested at mid-exponential (me), stationary (st), or death (d) phases. Error bars represent standard deviations for triplicate measurements. Adapted from Bidle et al.(53).
Figure 1.6. Effect of caspase inhibition on cell growth and *in vivo* caspase specific activity during incubation in low (1.5 M), optimal (2.1 M), and high (3.5 M) salt concentrations. *H. volcanii* cell growth (left panels; measured via optical density at 600 nm) and *in vivo* caspase specific activity (right panels) in the presence of either 25 µM z-VAD-FMK (solid symbols, solid lines) or 0.2% DMSO (open symbols, dotted lines). z-VAD-FMK was dissolved in DMSO. Asterisks in 2.1 M (left panel) indicate samples used for Western analysis in Figure 1.6. Standard errors for cell growth measurements (left panels) were smaller than symbol size. Adapted from Bidle et al. (53).
**Figure 1.7. Caspase activity and protein expression in representatives from archaeal and bacterial domains.** (A) Basal, caspase specific activity in cell extracts from different Euryarchaeota: haloarchaea (*H. volcanii, SC3, SC8*), a methanogen (*Methanosarcina* sp.), a hyperthermophile (*Pyrococcus furiosus*); an acidophilic Crenarchaeote (*Sulfolobus solfataricus*); halophilic bacteria (*2HS38b, 1HS38b*); mesophilic Gram-positive (*Escherichia coli*) and Gram-negative (*Bacillus subtilis*) bacteria. (B) Western blot analyses of aforementioned cell extracts probed with human caspase 8 antisera. Extracts were generated from cells grown under their respective optimal conditions. *E. coli* and *B. subtilis* were not included in Western blot analysis. Immunohybridization of *T. erythraeum* cell extracts to human caspase antibodies has been previously reported (91). Asterisks denote halophilic representatives. Adapted from Bidle et al. (53).
Figure 1.8. Original characterization of the model haloarchaeon, *Haloferax volcanii*. (A) Light micrograph of Ben Volcanii’s original Dead Sea ‘bacteria’ indicating pleomorphic microbes in the Dead Sea; Adapted from Wilkansky, 1936. (B) Effect of NaCl on growth of *H. volcanii* in the tryptone yeast autolysate salt medium; Adapted from Mullokhanbhai and Larsen (62).
**Figure 1.9. H. volcanii has a tractable genetic system.** Genetic transformation methods have been devised for *H. volcanii* making it an excellent model haloarchaeon to test the function of specific candidate genes. Routine genetic manipulation for creating mutants uses “pop-in, pop-out” deletion method using a variety of selectable markers most commonly using schema “b” (48, 80). Adapted from Allers and Mevarech (48).
Section 1.3 **REFERENCES**


CHAPTER 2

Specificity of archaeal caspase activity in the extreme halophile *Haloferax volcanii*

Mansha Seth-Pasricha¹, Kelly A. Bidle², Kay D. Bidle¹

¹Department of Marine and Coastal Sciences, Rutgers University, 71 Dudley Road, New Brunswick, NJ 08901
²Department of Biology and Behavioral Neuroscience, Rider University, Lawrenceville, NJ
Specificity of archaecal caspase activity in the extreme halophile *Haloferax volcanii*

Mansha Seth-Pasricha,1, Kelly A. Bidle2 and Kay D. Bidle1,*  
1Institute for Marine and Coastal Sciences, Rutgers University, New Brunswick, NJ, USA.  
2Department of Biology, Rider University, Lawrenceville, NJ, USA.

Summary

Caspase-like proteases are key initiators and executors of programmed cell death (PCD), which is initiated by environmental stimuli and manifests in organisms ranging from unicellular microbes to higher eukaryotes. *Archaea* had been absent from the caspase inheritance discussion due to a lack of gene homologues. We recently demonstrated extremely high, basal caspase-like catalytic activity in the model haloarchaeon, *Haloferax volcanii*, which was linked to the cellular stress response and was widespread among diverse *Archaea*. Here, we rigorously tested the catalytic specificity of the observed archaecal caspase-like activities using hydrolytic assays with a diverse suite of protease substrates and inhibitors compared with known model serine and cysteine proteases (trypsin, cathepsin, papain, and human caspase-8). Our experiments demonstrate that exponentially growing *H. volcanii* possesses a highly specific caspase-like activity that most closely resembles caspase-4, is preferentially inhibited by the pan-caspase inhibitor, zVAD-FMK, and has no cross-reactivity with other known protease families. Our findings firmly root the extremely high levels of caspase-like activity as the dominant proteolytic activity in this extreme haloarchaeon, thereby providing further support for housekeeping functions in *Halobacteria*. Given the deep archaecal roots of eukaryotes, we suggest that this activity served as a foundation for stress pathways in higher organisms.

Introduction

Caspases are a family of highly refined, intracellular cysteine proteases that cleave a wide variety of substrate proteins at the C-terminus of an aspartate residue within specific tetrapeptide motifs. They generally display strong conservation in amino acid sequence, structure and substrate specificity (Cohen, 1997; Stennicke and Salvesen, 1998; Thornberry and Lazebnik, 1998) possessing a conserved domain structure with a histidine- and cysteine-containing catalytic diad. A variety of caspases have been identified in different metazoan animals, ranging from *Hydra* to humans (Thornberry and Lazebnik, 1998; Cikala et al., 1999; Sanmartín et al., 2005) that are key initiators and executors of programmed cell death (PCD) or apoptosis, a genetically controlled, irreversible form of cell death that elicits specific morphological changes initiated by environmental stimuli (Lockshin and Williams, 1965; Kerr et al., 1972).

Although PCD was first discovered in multicellular organisms, which use it for development and defence, it has now been shown to be a ubiquitous trait throughout nature, spanning diverse prokaryotes, and both unicellular and multicellular eukaryotes. PCD is catalysed either by classic caspases or orthologous proteins, such as metacaspases, paracaspases, pytaspases, and sas-pases (Vaux and Korsmeyer, 1999; Koonin and Aravind, 2002; Coffeen and Wolpert, 2004; Riedl and Shi, 2004; Salvesen and Abrams, 2004; Sanmartín et al., 2005; Chichkova et al., 2010; Mace et al., 2010; Vartapetian et al., 2011). Genomic, morphological and biochemical evidence of caspase-mediated PCD has also been documented in widely diverse evolutionary lineages of prokaryotic and eukaryotic phytoplankton including cyanobacteria, coccolithophores, diatoms, and dinoflagellates (Vardi et al., 1999; Segovia et al., 2003; Berman-Frank et al., 2004; Bidle and Falkowski, 2004; Mohankar et al., 2006; Bidle et al., 2007; Bidle and Bender, 2008). These cumulative findings have peaked interest in exploring the molecular evolution of caspase-like proteins, their function, and the evolutionary drivers that have influenced their retention in different microbial lineages (Ameisen, 2002; Koonin and Aravind, 2002; Bidle and Falkowski, 2004).

Until recently, *Archaea* were absent from the discussion on the establishment, maintenance, and inheritance of apoptotic machinery, as their genomes lack clear homologues of these proteins (Koonin and Aravind, 2002). However, the incidence and roles of caspase-like proteins are only now beginning to be recognized in this third
Results and discussion

Specificity of caspase activity in H. volcanii cell extracts

We challenged cell extracts from exponentially growing H. volcanii cells with a diverse suite of canonical, fluorogenic caspase tetrapeptide substrates [YVAD- (caspase-1), VDVAD- (caspase-2), DEVD- (caspase-3), LEVD- (caspase-4), WEHD- (caspase-5), VEID- (caspase-6), IETD- (caspase-8), and LEHD-AMC (caspase-9)] to expand our coverage from previously documented caspase-8-like activity (Bidle et al., 2010) and to better elucidate the nature of caspase-like activity in this model haloarchaeon. Haloferax volcanii cell extracts displayed ~ 10- and ~ fivefold higher specific activities with LEVD-AMC (~ 232 relative fluorescence units (RFU) h⁻¹ mg protein⁻¹) and VEID-AMC (~ 301 RFU h⁻¹ mg protein⁻¹) substrates, respectively, compared with IETD-AMC (6909 RFU h⁻¹ mg protein⁻¹). AMC fluorescence calibration curves placed the corresponding substrate hydrolysis rates at 4124, 2000, and 383 nmol h⁻¹ mg protein⁻¹ (Fig. 1). Much lower specific activities were observed for extracts incubating with the other caspase substrates; VDVAD-, YVAD-, DEVD-, LEHD- and WEHD-based activities ranged from only 9% to 47% the IETD-based activity (data not shown). Notably, the same pattern of relative hydrolysis rates for LEVD-, VEID-, and IETD-AMC was observed in extracts from cells grown under low (1.5 M) and high (3.5 M) salt stress; VEID and IETD were 25–57% and 8–18% of these observed LEVD hydrolysis rates (Fig. 2). These findings may indicate the presence of either one or several proteins with overlapping activities and/or functions.

Our observed caspase-like-specific activities are by far the highest ever reported in a unicellular microbe, more notably in an exponentially growing archaeon. Reported caspase-specific (IETDase) activities in diverse unicellular marine phytoplankton, including cyanobacteria, diatoms, coccolithophores, and chlorophytes, are comparatively low for exponentially growing cells [10 s to 100 s of RFUs h⁻¹ mg protein⁻¹ (Segovia et al., 2003; Berman-Frank et al., 2004; Bidle and Falkowski, 2004; Bidle et al., 2007; Bidle and Bender, 2008; Bidle and Kwitny, 2012)], with rates increasing ~ 10-fold when cells have activated autocatalytic cell death pathways. A wider survey of hydrolysis rates in unicellular microbes using other canonical caspase substrates has only been examined with the unicellular chlorophyte, Dunaliella tertiolecta (Segovia et al., 2003), whereby hydrolytic cleavage of various caspase substrates was tested in cell extracts from stressed cells that were in different stages of light deprivation (0–5 days). Nonetheless, the maximum activity was seen for IETD and LEHD at ~ 300 RFU h⁻¹ mg protein⁻¹ after 5 days light deprivation and associated PCD activation. Overall, the activity rates for these
substrates ranged from 50 to 300 RFU h⁻¹ mg protein⁻¹ over the 5 day time-course. Our hydrolysis rates for exponentially growing *H. volcanii* cells, along with our previous measurements in diverse *Archea* (Bidle et al., 2010), comparatively dwarf these rates from other unicellular microbes (by more than an order of magnitude), putting them in a unique catalytic and physiological context.

A comparison of substrate specificity for *H. volcanii* cell extracts to four model proteases, trypsin, cathepsin, papain, and recombinant human caspase-8, helped to further place it in the caspase catalytic class (Fig. 1). These model proteases represent distinct classes of peptidases that have well characterized active sites, substrate requirements, and inhibitors (Otto and Schirmeister, 1997; Hedstrom, 2002; Barrett and Rawlings, 2007; Vartapetian et al., 2011). Not surprisingly, when challenged with a suite of fluorescent substrates [ARR-AMC (trypsin and cathepsin), LLVY-AMC (papain), and IETD-AMC, VEID-AMC, and LEVD-AMC (caspase-8, caspase-6, and caspase-4 respectively)], each model protease had the highest activity with its preferred substrate (Fig. 1). Caspase-8 demonstrated notable selectivity among the caspase substrates, with highest catalytic activity for IETD-AMC. We did notice higher IETDase activity in the present study for purified human caspase-8 compared with what had been previously reported (Bidle et al., 2011), likely due to a different source and batch of enzyme. *Haloferax volcanii* cell lysates had no measurable activity towards substrates of trypsin, cathepsin, and papain substrates under our incubation conditions at 42°C (Fig. 1).

We also tested for general leucine and aspartate aminopeptidase activity in *H. volcanii* extracts through hydrolysis of leucine-AMC (Leu-AMC) and aspartate-AMC (Asp-AMC), respectively, given the Haloarchaea genome (http://archaea.ucsc.edu/cgi-bin/hgGateway?db=haloVolc1) has 13 annotated aminopeptidases (HVO_0160, CAAX aminoterminal protease; HVO_1966, CAAX aminoterminal protease family; HVO_1829, aminopeptidase homologue; HVO_0836, aminopeptidase; HVO_1774, leucyl aminopeptidase; HVO_0826, deblocking aminopeptidase; HVO_0477, aminopeptidase II; HVO_0477, ampS aminopeptidase; HVO_0826, degrading aminopeptidase; HVO_0836, aminopeptidase; HVO_1774, leucyl aminopeptidase; HVO_1829, aminopeptidase homologue; HVO_1849, aminopeptidase; HVO_2600, methionine aminopeptidase; HVO_A0535, aminopeptidase putative; HVO_1966, CAAX aminoterminal protease family; HVO_1997, CAAX aminopeptidase; HVO_0082 CAAX amino terminal protease; HVO_0160, CAAX...
amino terminal protease). Leu-AMC is a commonly used aminopeptidase substrate for diverse environmental microbes (Hoppe et al., 2002) and has served as a model for general proteolytic activity. *Haloferax volcanii* grows at an acidic pH of ~5.1 and contains ~13 moles of Asp per 100 moles of total amino acids (Mullakhanbhai and Larsen, 1975; Hartman et al., 2010), so Asp-AMC incubations helped verify that our observed caspase cleavage were specific to the C-terminal aspartate residue in tetrapeptide motifs, as opposed to a general aspartate cleavage response. Very low Asp-AMC hydrolysis activity (2 ± 0.4 nmol h⁻¹ mg protein⁻¹; Fig. 1) was detected in *H. volcanii* cell extracts, putting it at 0.05% LEVD-ase activity, while no detectable Leu-AMC hydrolysis was detected. For comparison, little to no activity was observed in *Escherichia coli* cell extracts incubated with Asp-AMC, but Leu-AMC displayed very high activity (2201 nmol h⁻¹ mg protein⁻¹).

Our results confirmed the consistent and extremely high basal activity in exponentially growing *H. volcanii* cells and refined the activity as more caspase-4-like, and demonstrate that the previous findings using IETD actually underestimated *H. volcanii* caspase activity. In order to address the possibility that the absence of some enzyme activities (e.g. LLVY-, ARR-, and Asp-AMC hydrolysis) in our extracts may have been due to competing small molecules, we also independently assayed for these activities in cell extracts heated to 60°C and in partially purified proteins at 42°C obtained using standard ammonium sulfate (AmSO₄) precipitation techniques (20%, 40%, 60%, 80% and 100%; Fig. S1); they were generally enriched for general proteolytic activity (i.e. LEVDase) and was not detectable at physiological temperatures.

**Selectivity of caspase inhibition**

The specificity of caspase-like activity exhibited in *H. volcanii* was further verified by challenging cell extracts with a panel of protease inhibitors that target a diverse array of proteases including serine proteases (phenylmethylsulfonylfluoride, PMSF; aprotinin, Apr), cysteine proteases (leupeptin, Leu; E-64), metalloproteases (EDTA), and a cocktail of these inhibitors (pooled inhibitors, PI) or caspases (z-VAD-FMK). The aforementioned four model proteases, trypsin (serine protease), papain and cathepsin (cysteine proteases), and caspase-8 (caspase), were used to diagnose the specificity and efficacy of the chosen inhibitors when incubating with their preferred substrate.

**Fig. 2. Caspase-specific activities in cell extracts from *H. volcanii* cells grown in low (1.5 M), optimal (2.1 M), and high (3.5 M) NaCl concentrations and harvested at mid-exponential phase. Cell growth at respective salinities was performed as previously described (Bidle et al., 2008). Extract preparation and kinetic cleavage assays were conducted as described in Fig. 1, with cell extracts incubating with LEVD-, VEID-, and IETD-AMC. Numbers above VEID and IETD bar graphs indicate the percent of LEVD activity. Error bars represent standard deviations for triplicate measurements.**
As expected, each specific inhibitor exhibited between 84 and 100% inhibition of protease activity for its given substrate (Fig. S2). Interestingly, papain and cathepsin, a cysteine peptidase of the papain protease family, were both completely inhibited (99–100%) by zVAD-FMK, a pan-caspase inhibitor. This observation was not entirely surprising since it has been suggested that zVAD-FMK can indeed inhibit papain and papain-like peptidases (Rozman-Pungercar et al., 2003). However, unlike H. volcanii extracts, neither purified papain nor cathepsin demonstrated preferential catalytic cleavage of caspase tetrapeptide recognition sequences (Fig. 1 and Schotte et al., 1999).

We observed nearly complete inhibition (95%) of caspase-8-like activity (IETDase) when H. volcanii extracts were treated with zVAD-FMK (Fig. 3), corroborating our previous findings (Bidle et al., 2010). Likewise, zVAD-FMK severely inhibited (> 85%) H. volcanii caspase-4-like (LEVDase) and caspase-6-like (VEIDase) activities further supporting that this activity is produced by a caspase-like enzyme(s) despite possessing distinct sequence differences to classic caspase superfamily proteins (Bidle et al., 2010). Given zVAD-fmk is a broad pan-inhibitor of caspases that is capable of inhibiting different caspases, we further refined our analysis of caspase inhibition by challenging cell extracts with fluoromethylketone inhibitors with higher specificity to individual caspase activities (e.g. LEVD-fmk, VEID-fmk, and IETD-fmk). The response to these specific caspase inhibitors may shed more light on the specificity of the observed caspase activities. In each case, we observed a dose-dependent inhibition of LEVDase, VEIDase and IETDase activities by their respective inhibitors (Fig. 4), but the degree of inhibition with 50 uM of individual inhibitors was considerably lower than that observed for 20 uM zVAD-fmk (Fig. 3).

Higher inhibition with zVAD-CHO over specific caspase inhibitors (e.g. IETD-CHO) has also been reported for ‘phytaspases’, plant proteases with caspase activity (Chichkova et al., 2010). For comparison, application of zVAD-fmk and IETD-fmk inhibited the activity of purified, recombinant human caspase-8 by 99.9% and 99.7% respectively (Fig. 4). Our findings suggest that the observed caspase activities in H. volcanii are not as refined as in higher eukaryotes and may represent a
broader, more ancestral type of activity (Fig. 4). These results are also consistent with the presence of either one or several proteins with overlapping activities and/or functions.

Interestingly, EDTA partially inhibited the observed caspase activities derived from all three substrates tested (by 40–50%), indicating a possible requirement for magnesium ions (Rodriguez-Valera, 1995). *Haloferax volcanii* has a very high tolerance for MgCl₂ (Mullakhanbhai and Larsen, 1975), likely reflective of the abundant MgCl₂ concentrations found in the Dead Sea, which constitute 50.8% of the anhydrous chlorides on a weight percentage basis, compared with 30.4% for NaCl (Steinhorn, 1983). This is remarkably different from the composition of ocean seawater which is ~97% NaCl. Addition of 100 μM MgCl₂ to EDTA-treated cell extracts partially reconstituted caspase activity levels for all three substrates tested, but it was not able to restore all activity. MgCl₂ addition consistently elevated activity more than the addition of 100 μM CaCl₂ (Fig. 5) albeit slightly, thereby providing some support for a Mg²⁺ requirement.

**Potential cellular roles**

Our findings firmly root extremely high levels of specific caspase activity, particularly resembling caspase-4, as the dominant proteolytic activity in this extreme halophilic archaean. These findings, along with our previous work that linked very high basal caspase activity to the cellular salt stress response (Bidle et al., 2010) raise intriguing questions as to the identity, function, and evolution of these caspase-associated proteins. We suggest that this caspase-activity may have emerged in *Archaea*, for essential housekeeping functions and possibly served as a foundation for stress and PCD pathways in higher organisms. In higher eukaryotes, caspase-4 activity has been implicated in endoplasmic reticulum (ER) stress (Kim et al., 2006; Binet et al., 2010), which can activate the ‘unfolded protein response’ (UPR) under conditions that alter protein folding or calcium homeostasis (Patil and Walter, 2001). UPR leads to the increased expression of chaperones and folding enzymes, which prevent the aggregation of misfolded proteins and facilitate proper protein folding. UPR can also induce components of the protein degradative machinery to remove misfolded proteins, which are tagged by ubiquitin and degraded in the proteosome. This physiological response affords the cell time to survive a stressful insult, unless enhanced or prolonged stress initiates PCD. Interestingly, eukaryotic proteosomes have been shown to contain caspase-like sites and associated activity (Kisselev et al., 2003; Murata et al., 2009). Archaeal proteosomes, which demonstrate a high level of structural similarity to those found in eukaryotes (Maupin-Furlow et al., 2005) have been well studied in *H. volcanii* and play an integral role in the organisms’ cellular stress response (Zhou et al., 2008). Could the...
high-level caspase activity be linked to proteosomes and an ancestral UPR response? A UPR is present in all eukaryotes studied to date, but little is known if a similar response operates in Archaea. Regulation of proper protein folding in the face of very high extra- and intracellular salt environments encountered during growth is a critical concern for H. volcanii (and other extreme halophiles). Our results highlight the need to identify the protein(s) that confer caspase activity and function. Given the high specificity of H. volcanii to caspase catalytic activity our findings strongly suggest that recently developed in situ inhibitor trapping techniques (Tu et al., 2006; Mohr and Zwacka, 2007) are a promising avenue to identify these enigmatic proteins and begin to elucidate their physiological roles and molecular evolution.

Acknowledgements

We thank Dr Chris Brown, Liti Haramaty, and Frank Natale for technical assistance and helpful discussions. This research was supported by NSF grants (OCE-0927829 and OCE-1061883) to K. D. B.

References


© 2012 Society for Applied Microbiology and Blackwell Publishing Ltd. Environmental Microbiology Reports


Supporting information

Additional Supporting Information may be found in the online version of this article:
Supplemental Figure S1. Panel of proteolytic activities in partially purified protein fractions.

Hydrolysis rates were determined for cell extracts (CE) from exponentially growing \textit{H. volcanii} cells, partially purified proteins obtained by sequential ammonium sulfate (AmSO$_4$) precipitation (20%, 40%, 60%, 80% and 100%), and non-AmSO$_4$-precipitable proteins (soluble). AmSO$_4$ was sequentially added to cell extracts at 4°C with constant stirring at and incubated for 30 min at each step. Prior to each AmSO$_4$ addition, proteins were pelleted at 10,000x g. The supernatant, representing dissolved proteins, was transferred to a new tube and used for the next % AmSO$_4$ addition. Protein pellets were resuspended in 100 mM Tris-Cl (pH 7.4)/150 mM NaCl and activity was measured using the aforementioned assay conditions and panel of fluorogenic substrates. Note difference in scale between caspase substrates (upper panel) and other protease substrates (lower panel). Hydrolysis of Leu-AMC was also measured but no activity was detected in any fraction. Error bars represent standard deviations for triplicate measurements.
Supplemental Figure S2. Protease inhibitor efficacy on model protease substrates. To verify inhibitor effectiveness, purified proteases (recombinant human caspase 8, trypsin, papain, cathepsin B) were individually incubated with inhibitor treatments and compared to uninhibited control reactions. Each purified enzyme was pre-incubated with individual protease inhibitors at 37˚C for 1 h at the aforementioned concentrations (see Figure 3) prior to the addition of the preferred fluorogenic substrate for each respective enzyme: caspase-8 with IETD-, trypsin and cathepsin with ARR-, and papain with LLVY-AMC).
Section 2.2 **ACKNOWLEDGEMENTS**

This chapter, in its entirety, is a reprint of the material as it appears in the April 2013 issue of Environmental Microbiology Reports, Volume 5, Issue 2, pages 263-71. The dissertation author was the primary researcher and author, and the co-authors listed in this publication directed and supervised the research which forms the basis for this chapter.

CHAPTER 3

Isolation and Identification of Proteins Associated with Caspase Activity in *Haloferax volcanii*

Mansha Seth-Pasricha\(^1\), Stefan Senn\(^3\), Laura E. Sanman\(^4\), Matthew Bogyo\(^4\), Kelly A. Bidle\(^2\), Kay D. Bidle\(^1\)

\(^1\)Department of Marine and Coastal Sciences, Rutgers University, 71 Dudley Road, New Brunswick, NJ 08901
\(^2\)Department of Biology and Behavioral Neuroscience, Rider University, Lawrenceville, NJ
\(^3\)Christian Doppler Laboratory for Innovative Tools for Biosimilar Characterization, University of Salzburg, Hellbrunnerstrasse 34, 5020 Salzburg, Austria (current address)
\(^4\)Department of Pathology, Stanford University School of Medicine, 300 Pasteur Drive, Stanford, CA 94305-5324
Section 3.1 **ABSTRACT**

The model haloarchaeon, *Haloferax volcanii*, possesses an extremely high, highly specific basal caspase-like activity in exponentially growing cells that most closely resembles caspase-4, is preferentially inhibited by the pan-caspase inhibitor (zVAD-FMK), and has no cross-reactivity with other known protease families. It is the dominant proteolytic activity in this extreme haloarchaeon and may perform housekeeping functions in *Haloarchaea*, perhaps as part of the unfolded protein response. Unfortunately, the identities of the protein(s) that confer this activity in haloarchaea are currently unknown. Here, biochemical purification and in situ trapping with biotinylated fmk- and AOMK-based inhibitors combined with genome-enabled proteomics and structural alignments were collectively used to identify the protein(s) that are associated with caspase activity and begin to elucidate their cellular function in *Archaea*. Spectra analysis of LC-MS/MS data for trapped proteins was used to identify proteins that were preferentially enriched with caspase-specific inhibitors. Furthermore, modeled tertiary structures and structural alignments for 269 recovered proteins with the crystallized active site of human caspase-6 were used to identify promising candidates. Collectively, these analyses identified thermosomes, chaperonins, proteasomes, a cell division protein (an ATPase recently identified as an activator of proteasomal degradation), a putative nuclease, putative aminopeptidase, elongation factor αEF-2, and an ornithine cyclodeaminase as key proteins associated with caspase activity. Our findings biochemically connect caspase activity in *H. volcanii* to specific stress-related protein complexes, such as those involved in the unfolded protein response.
Section 3.2 INTRODUCTION

*Haloferax volcanii* is a moderate obligate haloarcheon that grows at salinities between 1.5-3.5 M (1), having been isolated from the Dead Sea. *H. volcanii* is considered a model haloarcheon due to its ecologically relevance, ease of growth in the laboratory in both batch liquid culture and on solid media, availability of a stable, sequenced genome and well-established genetic system. Our previous work firmly established very high caspase-like activity in *H. volcanii* (2) that is the dominant cellular proteolytic activity and is part of a cellular stress response (3). In order to better elucidate the nature of caspase activity in this model haloarchaeon, *H. volcanii* cell extracts were challenged with a diverse suite of canonical, fluorogenic caspase tetrapeptide substrates (2, 3). *H. volcanii* cell extracts displayed ~eight- and ~five-fold higher specific activities with LEVD-AMC (caspase-4) and VEID-AMC (caspase-6) substrates respectively, compared to IETD-AMC (caspase-8) (3). Much lower specific activities were observed for extracts incubated with the other caspase substrates; VDVAD-, YVAD-, DEVD-, LEHD- and WEHD-ase activities were only 0.09 to 0.47% of the IETD-based activity (3).

Subsequent work firmly established the specificity of this activity in *H. volcanii* by using a variety of inhibitors and model proteases (trypsin, cathepsin, papain, and recombinant human caspase-8), firmly placing it in the caspase catalytic class (3). These model proteases represented distinct classes of peptidases that have well characterized active sites, substrate requirements, and inhibitors (4-7). Not surprisingly, when challenged with a suite of fluorogenic substrates [ARR- (trypsin and cathepsin), LLVY- (papain), and IETD-, VEID-, and LEVD-AMC (caspases), each model protease had the highest activity with its preferred substrate. The
specificity of caspase activity exhibited in *H. volcanii* was further verified by challenging cell extracts with a panel of protease inhibitors that target a diverse array of protease (3) including serine proteases (phenylmethylsulfonylfluoride, PMSF; aprotinin, Apr), cysteine proteases (leupeptin, Leu; E-64), metalloproteases (EDTA), and a cocktail of these inhibitors, or caspases (z-VAD-fmk) (3). Nearly complete inhibition (85-95%) of caspase activity was observed when *H. volcanii* extracts were treated with z-VAD-fmk, corroborating previous findings (2).

*An in silico* analysis of the *H. volcanii* proteome even identified ~50 potential target substrate proteins containing the site-specific, caspase tetrapeptide sequences (IETD, VEID, LEVD) strongly suggesting that this catalytic activity is part of orchestrated cellular pathways. Indeed, the putative substrate proteins represented a broad suite of cellular functions [(2) and unpublished data], playing diverse roles in cell physiology such as energy metabolism, information transfer (central dogma), transport, membrane biosynthesis, nuclease activity, and protein folding (chaperonins). Potential proteolytic targets were also identified in proteins having putative roles in osmotic and temperature stress, allosteric regulation, and cell communication. Unfortunately, the identities of the protein(s) that confer the extremely high basal caspase activity in haloarchaea are currently unknown. It is intriguing to speculate that these enzymes in *H. volcanii* (and other Archaea) represent a unique, evolutionarily distinct subset of proteins since they clearly have sequence differences to classic caspases, yet share the diagnostic catalytic activity. Their lack of homologous caspase domain structures has made it intractable to uncover their identity using genome sequence information alone. Rather it requires unique approaches utilizing their documented biochemical activity (2, 3).
Here, biochemical purification and in situ trapping with biotinylated fmk- and AOMK-based inhibitors combined with genome-enabled proteomics and structural alignments were employed to identify candidate proteins associated with the high, basal caspase activity in H. volcanii cell extracts. The analysis revealed that a suite of functionally diverse proteins were associated with this dominant proteolytic activity. These included thermosomes, chaperonins, proteasomes, a cell division protein (an ATPase recently identified as an activator of proteasomal degradation), a putative nuclease, putative aminopeptidase, elongation factor αEF-2, and an ornithine cyclodeaminase as key proteins associated with caspase activity. These proteins were repeatedly identified by LC-MS/MS and either displayed differential peptide spectral counts with caspase-specific inhibitors or had structural similarities to the active site of human caspases. These diagnostic biochemical findings connect caspase activity in H. volcanii to specific stress-related protein complexes, such as those involved in the unfolded protein response (proteasomes, thermosomes, etc.) opening up an opportunity to empirically test their relationship to caspase activity and to assess their physiological roles and link to UPR with targeted gene knockouts.

Section 3.3 MATERIALS AND METHODS

Growth and generation of cell lysates. Haloferax volcanii wild-type strain DS70 was grown aerobically at 45°C in batch culture in an optimal 2.1M salt medium (2, 3) to mid-exponential phase. Cell lysates were prepared as described previously (3). Briefly, cells were pelleted (5000 x g, 15 min; RT; wet weight 3.874 g), dissolved in 13 ml of 100 mM Tris-Cl (pH7.4)/2M NaCl buffer and sonicated on ice (4x, 30 sec cycle⁻¹) using an Ultra Cell Disruptor (Microson) set at
Power 2. Cell lysate was further centrifuged (8800 x g, 10 min) to remove cell debris with supernatant material being used in various downstream procedures including caspase activity assays, ammonium sulfate [(NH₄)₂SO₄] precipitation, SDS-PAGE, Western Blot analyses and inhibitor trapping (see below).

**Ammonium sulfate precipitation.** Partially purified proteins were obtained from cell extracts by sequential ammonium sulfate [(NH₄)₂SO₄] precipitation (20%, 40%, 60%, 80% and 100%) using standard protocols (8). Briefly, (NH₄)₂SO₄ was sequentially added to cell extracts at 4°C with gentle rocking and incubated for 30 min or until dissolved, at each step. Prior to each (NH₄)₂SO₄ addition, proteins were pelleted by centrifugation (11,000 x g, 20 min; 4°C). The supernatant, representing dissolved proteins, was transferred to a new tube and used for the next (NH₄)₂SO₄ addition. Pellets at each stage were dissolved in 3 ml of 100 mM Tris-Cl (pH7.4)/2M NaCl buffer. The post-100%, non-(NH₄)₂SO₄-precipitable proteins (soluble) were dialyzed against buffer (3.5 kDa MWCO; 3 changes over 20 h; 4°C) to remove excess salt. Total protein concentration in each fraction was measured using the BCA assay (Thermo scientific), with bovine serum albumin serving as a standard.

**Size exclusion chromatography.** Size exclusion chromatographic separation of caspase-active proteins from *H. volcanii* extracts was performed using high-performance liquid chromatography (HPLC) and 7.8 mm diameter gel purification columns (SEC-3; Agilent), with both wide (5,000 to 1,250,000 Da; 300 mm length, 300 Å pore size) and narrow (500-150,000 Da; 150 mm length, 150 Å pore size) molecular weight ranges. Briefly, 1.6 mg total protein in
cleared cell extracts was suspended in 100 mM Tris-Cl (pH7.4)/2 M NaCl buffer and applied to the column with an aqueous phase buffer 50 mM NaPO₄ 150 mM NaCl (pH 7.0). Separation was achieved by HPLC (LC-10AT; Shimadzu) at a flow-rate of 0.5 ml min⁻¹. Collections were made every 2 min with the wide range column followed by every 20 seconds with the narrow range column. Standardization was done using gel filtration protein standards (Bio-Rad). Caspase activity assays and SDS-PAGE were performed on each fraction (see below).

**Fluorogenic protease activity assays.** Protein-normalized cleavage rates (nmol h⁻¹ mg protein⁻¹) were measured in 96-well microtiter plate format using a variety of different substrates. Caspase activities used a suite of canonical, fluorogenic substrates [LEVD-AMC (caspase-4), VEID-AMC (caspase-6), and IETD-AMC (caspase-8); Enzo Life Sciences] as described previously (2, 3). Trypsin/Cathepsin and Papain/Chymotrypsin activities used ARR-AMC and LLVY-AMC (Sigma), as described previously (3). Briefly, 15 μg of protein from each sample was incubated with individual fluorogenic substrates in a modified, assay buffer [50 mM HEPES (pH 7.4), 10% Sucrose, 0.1% CHAPS, 10 mM DTT, 1.5 M NaCl]. The kinetics of substrate cleavage (Ex 380 nm, Em 460 nm) was measured at 42°C (or 60°C as required) over 1-2 h at 3-5 min resolution using a Spectra Max M3 plate reader equipped with SoftMax Pro 6.2.1. Assays were performed in triplicate.

**SDS-PAGE and Western blot analysis.** Twenty μg of total protein was loaded onto either 4-20%, 8-16% gradient or 12% TGX precast gels (Bio Rad) followed by visualization with Coomassie Blue-based (Gel Code Blue, Thermo Scientific) or silver-based stains (Pierce Silver Stain Kit;
Thermo Scientific). Stained gels were washed in MilliQ water and incubated in 3% glycerol for 1 hour before drying under vacuum (2 hours, 80°C; Bio Rad Gel Dryer). For Western Blot analysis, proteins were transferred onto PVDF membranes (100 V; 1 h), blocked with ECL blocking reagent (GE Life Sciences), washed 4x with 1x PBS 0.05% Tween-20, and probed with an Avidin-HRP antibody (1:10,000 dilution) directed against biotin bound moieties and visualized using Amersham ECL Plus western blotting reagent (GE Life Sciences) and a CCD imager. (Chemidoc XRS+, Bio-Rad).

For control experiments using human caspase-8, fractions were run on 4-20% TGX precast gels (Bio Rad), transferred on PVDF membrane, and blocked with ECL blocking reagent. Samples used in biotinylated trapping experiments were probed either with Avidin-HRP antibody (as above) or probed with a monoclonal anti-human Caspase-8 mouse antibody (Enzo Lifesciences; 1:1500 dilution) and probed with a goat anti-mouse HRP secondary antibody (1:5000). Visualization was as described above.

**Fluoromethyl ketone-based biochemical inhibitors.** A suite of commercially available protease inhibitors (with some conjugated to biotin) containing a fluoromethyl ketone (fmk) reactive functional group was used (50 μM final concentration). They included: biotin-VAD-fmk (Calbiochem), z-VAD-fmk (Enzo Life Sciences), and biotin-FA-fmk inhibitors, with the VAD-based inhibitors being specific for caspase activity. Samples were incubated with inhibitors in caspase assay buffer for 1 hour at 42°C prior to addition of fluorogenic substrates and determination of protein-normalized cleavage rates (nmol h⁻¹ mg protein⁻¹) as previously described [above and (2, 3)]. In some cases, (NH₄)₂SO₄ samples were pre-treated with 50 μM biotin for 30 min at RT,
followed by addition of streptavidin-coated magnetic beads and to remove proteins that would
bind conjugated inhibitors at the biotin moiety and not at the functional group. These
precleared lysates were subsequently treated with biotinylated VAD-fmk and FA-fmk inhibitors
for activity assays.

**Acyloxymethyl ketone-based inhibitors.** A suite of protease inhibitors (with some conjugated
to biotin) containing acyloxymethyl ketone (AOMK) reactive groups were kindly provided by
Matt Bogyo (Stanford University). They included biotin-L-, biotin-R-, biotin-VEID-, biotin-VAD-, D-, KL-, VEID-AOMK, with biotin-VEID, biotin-VAD- and VEID-AOMK being more specific for
caspase activity. 15 μg of *H. volcanii* cell extracts were incubated with the inhibitors in caspase
assay buffer for 1 hour at 42°C prior to addition of fluorogenic substrates and determination of
protein-normalized cleavage rates (nmol h⁻¹ mg protein⁻¹) as previously described [above and
(2, 3)].

**Inhibitor trapping.** Biotinylated inhibitors were used to diagnostically bind and isolate proteins
associated with caspase activity. *H. volcanii* cell lysate and (NH₄)₂SO₄ samples were pre-treated
with 30-50 μM biotin for 30 min at RT, followed by addition of streptavidin-coated
superparamagnetic beads (Thermo Scientific) and to remove proteins that would bind
conjugated inhibitors at the biotin moiety and not at the functional group. This was followed by
incubation with only streptavidin-coated superparamagnetic beads at 4°C for 1 h to remove any
proteins that inherently contain biotin. Beads were prewashed in 100 mM Tris-Cl (pH 7.4)/150
mM NaCl prior to use. Pre-cleared samples were split and separately treated with individual
inhibitors for 1 h at 42°C, followed by incubation with fresh streptavidin-coated magnetic beads (100 μl) for 2 h at 4°C. Beads and associated inhibitor-bound proteins were collected using a Dynal magnetic separator rack (Life Technologies) and washed 3x in 200 μl 100 mM Tris-Cl (pH7.4)/2M NaCl buffer. Competitive elution with subsequent treatments of 2 mM biotin and 0.1 M acidic glycine was first used to remove bound proteins. Beads and any remaining proteins were resuspended in 100 mM Tris-Cl (pH7.4)/2M NaCl and analyzed via SDS PAGE and Western blot.

Inhibitor trapping experiments using biotinylated fmk-based inhibitors (b-VAD-fmk and b-FA-fmk) were performed on caspase active, 100% (NH$_4$)$_2$SO$_4$ fractions. Subsequently, experiments with biotinylated AOMK-based inhibitors (200 μM b-VAD-AOMK and 50 μM b-L-AOMK) used total protein *H. volcanii* cell extracts (500μg total protein), in order to have access to the entire suite of proteins for the purification procedure. Final concentrations were empirically determined in caspase activity assays in order to employ an effective dose. Caspase activity was measured after incubation with each biotinylated inhibitor to confirm degree of inhibition, prior to capture with streptavidin beads.

**Proteomic analysis.** The identity and representation (spectra counts) of proteins recovered from inhibitor trapping procedure was determined by LC-MS/MS analysis at the Biological Mass Spectrometry Facility (Center for Advanced Biotechnology and Medicine, Rutgers University; [http://www3.cabm.rutgers.edu/home.php](http://www3.cabm.rutgers.edu/home.php)). Sixty μl samples were loaded and run ~1cm into a Novex gel Bis-Tris 10% gel. The entire gel band was excised and proteins therein were reduced, carboxymethylated, and digested with trypsin using standard facility protocols. Peptides were
extracted, solubilized in 0.1% trifluoroacetic acid, and analyzed by nano LC-MS/MS using a RSLC system (Dionex) interfaced with a Velos-LTQ-Orbitrap (Thermo Fisher). Samples were loaded onto a self-packed 100 μm x 2 cm trap packed with Magic C18AQ, 5 μm 200 Angstrom (Michrom Bioresources Inc) and washed with ‘Buffer A’ (0.2% formic acid) for 5 min with flow rate of 10 μl min⁻¹. The trap was brought in-line with the homemade analytical column (Magic C18AQ, 3 μm 200 A°, 75 μm x 50 cm) and peptides fractionated at 300 nL min⁻¹ using multi-step gradients of ‘Buffer B’ (0.16% formic acid 80% acetonitrile) consisting of 4 to 25 % over 60 min and 25-55 % over 30 min). Mass spectrometry data was acquired using a data-dependent acquisition procedure with a cyclic series of a full scan acquired in Orbitrap with resolution of 60,000 followed by MS/MS scans (acquired in linear ion trap) of 20 most intense ions with a repeat count of two and the dynamic exclusion duration of 60 sec.

For identification of VAD-fmk modification on cysteine residues of human caspase-8, human caspase-8 control, caspase-8 bound to b-VAD-fmk, and Caspase-8 bound to z-VAD-fmk were digested overnight with 0.01 μg trypsin at 37°C. Digests were acidified with trifluoroacetic acid and analyzed by nano LC-MS/MS using a RSLC system (Dionex) interfaced with a Velos-LTQ-Orbitrap (Thermo Fisher). Alternately, the same samples of human caspase-8 were 1:3 diluted with matrix (10 mg/ml Sinapinic Acid in 50% acetonitrile, 0.1% trifluoroacetic acid) and deposited on opti-TOF 384 well insert for MALDI-TOFTOF (ABSciex) using dry-droplet method. The MALDI-TOF data were acquired using 4800 MALDI-TOFTOF (ABSciex) with linear mid mass positive mode. Data were acquired from 4 to 40 kDa mass range with external calibration by apomyogolobin doubly charged and dimer ions. The laser was fixed at 4600 and 500 laser shot were accumulated for each spectrum.
MS data were searched against the complete annotated *H. volcanii* protein databases (http://archaea.ucsc.edu/cgi-bin/hgGateway? db=haloVolc1) and CRAP.fasta (www.theGPM.org). For human caspase-8, MS data were searched against the human protein database, and CRAP.fasta (www.theGPM.org). Analyses used a local version of the Global Proteome Machine (GPM cyclone, Beavis Informatics Ltd, Winnipeg, Canada) with carbamidoethyl on cysteine as fixed modification and oxidation of methionine and tryptophan as variable modifications using a 10 ppm precursor ion tolerance and a 0.4 Da fragment ion tolerance. Key proteins in AOMK-pull downs with enriched spectral counts (>10) in the b-VAD-AOMK inhibitor trapping were assessed using online skyline quantification software (https://brendanxuw1.gs.washington.edu/labkey/_webdav/home/software/Skyline/%40files/tutorials/AbsoluteQuant-1_4.pdf) to identify proteins likely enriched in b-VAD-AOMK versus b-L-AOMK pull downs.

**Structural modeling and analysis.** Putative tertiary structures were generated for recovered proteins by homology modeling using Phyre2, an online structure prediction tool that uses a variety of methods to find an accurate match for the query protein (9). The Phyre2 server mines protein structures from the Structural Classification of Proteins (SCOP) database augmented with depositions in Protein Data Bank (PDB), PSI-BLAST (http://www.ncbi.nlm.nih.gov/blast/Blast.cgi?CMD=Web&PAGE=Proteins&PROGRAM=blastp&RUN_PSIBLAST=on) for sequence homologs, and three independent secondary structure prediction programs: Psi-Pred, SSPro, and JNet (9).

Two approaches were taken for structural alignments to known caspase proteins. Predicted
structures were matched to the resolved crystal structure of inhibitor (VEID-CHO)-bound caspase-6 (pdb ID 3P4U) using TopMatch, a structure alignment tool that uses established structure-matching techniques to align a query protein with a known structure and generate an alignment score based on alignment length and spatial deviation of superimposed structures (10). Each protein was examined individually for matches in the active site, initially making a ~17 angstrom sphere and then included the entire β sheet core structure, flanked by the α helices. Secondary structure alignment comparisons were evaluated from the calculated output generated by the software, and visual analyses were performed for maximum superimposition.

Secondly, Cytoscape, a network analysis tool, was used to examine the relationships of all recovered protein structures with the 3P4U crystal structure and the putative caspase fold comprising β sheets and turns. Cytoscape allows cellular network comparisons and molecular profiling (11, 12). Since protein structure modeling by Phyre2 generates protein structures based on sequence homology with proteins from sequenced genomes with annotated structures, Topmatch can calculate a predicted phylogenetic interaction with functionality, by identifying the best matches to the 3P4U query for a maximum length of alignment (L) and percent of identical residues (Is), along with a few other criteria like root mean square deviation of all the Cα distances in the alignment in comparison to 3P4U, the number of peptides identified by mass spectrometry for both inhibitors used, and the functionality of the protein identified.

**Human caspase-8.** Purified, recombinant, human caspase-8 (10-70 ng; Enzo Life Sciences) served as a control to test the efficacies of fmk- and AOMK-based VAD inhibitors to a known
caspase and validate their utility in trapping procedure. Caspase-8 was incubated with DMSO (negative control, given inhibitors were dissolved in DMSO) or respective inhibitors (20μM each of z-VAD-fmk and b-VAD-fmk for 1 h at 37°C followed by fluorogenic caspase assay with IETD-AMC (see above for details). For inhibitor trapping, 200-700 ng of human caspase-8 was suspended in 100 mM Tris-Cl (pH 7.4)/150 mM NaCl buffer (or 1x PBS) and incubated with 50μM of either z-VAD-fmk or B-VAD-fmk at 37°C for 1 h. A small aliquot was removed and tested for inhibition of caspase activity. Subsequent steps were identical to those described above for H. volcanii cell extracts except for buffer composition for washes [100 mM Tris-Cl (pH 7.4)/150 mM]. Fractions were run on 4-20% TGX precast gels (Bio Rad) and Western blot analysis performed as described.

**Staining and visualization of caspase active proteins with LE-22.** LE22, a Cy5 reporter tag bound to VEID-AOMK, was used to investigate binding with caspase-like proteins in H. volcanii cell extracts and purified caspases. Total H. volcanii protein (20 μg) was incubated with LE-22 at various concentrations (5-20 μM) at 42°C for 1 h in caspase assay buffer (with 1.5M NaCl). For negative controls, 20 μg H. volcanii cell extract was treated with either DMSO and or 20μM zVAD-fmk. Samples were loaded on precast 8-16% native Tris-HCl gels (BioRad) as well as denaturing SDS PAGE 4-15% precast gels (BioRad). Gels were visualized on a Typhoon 9410 variable mode imager (Ex 633 nm, Em 670 nm; Amersham) and a Biorad Chemidoc MP system (using epifluorecence mode). The efficacy of LE-22 binding was tested at various NaCl concentrations (0.15 M, 0.75 M, 1.5 M). Recombinant, purified caspase-6 (Enzo Lifesciences) was tested for binding with LE22 as above except incubations were at 37°C for 1 h.
Section 3.4 RESULTS AND DISCUSSION

Isolation and identification of proteins exhibiting caspase-like activity

Our previous work has shown that the extremely high, basal caspase-like catalytic activity in *H. volcanii*, is shared across diverse phyla of extremophilic Archaea (2), is linked to physiological stress responses (2), most closely resembles caspase-4 catalytic activity (3), has no cross-reactivity with other known protease families (3), and represents the dominant proteolytic activity in this model extreme haloarchaeon (3). Collectively, these findings provide further support for its housekeeping functions and, given the deep archaeal roots of eukaryotes, they suggest it possibly served as a foundation for stress pathways in higher organisms. Unfortunately, the identity of caspase-associated proteins and insight into their specific cellular roles, including possible interactions with other cellular proteins, still remains unresolved. In light of the highly specific catalytic activity and the clear lack of sequence homologs to known caspase superfamily proteins, we attempted to isolate and identify caspase associated proteins in *H. volcanii* by merging a suite of classic and activity-based biochemical purification with genome enabled proteomics.

Cell extracts from exponentially growing *H. volcanii* cells were first subjected to standard, sequential (NH₄)₂SO₄ precipitation (13) in order to reduce the protein complexity associated with the caspase specific activity for subsequent, more refined biochemical protocols. The protein-normalized caspase specific activity was assessed using a panel of canonical, fluorogenic substrates (LEVD-, VEID-, and IETD-AMC) for complex cell extracts, precipitated proteins (from 20%, 40%, 60%, 80% and 100% (NH₄)₂SO₄ fractions), and a post-
100% (NH₄)₂SO₄ soluble fraction (Figure 3.1). Data is only shown only for LEVDase activity because it represents the highest catalytic activity for these substrates and the patterns were generally the same across purification fractions for each substrate tested. Substantial protein-normalized caspase activity was observed in all (NH₄)₂SO₄ precipitated samples, as well as the post-100% soluble fraction, indicating that either a suite of proteins with different ‘salting out’ properties are responsible for the observed activity, or the responsible protein has a very wide (NH₄)₂SO₄ tolerance threshold, or the caspase-active protease is associated with a variety of cellular proteins in a complex. Activity was maximal in the 100% (NH₄)₂SO₄ fraction, with almost double the activity observed in the crude cell lysate and cleared cell extract, so it was chosen as source material for downstream purification procedures via inhibitor trapping with fmk-based inhibitors. The (NH₄)₂SO₄ fractions were also visualized on SDS-PAGE gels to see if a noticeable refinement in protein complexity could be visually distinguished. A rich array of diverse proteins was detected in each fraction with no obvious changes in the protein profile during (NH₄)₂SO₄ procedure (Figure 3.2). The presence of very similar profiles across all fractions may also help explain the observation of their reasonably high caspase specific activity (>1000 nmol h⁻¹ mg protein⁻¹; Figure 3.1). There was a slight increase in staining intensity for the 100% (NH₄)₂SO₄ and post-100% soluble fractions, even though lanes were loaded with equal protein; these samples displayed the highest two caspase activities measured.

Anion exchange chromatography was performed as a second step of purification, eluting bound proteins with increasing concentrations of NaCl (from 150 mM to 2 M), but use of this technique was discontinued as H. volcanii proteins denatured and clumped in low salt. It was reasoned that this could introduce artifacts and complicate interpretation of downstream
analyses. Size exclusion purification of caspase active proteins from \textit{H.volcanii} extracts was performed using high-performance liquid chromatography and Agilent gel purification columns with both wide (5,000 - 1,250,000 Da) and narrow (500 - 150,000 Da) molecular weight ranges; the caspase activity enriched fraction was found to contain proteins in the range of ~30 to 100 kDa (Figure 3.3), likely representing a very large cellular milieu. These findings didn’t really aid in narrowing the range of proteins and, combined with the fact that the collected fraction volume was extremely small (~166 µl collected every 20 seconds at a flow rate of 0.5ml/min), it precluded use in downstream inhibitor trapping procedures.

\textit{In-situ inhibitor trapping}

A very promising biochemical purification technique that specifically targets the diagnostic catalytic activity of caspase proteases and has been successfully employed with animal caspases is \textit{‘in situ} trapping’, whereby a biotinylated, fluoromethyl ketone (fmk)-based pan caspase inhibitor (e.g. b-VAD-fmk) in combination with streptavidin coated beads, are used to specifically ‘trap’ and ‘pull down’ active caspases proteins (14, 15). The irreversible inhibition of caspase activity with fmk-based peptides (like VAD-fmk) stems from the fact that they resemble the cleavage site of caspase substrates and irreversibly alkylate the cysteine residue in caspase active sites. Therefore, the successful inhibition and recovery of candidate proteins would be diagnostic of caspase activity. This \textit{in-situ} trapping strategy was applied to the caspase-active, 100\% (NH\textsubscript{4})\textsubscript{2}SO\textsubscript{4} fractions from \textit{H. volcanii} using b-VAD-fmk. Notably, we had already established that 25 µM z-VAD-fmk effectively abolished the high \textit{in vivo} caspase activities in \textit{H. volcanii} cells (2) and in \textit{H.volcanii} cell extracts [Chapter 2; (3)], providing
confidence that this approach could work effectively.

Prior to pull downs, the efficacy of the b-VAD-fmk inhibitor on *H. volcanii* cell extracts and on purified recombinant human caspase-8 was tested relative to that of z-VAD-fmk and a biotinylated FA-fmk (b-FA-fmk), a known cathepsin and papain inhibitor (16, 17) and, hence, a negative control for our *in situ* trapping of caspase active proteins. b-VAD-fmk inhibited caspase-like activity in *H. volcanii* cell extracts similarly to z-VAD-fmk; respective inhibition was between 83-88% and 90-95%, suggesting that b-VAD-fmk would be effective in inhibitor trapping of caspase active fractions. The b-FA-fmk inhibitor was not effective at inhibiting caspase activity, which is consistent with its reported specificity for cathepsin activity, not caspase activity (17). As such, and by virtue of it also having a conjugated biotin group, it would target a distinctly different repertoire of trapped proteins (Figure 3.4). The relative efficacy of z-VAD-fmk, b-VAD-fmk, and b-FA-fmk was also tested on recombinant, purified human caspase-8 as an established, bonafide caspase (Figure 3.5). Both VAD-fmk inhibitors completely inhibited human caspase-8 as no catalytic hydrolysis of IETD-AMC was detected in spectrophotometric kinetic assays. No inhibition was reported for b-FA-fmk.

Having established inhibitor efficacy, the successful use of *in situ* inhibitor trapping with streptavidin beads was first tested with recombinant purified human caspase-8 and probed with a monoclonal antibody to human caspase-8 or Avidin-HRP (Figure 3.6). The former assessed whether human caspase-8 itself was being recovered in bead fractions; the latter assessed whether the biotinylated inhibitor was binding to human caspase-8 and ultimately being recovered in the bead fraction. Both independent methods of detection corroborated each other and showed that human caspase-8 is bound to b-VAD-fmk and recovered in the
bead fraction. They specifically revealed the successful trapping ~30 kDa heterodimer and the active site-containing ~20 kDa monomer of human caspase-8. This helped to establish the biochemical specificity of the approach to pull down caspase-active proteases using a biotinylated inhibitor.

For inhibitor trapping of caspase-active proteins in *H. volcanii*, the 100% (NH₄)₂SO₄ fractions were first pre-treated with 50 μM biotin for 30 min at RT, followed by addition of streptavidin-coated beads to remove proteins that would bind conjugated inhibitors at the biotin moiety and not at the functional group. This was followed by incubation with only streptavidin-coated beads to remove any proteins that inherently contain biotin. Cleared lysates were then treated with either b-VAD-fmk or b-FA-fmk and assayed for caspase-like activity to diagnostically confirm specific inhibition and binding of caspase-active proteins (Figure 3.7). Comparative inhibition of caspase activity was ~87% for b-VAD-fmk compared to only 1.5% for b-FA-fmk. The respective protein pools associated with b-VAD-fmk (caspase associated) and b-FA-fmk (cathepsin associated) were affinity purified with streptavidin magnetic beads.

Prior to LC MS/MS analysis of recovered proteins, the effectiveness of *in situ* trapping was monitored by Western blot analysis using an avidin-HRP antibody and mass spectrometry-compatible silver stain (Figure 3.8, Figure 3.9 and Figure 3.10). This allowed us to specifically detect and follow the biotinylated protein pool through the isolation procedure (via avidin-HRP detection) and assess the complexity of the total protein pools (via silver staining) upon b-VAD-fmk and b-FA-fmk treatment. Cell extracts challenged with both inhibitors displayed an increase in biotin-associated proteins, as detected by avidin-HRP hybridization, indicative of a pool of
proteins with affinity for either VAD- and FA-amino acids. At the same time, silver staining revealed a strong refinement in the total protein complexity, consistent with effective purification. The presence of a prominent, biotin-containing ~75 kDa protein in *H. volcanii* cell extracts was consistently detected, some of which was removed onto beads during pre-clearing step. The substantial increase in proteins reacting to avidin-HRP in pre-cleared extracts treated with b-VAD-fmk and b-FA-fmk inhibitors demonstrates that a suite of abundant proteins, many of which share visual similarities and could possibly be part of a large complex, incorporate these inhibitors and that these proteins are successfully recovered with streptavidin magnetic beads. Our inhibitor-based trapping approach allowed us to determine whether proteases bound to either inhibitor are affiliated with the same complex(es) and possibly perform basic housekeeping function(s). The detection of bound protein by either method in the b-FA-fmk inhibitor bound fraction suggested that there were some chymotrypsin or trypsin like proteins in *H. volcanii*. This is consistent not only with our previous work documenting this biochemical activity in post-100% soluble fractions (3), but also findings that proteasomes exhibit chymotrypsin activity at elevated temperatures in *H. volcanii* (18).

A chemically, distinct class of acyloxyethyl ketones (AOMK)-based caspase inhibitors were also tested as independent functional ‘warheads’ to inhibit and *in situ* trap caspase-associated proteins in *H. volcanii*. These AOMK probes are also conjugated to inhibitor peptides (like VAD) and have been reported to be highly cross-reactive to caspases and more specific than the fmk-based peptides (19). AOMK-associated caspase and cysteine-protease inhibitors are a part of a new class of protein detection molecules formulated on activity based protein profiling (ABPP). ABPP is a chemical proteomic strategy for understanding protein function. It
comprises an active-site directed reactive probe that binds to and covalently labels enzymes and can have a reporter tag like a fluorophore or biotin, for enrichment of the probe bound enzyme (20). AOMK probes have been reported to be far more selective for cysteine proteases and displayed low reactivity for weak nucleophiles. Additionally the AOMK bound peptides were reported to be far more sensitive to both active and inactive forms of caspases, lending it higher specificity for the substrate binding site of caspases (21).

A broad variety of AOMK-based inhibitors (biotin-L-AOMK, biotin-R-AOMK, biotin-VEID-AOMK, biotin-VAD-AOMK, D-AOMK, KL-AOMK, VEID-AOMK) were first empirically tested for their efficacy in inhibiting human caspase-8 and caspase-active proteins in H. volcanii. Given these inhibitors are not commercially available and instead have been synthesized by the Bogyo Lab (Stanford University), the array of inhibitors used was dependent on what was available and kindly sent to us for testing. Several of these inhibitors contained biotin conjugates making them compatible with in situ trapping. Among the suite of inhibitors, b-VAD-AOMK inhibited human caspase-8 by 85% at a concentration of 20 µM (Figure 3.11); for comparison, complete inhibition of activity was observed by z-VAD-fmk at 20µM. Initial tests on H. volcanii cell extracts employed b-L-, b-R-, b-VEID-, D-, KL-, VEID-AOMK inhibitors at a range of concentrations (1, 5, 25, and 50 µM). A 77% reduction was observed in extracts treated with 50 µM b-VEID-AOMK, while the other inhibitors displayed modest (28-45%) inhibition compared to untreated cell extracts (Figure 3.12). Notably, significant inhibition (35-45%) of caspase activity was also observed with non-canonical, peptide-AOMK conjugates (i.e., b-R-AOMK, D-AOMK, and L-AOMK). The AOMK group has been reported to have some cross-reactivity with non-
caspase substrates but this cross-reactivity is reported to be significantly less than the fmk functional group (21).

The VEID-AOMK inhibitors effectively inhibited caspase-like activity in *H. volcanii* at concentrations >25 μM, with b-VEID-AOMK being a more potent inhibitor of activity at 77% versus 40% inhibition. Modest inhibition of caspase activity in *H. volcanii* cell extracts was previously observed with the specific IETD-, VEID-, and LEVD-fmk caspase inhibitors at 20 μM (10-40%) and 50 μM (46-59%) concentrations [see Figure 4 in (3)], perhaps suggesting an ancestral and less refined caspase active protease(s) in Archaea for which specific substrate preferences have not yet evolved. The broad-spectrum b-VAD-AOMK inhibitor was also tested on *H. volcanii* extracts at a range of concentrations (10 to 200 μM; Figure 3.13). The degree of inhibition was dose dependent with 200 μM b-VAD-AOMK showing the most complete inhibition at 77%. A similar degree of inhibition was observed in *H. volcanii* cell extracts for b-VAD-fmk at 20 μM, which differs in having an active fluoromethyl ketone group (as opposed to the acyloxy methyl ketone group) and that it was commercially obtained (Calbiochem). The large discrepancy in % inhibition for these two inhibitors with otherwise identical biotinylated substrate group (b-VAD) was surprising and likely stems from different relative chemical reactivities of the functional group moieties in our halophilic *H. volcanii* protein system.

Given these observations, b-VAD-AOMK was used for in situ trapping of caspase active proteins, in a manner similar to the aforementioned experiments with b-VAD-fmk, with b-L-AOMK serving as a negative control. The latter inhibitor has a more general leucine amino acid residue conjugated to AOMK and only inhibited caspase activity by 32%. The effectiveness of in situ trapping was monitored by Western blot analysis using an avidin-HRP antibody and MS-
compatible silver staining. Data is only shown for silver staining of the final recovered bead fraction for both inhibitors (Figure 3.14). As with the b-VAD-fmk and b-FA-fmk trapped samples, b-VAD-AOMK and b-L-AOMK bound treatments recovered similarly diverse proteins on the streptavidin-coated beads with most ranging between 75 - 100 kDa in size.

Identification of caspase-associated proteins by mass spectrometry

LC-MS/MS analyses of inhibitor-trapped proteins with b-VAD-fmk and b-FA-fmk served as an initial window into the diversity and identity of caspase-associated proteins (Table 1). A surprisingly high number (total of 269 proteins; ~5% of the H. volcanii proteome) of functionally different H. volcanii proteins were detected and identified from the LC-MS/MS peptide fragment data at a log-e score cutoff < -20. While many of the identified proteins were involved in amino acid and carbohydrate metabolism, several criteria were applied to identify a subset of proteins for subsequent analysis. Given the well documented caspase-4 activity in H. volcanii (3) and its possible connection to an ER stress-like response and UPR, recovered proteins with putative roles in stress- and degradation-related processes and complexes were targeted (Table 3.1). These included all three units of the archaeal thermosome chaperonins [cct1, cct2, and cct3 (22)] and components of the proteosomal protein degradation machinery [psm A, psmC, and panA (23)]. Other interesting proteins identified in the inhibitor pull-down experiments include a pair of universal stress proteins, a cell division protein 48 (Cdc48), translation factors α EF1 and α EF2, known to be associated with the unfolded protein response (UPR) in higher eukaryotes and, more recently, with SAMP2 in H. volcanii (24), a putative CBS domain pair protein, and DnaK, a chaperone protein. Most of these proteins, or their homologues, are
reported to be involved in protein folding or degradation in *H. volcanii* and other cell systems.

As previously mentioned, the b-FA-fmk inhibitor was used as a negative control to help refine and disentangle which recovered proteins were more specific to and enriched in the b-VAD-fmk caspase inhibitor fractions. Given b-FA-fmk was not effective at inhibiting caspase activity and instead has a reported specificity for cathepsin activity, it would target a distinctly different repertoire of trapped proteins. A comparative analysis of LC MS/MS data with both fmk-based inhibitors did not greatly reduce the number of candidate proteins; the same suite of proteins, and at similar levels of representation (as determined by spectral counts of peptides for recovered proteins) were also generally seen with in the b-FA-fmk pull downs despite a negligible reduction (1.5%) in caspase catalytic activity. *In situ* inhibitor trapping was repeated several times under varying conditions, for example change in salt of the buffer used in pull down in a range from 150 mM to 2M, with these inhibitors to see if the relative differences could be accentuated, and use of avidin versus streptavidin beads. However, despite consistent differences in their ability to inhibit caspase biochemical activity, the b-VAD-fmk and B-FA-fmk consistently ‘trapped’ the same suite of proteins and with similar spectral counts.

The puzzling and unexpected observations of approximately equivalent spectral counts might be explained by a few scenarios: 1) the biotin moiety and/or inhibitory functional groups (fmk or AOMK) impart an inherent non-specificity to the assay; 2) the very high salt environment (1.5 - 2.1 M) of our *H. volcanii* cell extracts and the corresponding biochemical properties of associated proteins is such that some proteins have a tendency to naturally ‘salt out’, clump and precipitate under our experimental conditions, ultimately being recovered in
bead fractions as an artifact; or 3) proteins with activities that are inhibited by both b-VAD-fmk (caspase-associated) and b-FA-fmk (cathepsin-associated) interact in protein complexes that were simultaneously pulled down. The fact that cell extracts were treated sequentially with both biotin (to remove proteins that would bind conjugated inhibitors at the biotin moiety and not at the functional group) and streptavidin-coated beads (to remove any proteins that inherently contain biotin) should have greatly reduced complications from scenario #1. Likewise, the fact that all of the biochemical procedures were performed in buffer containing a background 2 M NaCl should have minimized scenario #2. It is possible that the salt composition of the buffer could be critical. *H. volcanii* has a very high tolerance for MgCl$_2$ (1), likely reflective of the abundant MgCl$_2$ concentrations found in the Dead Sea, which constitute 50.8% of the anhydrous chlorides on a weight percentage basis, compared with 30.4% for NaCl (25). This is remarkably different from the composition of ocean seawater which is ~ 97% NaCl. Future work may address this issue. Given that scenarios #1 and #2 were largely controlled for, it is probable that recovered proteins are indeed components of interacting complexes and/or complexes that possess multiple proteolytic activities (e.g. proteasomes).

Additional support for this scenario comes from an examination of chymotrypsin/papain (LLVY-AMC) biochemical activity in the 80% and 100% (NH$_4$)$_2$SO$_4$ precipitated proteins as tested previously (3), along with corresponding b-FA-fmk inhibition of this activity (Figure 3.15). B-FA-fmk completely inhibited the LLVY-ase activity of *H. volcanii* in both the 80% and 100% (NH$_4$)$_2$SO$_4$ fractions, suggestive that the b-FA-fmk inhibitor effectively traps proteins with chymotrypsin/papain-like activities. *H. volcanii* proteasomes have been previously reported to have chymotrypsin-like activity, detected by hydrolysis of LLVY-fluorogenic substrates (18).
Notably, proteasomes in other cell systems have been reported to have chymotrypsin, trypsin, and caspase-like proteolytic activity for protein degradation (26, 27).

Independent inhibitor trapping experiments with b-VAD-AOMK and b-L-AOMK inhibitors recovered essentially the same pool of proteins as recovered with –fmk based inhibitors. There was a more defined distinction in the peptide spectral counts between proteins recovered with the different AOMK-based inhibitors, which helped to reinforce the enrichment of protein turnover-related proteins (i.e., thermosomes, cdc48, DNA chaperone, elongation factors) associated with caspase activity, as well as highlighted a few new promising candidates (Table 3.2).

Proteins exhibiting a difference of >10 in spectral counts in the b-VAD-AOMK compared to b-L-AOMK were subjected to additional analysis by Skyline quantification software (https://brendanxuw1.gs.washington.edu/labkey/_webdav/home/software/Skyline/%40files/tutorials/AbsoluteQuant-1_4.pdf) to determine if these spectral differences were significant. The Skyline software compared absolute abundance of distinct target peptides for each protein in the b-VAD-AOMK:b-L-AOMK trapped samples with ratios >1 noting significant enrichment for caspase inhibited proteins and, hence, a candidate for further exploration. This analysis identified elongation factor αEF-2 (HVO_0356) and ornithine cyclodeaminase (HVO_2879) as additional promising candidates (Table 3.3). Intriguingly, elongation factor αEF-2 has been identified in protein sampylation and proteosome function (28) and ornithine cyclodeaminase is involved in amino acid metabolism, catalyzing the conversion of ornithine to proline and in arginine and proline biosynthesis (29).
**Structural mining for caspase-associated H. volcanii proteins**

Given the *in situ* trapping approach was based on the diagnostic biochemical activity of caspase-like proteases, the suite of recovered *H. volcanii* proteins was further interrogated for structural homologues with known crystal structures of caspase active site structures. Hence, our approach was to combine our targeted purification with a ‘structure mining approach’ in order to identify proteins that have a similar active site folds to caspases. This active site, structural-based approach was further predicated on the fact that no obvious primary sequence homologues to known caspase proteins exist in the *H. volcanii* genome. To complicate matters, only 11 resolved crystal structures of *H. volcanii* proteins, representing only six distinct proteins, have been submitted to the Protein Data Bank (PDB; [http://www.rcsb.org/pdb/home/home.do](http://www.rcsb.org/pdb/home/home.do)). Consequently, putative tertiary structures were generated for our recovered proteins by homology modeling using Phyre2 (Figure 3.14), an online structure prediction tool that uses a variety of methods to find an accurate match for the query protein (9). The rationale was that *H. volcanii* proteins with active sites structures similar to caspase-4 would serve as promising candidates for downstream gene knock out experiments (see Chapter 4).

Resolved crystal structures for caspase-4 are not currently available in the PDB (30). However, human caspase-6 has a resolved crystal structure bound to its inhibitor [VEID- CHO; 3P4U Muellar et al., unpublished], which provided the opportunity to use its active site structures and motif as a structural alignment query. The structural alignment of the modeled caspase-4 tertiary structure, generated in Phyre2, was first confirmed with that of the inhibitor-bound caspase-6 crystal structure using TopMatch (see Methods). This provided confidence
that subsequent alignments of putative tertiary structures for *H. volcanii* proteins to that of caspase-4 would incorporate a known caspase active site. Structures for all the ~250 proteins that were seen in the pull downs were generated and evaluated from the calculated output generated by the TopMatch software, followed by visual analyses for maximum superimposition. This first level analysis indicated that the thermosome subunits *cct*1, *cct*2, and *cct*3 (*HVO* _0778, *HVO* _0455, and *HVO* _0133_ respectively), proteosome subunit alpha 1 *psmA* (*HVO* _1091_), cdc48 cell division control protein (*HV* _2700_), putative nuclease (*HVO* _2889_), and aminopeptidase (*HVO* _0836_) (Figure 3.16), among others [e.g., *ftsZ* (*HVO* _2204_), *dnaK* (*HVO* _1590_), and the translation elongation factors (*HVO* _0356_ and *HVO* _0359_); not shown)] were promising candidates.

Subsequent analyses used Cytoscape to quantitatively compare all 269 modeled structures in the active site with the inhibitor-bound caspase 6 crystal structure and the putative caspase fold comprising β sheets and turns (Figure 3.17) providing an alignment of the best matches using length of the query and percent identical residues. This analysis, combined with additional criteria (type of protein, spectral representation in MS data, etc.) was used to assess for logical candidates. This analysis showed that the putative nuclease and putative aminopeptidase were the best candidate proteins even though their detected peptide spectral counts were relatively low in the MS data (Table 3.4). Other proteins like an ATP synthase (*HVO* _0317_), while having high matches based on the selected criteria [i.e. length vs identical residues (87, 13)], were disregarded based on a high root mean square deviation from all the alpha C atom, as well as the fact that we were looking for a form of hydrolase not a synthase that would catalyze condensation reactions. Similarly, an arginine synthase ([*HVO* _0324_];
Length versus identical residues 85,9] was not selected as a candidate despite a lower root mean square deviation than the putative aminopeptidase because the protein was a synthase, had only one spectral count reported in pull downs with both B-VAD-fmk and B-FAfmk, and was not detected in AOMK pull downs (see below). The Conserved Domain Database search at NCBI (www.ncbi.nlm.nih.gov/Structure/cdd/cdd.html) was searched for conserved domains that might match the caspase protein domains, no matches were found with the caspase protein domains either in the H.volcanii proteins pulled down. Interestingly, the assessment of structural alignments via TopMatch, and the examination of relationships via cytoscape indicated a significant distance from caspase-6 structure aligned with itself and caspase-4 (Figure 3.17). Nevertheless, the putative nuclease and aminopeptidase were chosen as best-matched candidates for further testing.

As an independent diagnostic to identify caspase-like proteins, the LC MS/MS and MALDI-TOF-MS peptide data for cysteine residues that had conjugated b-VAD-fmk was searched, as would be reflected in a characteristic additional mass change. The VAD-fmk inhibitor has been reported to bind covalently and irreversibly to the active-site cysteine in most caspases (15). The molecular weight of b-VAD-fmk is 672.81 Da; binding to active site cysteine resulting in the loss of a water molecule (20.3 Da), making the expected net mass change 652.5 Da. Analyses were performed for proteins under both denaturing (proteins treated with SDS-PAGE buffer prior to trypsin digestion) and non-denaturing conditions (direct digestion of proteins with trypsin), but the diagnostic mass change could not resolved, even though there was successful documentation of effective inhibition of caspase activity by b-VAD-fmk. This analytical procedure was also tested on in situ trapped human caspase-8 as a positive
control with b-VADfmk and z-VAD-fmk (both of which abolish caspase activity; Figure 3.5).

Binding of z-VADfmk (MW = 453.5 Da) to the active site cysteine and the resulting loss of a water molecule yield a mass change of 433.2 Daltons. Binding of z-VAD-fmk did yield a detectable mass change of 433.2 Da on a cysteine residue in peptide 314-353 but no corresponding mass change was detected for b-VAD-fmk. Notably, the observed mass change was not detected on active site cysteine residue but on the cysteine residue flanking it. Despite repeating the mass spectrometry with a higher amount of human caspase enzyme and running the mass spectrometric analysis on non-denatured protein, a change in mass could not be detected (data not shown) even though it has been previously reported (31).

**Gel visualization of caspase active proteins**

A Cy5-bound VEID-AOMK fluorescent inhibitor tag (LE22), also provided by the Bogyo lab at Stanford University, was used to bind and visualize caspase-active proteins in *H. volcanii* using both native PAGE and SDS-PAGE. LE22 has previously been used on a human colorectal cancer cell line (COLO205) to identify initiator caspases that activate caspase-6 in apoptotic pathways (32). Successful binding in the *H. volcanii* system would provide useful information on protein size and further aid refining our search among candidate proteins. Native PAGE indicated a LE22-staining protein complex of ~25 kDa (data not shown), while denaturing SDS-PAGE revealed fluorescent bands between 10-15 kDa (Figure 3.18A). However, given a negative control sample consisting of only LE22 co-migrated with the stained *H. volcanii* extracts, indicative that a bulk of LE22 did not incorporate into caspase active proteins. If a small amount of LE22 did actually bind and label to caspase active proteins, it was below the level of
Human caspase-6 was also treated with LE22 as a positive control, since its preferred substrate contains the VEID tetrapeptide. The bulk of LE22 signal also migrated to a low MW position (~10-15 kDa), suggestive that it was in excess and did not incorporate into proteins. However, there was detectable fluorescence with an ~20 kDa and ~37 kDa protein bands in both native and denaturing PAGE, which correspond with the respective sizes of the active and inactive subunits (Figure 3.18B; data reported only for SDS-PAGE). The successful binding of human caspase 6 to LE22 was not due to the relatively low salt conditions (150 mM NaCl); successful binding was also observed at 150 mM, 750 mM and 1.5 M NaCl (data not shown), demonstrating that the high salt in H. volcanii cell extracts was not inhibitory to binding.

Section 3.5 CONCLUSIONS

In situ trapping using biotinylated fmk- and AOMK-based inhibitors was successfully employed in combination with genome-enabled proteomics and structural alignments to identify candidate proteins associated with the high, basal caspase activity in H. volcanii cell extracts. Our analysis revealed that a suite of diverse proteins were associated with this proteolytic activity, which is the dominant proteolytic activity in the cell. Collectively, these analyses identified thermosomes, chaperonins, proteasomes, a cell division protein, an ATPase, recently identified as an activator of proteasomal degradation, a putative nuclease, putative aminopeptidase, elongation factor αEF-2, and an ornithine cyclodeaminase as key proteins associated with caspase activity. These proteins were repeatedly identified by LC-MS/MS and either displayed differential peptide spectral counts with caspase-specific inhibitors or had
structural similarities to the active site of human caspases. Our results are consistent with previous observations on the nature of the caspase catalytic activity, namely that it is most similar to caspase 4, and linked to cellular stress responses. Our diagnostic biochemical work connects caspase activity in *H. volcanii* to specific stress-related protein complexes, such as those involved in the unfolded protein response. Our results are also consistent with observations in other cell systems that these complexes have a mixture of key proteolytic activities (chymotrypsin and papain). Future work will generate knockouts of these candidate genes to empirically test their relationship to caspase activity and to assess their physiological roles and link to UPR. We hypothesize that loss of this activity or reduction from a critical threshold level, and the impact on the proteins directly associated downstream, will have an important consequence to organismal fitness.
Section 3.6. **ACKNOWLEDGEMENTS**

We want to thank Stefan Senn (currently at University of Salzburg) for his help with the structural alignment analysis of *H. volcanii* proteins in cytoscape, Chris Brown for useful tips and discussions about protein purifications, and Liti Haramaty, Frank Natale, and Kevin Wyman for technical assistance. We also thank Michael Maniscalco and Kelly Bidle for providing the initial *H. volcanii* pellets for the purification. This research was supported in part by a grant from The Gordon and Betty Moore Foundation.

Author Contributions: M.S.P. designed and performed research, M.S.P, K.A.B., and K.D.B. analyzed data, S.S. analyzed the data for structural alignments, L.E.S. and M.B. synthesized and provided the AOMK inhibitors, M.S.P., K.A.B., and K.D.B. wrote the chapter.
Section 3.7 TABLES and FIGURES

*Table 3.1:* Rank order of a subset of *H. volcanii* proteins recovered in biotinylated fmk-based inhibitor trapping and identified by LC MS/MS analysis based on the sum total of spectral counts of unique peptides identified. Table lists a subset of proteins recovered based on their putative roles in stress pathways and protein folding and/or degradation, as well as proteins whose predicted tertiary structure had the strongest structural alignment to caspase active sites via TopMatch and Cytoscape analysis.

<table>
<thead>
<tr>
<th>Gene ID*</th>
<th>Putative Function</th>
<th>Protein log e</th>
<th># b-VAD-fmk</th>
<th># b-FA-fmk</th>
</tr>
</thead>
<tbody>
<tr>
<td>HVO_0778</td>
<td>cctA Thermosome subunit 3</td>
<td>-609.8</td>
<td>184</td>
<td>172</td>
</tr>
<tr>
<td>HVO_0133</td>
<td>cctA Thermosome subunit 1</td>
<td>-396.8</td>
<td>151</td>
<td>152</td>
</tr>
<tr>
<td>HVO_0455</td>
<td>cctB Thermosome subunit 2</td>
<td>-377.4</td>
<td>92</td>
<td>93</td>
</tr>
<tr>
<td>HVO_0359</td>
<td>Tuf translation elongation factor EF-1</td>
<td>-281</td>
<td>80</td>
<td>80</td>
</tr>
<tr>
<td>HVO_1955</td>
<td>aconitate hydratase putative</td>
<td>-317.9</td>
<td>64</td>
<td>56</td>
</tr>
<tr>
<td>HVO_0356</td>
<td>Translation elongation factor eEF-2</td>
<td>-223.2</td>
<td>36</td>
<td>34</td>
</tr>
<tr>
<td>HVO_1091</td>
<td>psmA proteasome subunit alpha1</td>
<td>-121.2</td>
<td>31</td>
<td>33</td>
</tr>
<tr>
<td>HVO_2384</td>
<td>CBS domain pair putative</td>
<td>-153.5</td>
<td>28</td>
<td>29</td>
</tr>
<tr>
<td>HVO_2204</td>
<td>ftsZ cell division GTPase</td>
<td>-137.5</td>
<td>24</td>
<td>24</td>
</tr>
<tr>
<td>HVO_0450</td>
<td>Molecular chaperone small heat shock protein</td>
<td>-83.3</td>
<td>16</td>
<td>19</td>
</tr>
<tr>
<td>HVO_1590</td>
<td>dnaK chaperone protein DnaK</td>
<td>-110.5</td>
<td>18</td>
<td>17</td>
</tr>
<tr>
<td>HVO_2156</td>
<td>universal stress protein A</td>
<td>-94.6</td>
<td>18</td>
<td>16</td>
</tr>
<tr>
<td>HVO_2889</td>
<td>putative nuclease</td>
<td>-71.4</td>
<td>13</td>
<td>16</td>
</tr>
<tr>
<td>HVO_1198</td>
<td>universal stress protein</td>
<td>-83.3</td>
<td>15</td>
<td>14</td>
</tr>
<tr>
<td>HVO_2700</td>
<td>cdo48 cell division control protein 48</td>
<td>-57.7</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>HVO_0850</td>
<td>PanA proteasome regulatory subunit</td>
<td>-45.9</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>HVO_2879</td>
<td>predicted ornithine cyclodeaminase</td>
<td>-46.9</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>HVO_2923</td>
<td>psmC Proteasome subunit alpha2</td>
<td>-35.9</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>HVO_0836</td>
<td>aminopeptidase</td>
<td>-27.3</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

*http://archaea.ucsc.edu/cgi-bin/hgGateway?db=haloVolc1

* # - sum total of the spectral counts of unique peptides for each identified protein. Confidence is provided by log e scores; 269 proteins had a log e < -20
Table 3.2: Rank order of a subset of *H. volcanii* proteins recovered in biotinylated AOMK-based inhibitor trapping and identified by LC-MS/MS analysis based on the sum total of spectral counts of unique peptides identified. Table lists an independent, comparative spectral analysis for a subset of recovered proteins compared to that with biotinylated fmk-based inhibitors (Table 3.1; rank order of proteins is maintained).

<table>
<thead>
<tr>
<th>Gene ID*</th>
<th>Putative Function</th>
<th>Protein log e</th>
<th># b-VAD-AOMK</th>
<th># b-L-AOMK</th>
</tr>
</thead>
<tbody>
<tr>
<td>§ HVO_0778</td>
<td>cctA Thermosome subunit 3</td>
<td>-701.2</td>
<td>297</td>
<td>249</td>
</tr>
<tr>
<td>§ HVO_0133</td>
<td>cctA Thermosome subunit 1</td>
<td>-447</td>
<td>171</td>
<td>162</td>
</tr>
<tr>
<td>§ HVO_0455</td>
<td>cctB Thermosome subunit 2</td>
<td>-399.9</td>
<td>95</td>
<td>79</td>
</tr>
<tr>
<td>§ HVO_0359</td>
<td>Tuf translation elongation factor EF-1</td>
<td>-268.1</td>
<td>90</td>
<td>69</td>
</tr>
<tr>
<td>HVO_1955</td>
<td>aconitate hydratase putative</td>
<td>-415</td>
<td>113</td>
<td>120</td>
</tr>
<tr>
<td>§ HVO_0356</td>
<td>Translation elongation factor eEF-2</td>
<td>-132.6</td>
<td>29</td>
<td>18</td>
</tr>
<tr>
<td>§ HVO_1091</td>
<td>psmA proteasome subunit alpha1</td>
<td>-150</td>
<td>33</td>
<td>39</td>
</tr>
<tr>
<td>HVO_2384</td>
<td>CBS domain pair putative</td>
<td>-161.4</td>
<td>21</td>
<td>20</td>
</tr>
<tr>
<td>§ HVO_2204</td>
<td>ftsZ cell division GTPase</td>
<td>-206.2</td>
<td>54</td>
<td>32</td>
</tr>
<tr>
<td>HVO_0450</td>
<td>Molecular chaperone small heat shock protein</td>
<td>-85.4</td>
<td>22</td>
<td>25</td>
</tr>
<tr>
<td>§ HVO_1590</td>
<td>dnaK chaperone protein DnaK</td>
<td>-204.5</td>
<td>47</td>
<td>43</td>
</tr>
<tr>
<td>§ HVO_2879</td>
<td>predicted ornithine cyclodeaminase</td>
<td>-121.8</td>
<td>38</td>
<td>15</td>
</tr>
<tr>
<td>HVO_2156</td>
<td>universal stress protein A</td>
<td>-139.6</td>
<td>31</td>
<td>27</td>
</tr>
<tr>
<td>HVO_2889</td>
<td>putative nuclease</td>
<td>-57.3</td>
<td>12</td>
<td>7</td>
</tr>
<tr>
<td>HVO_1198</td>
<td>universal stress protein</td>
<td>-79.5</td>
<td>22</td>
<td>10</td>
</tr>
<tr>
<td>§ HVO_2700</td>
<td>cdc48 cell division control protein 48</td>
<td>-125.3</td>
<td>27</td>
<td>15</td>
</tr>
<tr>
<td>HVO_0850</td>
<td>PanA proteasome regulatory subunit</td>
<td>-54.9</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>HVO_2923</td>
<td>psmA Proteasome subunit alpha2</td>
<td>-17.8</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>HVO_0836</td>
<td>aminopeptidase</td>
<td>-41.5</td>
<td>8</td>
<td>9</td>
</tr>
</tbody>
</table>

*http://archaea.ucsc.edu/cgi-bin/hgGateway?db=haloVolc1

# -Sum total of spectral counts of unique peptides for each protein identified. Confidence is provided by protein log e scores of peptide fragments; Note, the order of genes is the same as Table 1 though # detected is different from Table 1; §-Identifies proteins that had a >10 enrichment in spectral counts for b-VAD-AOMK inhibitor trapping. These proteins were subjected to Skyline analysis to verify if enrichment was statistically significant; ⌊⌋-DnaK chaperone protein was also analyzed as it has 54% identity and 70% similarity to yeast ER chaperone protein, BiP.
Table 3.3: Skyline absolute quantification analysis of a subset of *H. volcanii* proteins recovered in biotinylated AOMK-based inhibitor trapping. Table shows an independent spectral count analysis of proteins recovered with b-VD-AOMK or b-L-AOMK (from Table 3.2). Proteins with significant enrichment in b-VD-AOMK have median scores $>1.0$.

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Peptide sequence</th>
<th>b-VD-AOMK</th>
<th>b-L-AOMK</th>
<th>Total</th>
<th>Ratio</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>HVO_0133</td>
<td>AIESLGVVVR</td>
<td>52358328</td>
<td>47341788</td>
<td>49850058</td>
<td>0.90</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td>EFADSVGGR</td>
<td>1595998</td>
<td>1229343</td>
<td>1412670.5</td>
<td>0.77</td>
<td></td>
</tr>
<tr>
<td></td>
<td>EGISLR</td>
<td>11983206</td>
<td>10532040</td>
<td>11257623</td>
<td>0.88</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GGTGRTVESDLR</td>
<td>31525034</td>
<td>24029969.5</td>
<td>27777501.75</td>
<td>0.76</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IAATMNGK</td>
<td>34773940</td>
<td>34036592</td>
<td>34405266</td>
<td>0.98</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IFVVEDADAK</td>
<td>24693920</td>
<td>23733300</td>
<td>24213610</td>
<td>0.96</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MLVDGGQVYVTDGVTILK</td>
<td>1392083</td>
<td>3677780</td>
<td>2534931.5</td>
<td>2.64</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MOQGQPMIILGEDSQK</td>
<td>5683791</td>
<td>5341777</td>
<td>5512784</td>
<td>0.94</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SVTLRIR</td>
<td>11193592</td>
<td>12257830</td>
<td>11725711</td>
<td>1.10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TLAENAGLDPIDSLVLR</td>
<td>146868948</td>
<td>16210922</td>
<td>154489088</td>
<td>1.10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TQAIESATEAAMIRL</td>
<td>25565258</td>
<td>39210660</td>
<td>32387959</td>
<td>1.53</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TSGQDAOMNTIGA</td>
<td>69364912</td>
<td>65571832</td>
<td>67468372</td>
<td>0.95</td>
<td></td>
</tr>
<tr>
<td></td>
<td>VLPQPGPETTORS</td>
<td>2685533</td>
<td>2105147</td>
<td>2395340</td>
<td>0.78</td>
<td></td>
</tr>
<tr>
<td></td>
<td>VVSSLDDIEADLGFSVGQK</td>
<td>56611572</td>
<td>62908654</td>
<td>59760113</td>
<td>1.11</td>
<td></td>
</tr>
<tr>
<td>HVO_0317</td>
<td>DIVGAAINPYSR</td>
<td>2495542</td>
<td>2016480</td>
<td>2256011</td>
<td>0.81</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td>DLNSQRIPPINPLPSLR</td>
<td>4047115</td>
<td>3515041</td>
<td>3781078</td>
<td>0.87</td>
<td></td>
</tr>
<tr>
<td></td>
<td>EYPEEFIQGTSGAIDGMNLVR</td>
<td>444334</td>
<td>2637362.5</td>
<td>1540848.25</td>
<td>5.94</td>
<td></td>
</tr>
<tr>
<td></td>
<td>FLGETLK</td>
<td>2451881</td>
<td>1631771</td>
<td>2041826</td>
<td>0.67</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GYPGYMTDLAQQLYER</td>
<td>1323365</td>
<td>3338250.5</td>
<td>2330807.75</td>
<td>2.52</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LPYFSAGPHNLALQIAR</td>
<td>1365527</td>
<td>7901559</td>
<td>4633543</td>
<td>5.79</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MPVTQDLQGR</td>
<td>1706168</td>
<td>1617639</td>
<td>1661903.5</td>
<td>0.95</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SVVFTNLADDPAPV</td>
<td>72478776</td>
<td>58012756</td>
<td>65245766</td>
<td>0.80</td>
<td></td>
</tr>
<tr>
<td>HVO_0356</td>
<td>ALIAGQVK</td>
<td>808415</td>
<td>1286245</td>
<td>1047330</td>
<td>1.59</td>
<td>1.60</td>
</tr>
<tr>
<td></td>
<td>ASWNTENAGFR</td>
<td>1176568</td>
<td>1881734</td>
<td>1529151</td>
<td>1.60</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AVDGALVVVDAGAMPIETVLR</td>
<td>389651</td>
<td>4139770</td>
<td>2264710.5</td>
<td>10.62</td>
<td></td>
</tr>
<tr>
<td></td>
<td>FYITAEPQLDQVDSIK</td>
<td>817530</td>
<td>2002032</td>
<td>1409781</td>
<td>2.45</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GPAQVIPATR</td>
<td>2464542</td>
<td>3804178</td>
<td>3134360</td>
<td>1.54</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LLEPIQVR</td>
<td>1730741</td>
<td>2627152</td>
<td>2178946.5</td>
<td>1.52</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LVVDISDVAELR</td>
<td>3075899</td>
<td>5804562</td>
<td>4440230.5</td>
<td>1.89</td>
<td></td>
</tr>
<tr>
<td></td>
<td>VLSNQLO</td>
<td>1890680</td>
<td>2766294</td>
<td>3238487</td>
<td>1.46</td>
<td></td>
</tr>
<tr>
<td>HVO_0359</td>
<td>DMMGTIAAGK</td>
<td>12981130</td>
<td>17612744</td>
<td>15296937</td>
<td>1.36</td>
<td>1.25</td>
</tr>
<tr>
<td></td>
<td>EHVLAR</td>
<td>5895993</td>
<td>5414208</td>
<td>5655100.5</td>
<td>0.92</td>
<td></td>
</tr>
<tr>
<td></td>
<td>FNSDATYVPISAFEDNIAER</td>
<td>21082921</td>
<td>29828494</td>
<td>25455707.5</td>
<td>1.41</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IETGTNPGDNVFQPSDVGGVEK</td>
<td>39165791</td>
<td>42845633</td>
<td>41005712</td>
<td>1.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LDPASGEVAENPFKIK</td>
<td>4832544</td>
<td>5552775</td>
<td>5192659.5</td>
<td>1.15</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LPQYDVTISGTVPVGR</td>
<td>1857372.5</td>
<td>1762263</td>
<td>1809817.75</td>
<td>0.95</td>
<td></td>
</tr>
<tr>
<td></td>
<td>NMMITGAQADNVLVAAADDGAVPQTR</td>
<td>13076889.5</td>
<td>20962693.5</td>
<td>17019791.5</td>
<td>1.60</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SGDAAVTVPQXPSIEPSEIPELGSFAVR</td>
<td>880161</td>
<td>6404288</td>
<td>3642224.5</td>
<td>7.28</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TLGIGEILAVNK</td>
<td>3814199</td>
<td>4749756</td>
<td>4281977.5</td>
<td>1.25</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TVEMHHVEVDQAGPDNVGFNVR</td>
<td>8541452</td>
<td>15279204</td>
<td>11910328</td>
<td>1.79</td>
<td></td>
</tr>
<tr>
<td></td>
<td>VLEVNER</td>
<td>44522060</td>
<td>51272504</td>
<td>47897282</td>
<td>1.15</td>
<td></td>
</tr>
<tr>
<td>HVO_0455</td>
<td>AGLNVFTEVEDAFDAGVETAHAK</td>
<td>2167717</td>
<td>12727690</td>
<td>7447703.5</td>
<td>5.87</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ELLADLVR</td>
<td>886485</td>
<td>1341213</td>
<td>1113849</td>
<td>1.51</td>
<td></td>
</tr>
<tr>
<td></td>
<td>EQAVASASEAANLVLK</td>
<td>2671309</td>
<td>3135228</td>
<td>2903268.5</td>
<td>1.17</td>
<td></td>
</tr>
<tr>
<td>Sequence</td>
<td>M/G</td>
<td>p-MS</td>
<td>p-MS/MS</td>
<td>p-MS/p-MS</td>
<td>p-p-MS</td>
<td>p-MS/p-MS</td>
</tr>
<tr>
<td>----------</td>
<td>-----</td>
<td>------</td>
<td>---------</td>
<td>-----------</td>
<td>--------</td>
<td>-----------</td>
</tr>
<tr>
<td>EQLAVEAYADALELVPR</td>
<td>426994</td>
<td>2097072</td>
<td>1262033</td>
<td>4.91</td>
<td>0.98</td>
<td></td>
</tr>
<tr>
<td>FLDQEEAQKL</td>
<td>24473652</td>
<td>24852278</td>
<td>24662965</td>
<td>1.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GFNLASEK</td>
<td>14561371</td>
<td>13580778</td>
<td>14071074.5</td>
<td>0.93</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GIDDIAHYLAK</td>
<td>21549657.5</td>
<td>24404538</td>
<td>22977097.75</td>
<td>1.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GSTDHVDELER</td>
<td>4846047.5</td>
<td>3232152</td>
<td>4039099.75</td>
<td>0.67</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GVQDALDVVASTVADGR</td>
<td>5213215</td>
<td>4494657</td>
<td>4862936</td>
<td>0.86</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NYADSVGAR</td>
<td>1180201</td>
<td>883216</td>
<td>1031708.5</td>
<td>0.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>QGILAVR</td>
<td>1700226</td>
<td>1465057</td>
<td>1582641.5</td>
<td>0.86</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VLAGGGAEVELASR</td>
<td>4862699</td>
<td>9243080</td>
<td>7052889.5</td>
<td>1.90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HVO_0778</td>
<td>AIEDAVDVVVAIDK</td>
<td>3649348</td>
<td>3461174.5</td>
<td>3555261.25</td>
<td>0.95</td>
<td>0.98</td>
</tr>
<tr>
<td>ALADAGDVVVCFCTK</td>
<td>9022000.5</td>
<td>11503547</td>
<td>10262774</td>
<td>1.28</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AVTFLR</td>
<td>52073420</td>
<td>47473216</td>
<td>49773318</td>
<td>0.91</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EAISETAEATMIWKR</td>
<td>14549728.5</td>
<td>14627647.5</td>
<td>14588688</td>
<td>0.76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GGTAVDVAEL</td>
<td>31525034</td>
<td>24029969.5</td>
<td>27777501.75</td>
<td>0.98</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HVO_0960</td>
<td>HDATPQQIALAWLHR</td>
<td>24103</td>
<td>1279483</td>
<td>651793</td>
<td>53.08</td>
<td>1.67</td>
</tr>
<tr>
<td>HVGLSNVSVEQFDEAR</td>
<td>2417122</td>
<td>3329466</td>
<td>2872894</td>
<td>1.38</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ITGDDIEPDDVENAK</td>
<td>2566189</td>
<td>4295991</td>
<td>3431070</td>
<td>1.67</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LGFGAMR</td>
<td>1851326</td>
<td>3280240</td>
<td>2565783</td>
<td>0.77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LIGEALAPYPDDLVVATK</td>
<td>20657475.5</td>
<td>3455843</td>
<td>27607979.25</td>
<td>1.67</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NTGDWPKNGPDDYLR</td>
<td>5993285</td>
<td>8786451</td>
<td>7389868</td>
<td>1.47</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HVO_2204</td>
<td>ADLLGK</td>
<td>1247556</td>
<td>1824619</td>
<td>1536087.5</td>
<td>1.46</td>
<td>1.37</td>
</tr>
<tr>
<td>ALLVLAGPPEHLNR</td>
<td>8548318</td>
<td>13769319</td>
<td>11158818.5</td>
<td>1.61</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GGDYPGIAEK</td>
<td>2493789</td>
<td>2940997</td>
<td>2717393</td>
<td>1.18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ITSIR</td>
<td>1970724</td>
<td>2176809</td>
<td>2073766.5</td>
<td>1.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IYTEPVYGLGILPGSDEGGIYTLNAAR</td>
<td>2013328.5</td>
<td>10119940</td>
<td>6066634.25</td>
<td>5.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LTGGEDEPPDNLDTAHTTNR</td>
<td>19781784</td>
<td>22321068</td>
<td>21046926</td>
<td>1.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VAGVILLGTVTNSVR</td>
<td>5550471</td>
<td>10147317</td>
<td>7848894</td>
<td>1.83</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VLQGSR</td>
<td>4438573</td>
<td>5367782</td>
<td>4903177.5</td>
<td>1.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WIEEQTGMSMER</td>
<td>599842</td>
<td>1241478</td>
<td>920660</td>
<td>2.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HVO_2380</td>
<td>AGGGITYEDEGGTNEIQR</td>
<td>15342394</td>
<td>4619635</td>
<td>998104.5</td>
<td>0.30</td>
<td>1.06</td>
</tr>
<tr>
<td>AVANETNANFISVR</td>
<td>7368217</td>
<td>7826010</td>
<td>7597113.5</td>
<td>1.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dataset</td>
<td>Peptide</td>
<td>RAW Score</td>
<td>1663.44</td>
<td>2956.03</td>
<td>2396.89</td>
<td>1.78</td>
</tr>
<tr>
<td>---------</td>
<td>---------</td>
<td>-----------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>------</td>
</tr>
<tr>
<td>HVO_2700</td>
<td>AVANETSASFFSIAGPEIISK</td>
<td>166344.5</td>
<td>2956033.5</td>
<td>2309689</td>
<td>1.78</td>
<td></td>
</tr>
<tr>
<td></td>
<td>EGLSPDTR</td>
<td>1111434</td>
<td>773315</td>
<td>942374.5</td>
<td>0.70</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GPQLLSK</td>
<td>2187252</td>
<td>1663991</td>
<td>1925621.5</td>
<td>0.76</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GVLLYGPPGTGK</td>
<td>6395791</td>
<td>6823207</td>
<td>6609499</td>
<td>1.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LDPTLHLK</td>
<td>2179674</td>
<td>2500541</td>
<td>2340107.5</td>
<td>1.15</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LGIEPPGVLLHPGPPTGK</td>
<td>5537250</td>
<td>6322725</td>
<td>5929987.5</td>
<td>1.14</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LSPGDIIEIEGGK</td>
<td>4179897</td>
<td>4250580</td>
<td>4215238.5</td>
<td>1.02</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MNEVQLEVAK</td>
<td>1852136</td>
<td>1729507</td>
<td>1790821.5</td>
<td>0.93</td>
<td></td>
</tr>
<tr>
<td></td>
<td>VTWDVGGLEGPK</td>
<td>2101168</td>
<td>2290572</td>
<td>2195980</td>
<td>1.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AVANEDASFIHTISGPEIMSK</td>
<td>1245861</td>
<td>2356432</td>
<td>1801146.5</td>
<td>1.89</td>
<td></td>
</tr>
<tr>
<td>HVO_2759</td>
<td>EGLSPDTR</td>
<td>501647</td>
<td>1559190</td>
<td>1030418.5</td>
<td>3.11</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SIEPSALR</td>
<td>1718445</td>
<td>1286298</td>
<td>1502371.5</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td></td>
<td>VTIALPQNLR</td>
<td>2607025</td>
<td>2722930</td>
<td>2664977.5</td>
<td>1.04</td>
<td></td>
</tr>
<tr>
<td></td>
<td>VWPGYPEDQGTGIIR</td>
<td>3650159</td>
<td>4281005</td>
<td>3965582</td>
<td>1.17</td>
<td></td>
</tr>
<tr>
<td>HVO_2879</td>
<td>AFLEDLLTTPSGTEVAGQR</td>
<td>787777</td>
<td>1333884.5</td>
<td>1060830.75</td>
<td>1.69</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DIDDGDFVR</td>
<td>1345101</td>
<td>1186723</td>
<td>1265912</td>
<td>0.88</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GSANPHNPVALAR</td>
<td>6088236</td>
<td>6240511</td>
<td>6164373.5</td>
<td>1.03</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IGPIGGADR</td>
<td>3516325</td>
<td>2973466</td>
<td>3244895.5</td>
<td>0.85</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SGIPSINIPNR</td>
<td>2612665</td>
<td>2132582</td>
<td>2372623.5</td>
<td>0.82</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TGDADAFYTSR</td>
<td>6232065</td>
<td>5788830</td>
<td>6010447.5</td>
<td>0.93</td>
<td></td>
</tr>
<tr>
<td></td>
<td>VWTWSAEGLR</td>
<td>3244089</td>
<td>2653577</td>
<td>2948833</td>
<td>0.82</td>
<td></td>
</tr>
<tr>
<td></td>
<td>VWPQYPEDQGTGIIR</td>
<td>3650159</td>
<td>4281005</td>
<td>3965582</td>
<td>1.17</td>
<td></td>
</tr>
</tbody>
</table>

The scores in the b-VAD-AOMK and b-L-AOMK columns are RAW scores indicating abundance of each peptide graphed by skyline quantification software as the area covered by each peak in a Gaussian distribution (https://brendanxuw1.gs.washington.edu/labkey/_webdav/home/software/Skyline/%40files/tutorials/AbsoluteQuant-1_4.pdf). Ratio refers to the absolute abundance of each target peptide in b-VAD-AOMK spectral data to that in the b-L-AOMK. Median is the mean of all ratios for one protein.
Table 3.4: Rank order of the top twenty-five structural alignments for predicted tertiary structures of *H. volcanii* proteins to the crystal structure of the Caspase-6 inhibitor-bound active site (VEID-CHO; 3P4U) via TopMatch and Cytoscape analysis. Proteins listed for *H. volcanii* were recovered in biotinylated fmk-based inhibitor trapping and MALDI-TOF MS analysis. Some *H. volcanii* candidate proteins, while they had reasonable Cytoscape scores, were removed from consideration based on other criteria (e.g. number of unique peptides identified in mass spectrometry between VAD-fmk, FA-fmk, VAD-AOMK, and L-AOMK analysis, root mean square deviation from the Cα carbon atom, or putative role in the cell).

<table>
<thead>
<tr>
<th>Query</th>
<th>Putative Function</th>
<th>L</th>
<th>Is</th>
</tr>
</thead>
<tbody>
<tr>
<td>3p4u,A</td>
<td>3P4U.pdb</td>
<td>145</td>
<td>100</td>
</tr>
<tr>
<td>3p4u,A</td>
<td>Casp6.pdb</td>
<td>130</td>
<td>98</td>
</tr>
<tr>
<td>3p4u,A</td>
<td>caspase4.pdb</td>
<td>127</td>
<td>28</td>
</tr>
<tr>
<td>3p4u,A</td>
<td>Aminopeptidase HVO_0836</td>
<td>91</td>
<td>8</td>
</tr>
<tr>
<td>3p4u,A</td>
<td>Molybpterin Biosyn HVO_2305</td>
<td>88</td>
<td>7</td>
</tr>
<tr>
<td>3p4u,A</td>
<td>Arg Synth HVO_0324</td>
<td>85</td>
<td>9</td>
</tr>
<tr>
<td>3p4u,A</td>
<td>Putative Nuclease HVO_2889</td>
<td>83</td>
<td>11</td>
</tr>
<tr>
<td>3p4u,A</td>
<td>Aldehyde DH HVO_B0371</td>
<td>82</td>
<td>11</td>
</tr>
<tr>
<td>3p4u,A</td>
<td>Ribose ABC trans HVO_1401</td>
<td>79</td>
<td>9</td>
</tr>
<tr>
<td>3p4u,A</td>
<td>ATP syn HVO_0317</td>
<td>87</td>
<td>13</td>
</tr>
<tr>
<td>3p4u,A</td>
<td>Cys tRNA synthase HVO_1125</td>
<td>83</td>
<td>5</td>
</tr>
<tr>
<td>3p4u,A</td>
<td>Acetate CoA lig HVO_1236</td>
<td>80</td>
<td>5</td>
</tr>
<tr>
<td>3p4u,A</td>
<td>Cystathione Gamma Sulfate HVO_2946</td>
<td>79</td>
<td>11</td>
</tr>
<tr>
<td>3p4u,A</td>
<td>fabG reductase HVO_0884</td>
<td>77</td>
<td>13</td>
</tr>
<tr>
<td>3p4u,A</td>
<td>Csd Amintransfer HVO_0109</td>
<td>81</td>
<td>7</td>
</tr>
<tr>
<td>3p4u,A</td>
<td>Suc CoA synthase HVO_2465</td>
<td>78</td>
<td>8</td>
</tr>
<tr>
<td>3p4u,A</td>
<td>Succinyl CoA synthase HVO_2464</td>
<td>77</td>
<td>6</td>
</tr>
<tr>
<td>3p4u,A</td>
<td>Putative Dehydrogen HVO_1311</td>
<td>77</td>
<td>10</td>
</tr>
<tr>
<td>3p4u,A</td>
<td>Hydroxy acyl CoA dehydrogenase HVO_1444</td>
<td>76</td>
<td>4</td>
</tr>
<tr>
<td>3p4u,A</td>
<td>tnaA beta Eliminating Factor HVO_0009</td>
<td>73</td>
<td>11</td>
</tr>
<tr>
<td>3p4u,A</td>
<td>ATP synthase HVO_0316</td>
<td>82</td>
<td>6</td>
</tr>
<tr>
<td>3p4u,A</td>
<td>pyrG CTPsynth HVO_2624</td>
<td>76</td>
<td>11</td>
</tr>
<tr>
<td>3p4u,A</td>
<td>Hydroxyacyl CoA dehydrogen HVO_2815</td>
<td>75</td>
<td>7</td>
</tr>
<tr>
<td>3p4u,A</td>
<td>Prolyl Oligopep HVO_0829</td>
<td>74</td>
<td>9</td>
</tr>
</tbody>
</table>

“L” is the length of the subject covered with respect to the query, 3P4U; “Is” is the % identical residues in the alignment generated; “Er” is the root mean square deviation from the Cα carbon atom.
Figure 3.1: Caspase specific activity during initial purification with ammonium sulfate [(NH$_4$)$_2$SO$_4$]. Substantial protein-normalized caspase activity was observed in all (NH$_4$)$_2$SO$_4$ precipitated samples, as well as the post-100% soluble fraction, indicating that either a suite of proteins with different ‘salting out’ properties are responsible for the observed activity, or the responsible protein has a very wide (NH$_4$)$_2$SO$_4$ tolerance threshold. Activity was maximal in the 100% (NH$_4$)$_2$SO$_4$ fraction (*), so it was chosen as source material for downstream purification procedures via inhibitor trapping with fmk-based inhibitors. Error bars represent the standard deviation between technical triplicates.
Figure 3.2: Visualization of total protein profiles during initial (NH₄)₂SO₄ purification.

Coomassie-stained SDS-PAGE of total protein from crude and cleared cell extracts, (NH₄)₂SO₄ precipitated and post-100% (NH₄)₂SO₄ soluble proteins (see methods) showing a rich array of diverse proteins in each fraction and no obvious changes in the protein profile during (NH₄)₂SO₄ procedure. The presence of very similar profiles across all fractions may also help explain the observation of their reasonably high caspase specific activity (>1000 nmol h⁻¹ mg protein⁻¹). There was a slight increase in staining intensity for the 100% and soluble fractions, even though lanes were loaded with equal protein, with these samples displaying the highest two caspase activities measured.
Figure 3.3: Visualization of proteins fractionated by size exclusion chromatography from *H. volcanii* cleared cell extracts. Cleared cell extracts were subjected to size exclusion purification using high-pressure liquid chromatography. Bars above 3 wells correspond to a 1 min window indicative of fractions collected every 20 sec. Each fraction was tested for caspase activity. Highest activity was reported in fractions collected at 12-12’ 20” and 12’20” to 12,40” highlighted (▼). The two lanes indicate proteins in the relatively broad range of ~30-100 kDa, not reflective of a refined purification.
Figure 3.4: Test of efficacy of fmk-based inhibitors on caspase specific activity in H. volcanii.

Ten μg of protein from the 100% (NH₄)₂SO₄ fraction was treated with 50μM each of z-VAD-fmk, b-VAD-fmk, and b-FA-fmk, the latter two having a biotin conjugate for use in downstream inhibitor trapping procedure. Activity of each fraction is expressed as the mean % activity relative to the uninhibited control. The percent inhibition is depicted above each bar. Note effective inhibition only for both VAD inhibitors including biotinylated version, suggesting that it will be effective in inhibitor trapping of caspase active fractions. In contrast, b-FA-fmk was not effective at inhibiting activity; this is consistent with its reported specificity for cathepsin activity, not caspase activity. Error bars represent percent standard error.
**Figure 3.5 Efficacy of fmk-based inhibitors on human caspase-8 activity.** Recombinant, purified human caspase-8 was completely inhibited by 20 μM of z-VAD-fmk and b-VAD-fmk at 37°C compared to the untreated enzyme. In contrast, activity increased by 26% relative to untreated control on treatment with b-FA-fmk, a cathepsin and papain inhibitor. Activity of each fraction is expressed as the mean % activity relative to uninhibited control. Each assay was done in technical triplicates and the error bars indicate percent standard error.
Figure 3.6: Visualization of b-VAD-fmk-based inhibitor trapping of human caspase-8.
Recombinant purified human caspase-8 was subjected to in situ inhibitor trapping, where it was incubated with b-VAD-FMK and captured using avidin monomeric beads. Recovered beads were washed with buffer followed by elution with either 2 mM biotin or a low pH glycine buffer (see lane designations; purification flow is indicated by arrow). Recovered samples from the trapping procedure were probed with either monoclonal antibody to human caspase-8 (A) or Avidin-HRP (B). The former assessed whether human caspase-8 (*) itself was being recovered in bead fractions; the latter assessed whether the biotinylated inhibitor was binding to human caspase-8 and ultimately being recovered in the bead fraction. Both antibodies revealed the ~30 kDa heterodimer and the ~20kDa monomer of human caspase-8. The 11 kDa monomer of human caspase-8 is revealed only by the monoclonal antisera and not Avidin-HRP, indicating binding of b-VAD-fmk to the active site containing ~20 kDa monomer. These two independent methods of detection corroborated each other and showed that human caspase-8 is bound to b-VAD-fmk and recovered in the bead fraction.
**Figure 3.7:** Efficacy of biotinylated-fmk-based inhibitors on caspase specific activity in pre-cleared *H. volcanii* (NH₄)₂SO₄ cell extracts prior to inhibitor trapping with streptavidin beads. Incubation of caspase-active cleared extracts with 50 μM of biotinylated fmk-based inhibitors, b-VAD-fmk, and b-FA-fmk, for 1 h at 42°C, served as a diagnostic proxy for their respective effectiveness in inhibitor trapping of caspase active proteins. b-VAD-fmk inhibited caspase specific activity by ~87% compared to no inhibition by b-FA-fmk. Activity of each fraction is expressed as the mean % activity relative to the uninhibited control. These extracts correspond to the inhibitor trapping and proteomic data reported in Table 1.
Figure 3.8: Visualization of b-VAD-fmk-based inhibitor trapping of caspase-associated proteins in H. volcanii cell extracts. Partially purified H. volcanii cell extracts (100% (NH₄)₂SO₄ fractions) that were enriched in caspase-specific activity were subjected to in situ inhibitor trapping, where they were incubated with b-VAD-FMK and captured using avidin monomeric beads. Recovered beads were washed with buffer followed by elution with either 2 mM biotin or a low pH glycine buffer (see lane designations; purification flow is indicated by arrow). Recovered samples from the trapping procedure were probed with Avidin-HRP (A) and silverSNAP silver staining (B). Avidin-HRP probing identified biotin-containing proteins, while silver staining revealed the total pool of proteins recovered during the procedure. Prior to the in situ trapping, extracts had first been subjected to ammonium sulfate precipitation with caspase-enriched fractions being identified by LEVD-AMC cleavage. These fractions were ‘pre-cleared’ with avidin-conjugated beads prior to the addition of b-VAD-FMK in order to remove naturally occurring biotin-containing proteins. Note the presence of a prominent ~75 kDa biotin-containing protein that was removed onto beads during pre-clearing step. A variety of proteins were recovered on avidin beads, some of which were eluted with biotin and acidic glycine treatment.
Figure 3.9: Visualization of b-FA-fmk-based inhibitor trapping of non-caspase-associated proteins in \textit{H. volcanii} cell extracts. Partially purified \textit{H. volcanii} cell extracts (100% (NH$_4$)$_2$SO$_4$ fractions) that were enriched in caspase-specific activity were subjected to \textit{in situ} inhibitor trapping, where they were incubated with b-FA-FMK and captured using streptavidin magnetic beads. Recovered beads were washed and proteins eluted with biotin followed by glycine (see lane designations; purification flow is indicated by arrow). Recovered proteins were either visualized by silver staining (A; for total proteins recovered) or with Avidin-HRP (B; for identifying biotin-containing proteins and those successfully ‘trapped’ by FA-fmk). A variety of proteins were recovered on avidin beads, some of which were eluted with biotin and acidic glycine treatment. These proteins were submitted for LC MS/MS analysis at the Rutgers’ Biological Mass Spectrometry Facility.
Figure 3.10: Visual comparison of biotinylated fmk-based inhibitor trapping. Western blot (probed with avidin-HRP) of pre-cleared cell extracts either: untreated, (pre-cleared extract); bound to b-VAD-fmk (+b-VAD-fmk) and b-FA-fmk (+b-FA-fmk). Also shown are captured proteins either removed from competitive elution with biotin (Biotin elute) or remaining on beads (beads) (see lane designations; purification flow is indicated by arrows). Note that the untreated, \( H. \) *volcanii* precleared cell extracts contains a naturally biotinylated protein at \(~75\) kDa; this has been consistently observed on Western blots probed with avidin-HRP. The substantial increase in proteins reacting to avidin-HRP in pre-cleared extracts treated with b-VAD-fmk and b-FA-fmk inhibitors demonstrates that a suite of abundant proteins, possibly part of a large complex, incorporate these inhibitors and that these proteins are successfully recovered with streptavidin magnetic beads.
Fig 3.11 Efficacy of AOMK-based inhibitors on the activity of human caspase-8. When challenged with a variety of AOMK-based protease inhibitors (b-L-AOMK, b-R-AOMK, b-VAD-AOMK, D-AOMK, KL-AOMK; 20 μM incubated at 37°C for 1 h), human caspase-8 activity was only inhibited (~85%) by b-VAD-AOMK. For comparison, treatment with 20μM of z-VAD-fmk, completely inhibited activity and was used as a positive control of inhibitor efficacy.
Figure 3.12 Efficacy of AOMK-based inhibitors on caspase activity of *H. volcanii* cell extracts.

When challenged with a variety of AOMK-based protease inhibitors (b-L-AOMK, b-R-AOMK, b-VEID-AOMK, D-AOMK, KL-AOMK, VEID-AOMK; 50 µM; incubated at 42°C for 1 h), caspase activity in *H. volcanii* cell extracts was inhibited (~77%) by b-VEID-AOMK. Modest inhibition (28-45%) was also observed for the other inhibitors, which was puzzling given they are thought to be more generic.
Figure 3.13 Efficacy of b-VAD-AOMK on caspase activity of *H. volcanii* cell extracts. Caspase activity in *H. volcanii* cell extracts was challenged with the biotinylated b-VAD-AOMK inhibitor at a range of concentrations (10, 20, 50, 100, and 200 µM). The degree of inhibition was dose dependent with 200 µM B-VAD-AOMK concentration showing the most complete inhibition at 77%. A similar degree of inhibition was observed in *H. volcanii* cell extracts for b-VAD-fmk at 20 µM, which differs in its active fluoromethyl ketone group (as opposed to the acyloxymethyl ketone group) and that it was commercially obtained (Calbiochem). The large discrepancy in % inhibition for these two inhibitors with otherwise identical biotinylated substrate group (b-VAD) was surprising and likely stems from different relative chemical reactivities of the functional group moieties.
**Figure 3.14 Visualization of AOMK-based inhibitor trapping of proteins from *H. volcanii* cell extracts.** Partially purified *H. volcanii* cell extracts (100% (NH₄)₂SO₄ fractions) were subjected to *in situ* inhibitor trapping where they were incubated with either b-VAD-AOMK and b-L-AOMK and captured using streptavidin magnetic beads. After several rounds of washing, proteins on recovered beads were subject to 8-16% SDS-PAGE and visualized with by silver staining. These proteins were submitted for LC MS/MS analysis at the Rutgers’ Biological Mass Spectrometry Facility to identify AOMK-based inhibitor bound proteins.
Figure 3.15 Efficacy of b-FA-fmk inhibitor on the hydrolysis of LLVY-AMC activity in the H. volcanii 80% and 100% (NH₄)₂SO₄ salt precipitated samples. Previously purified 80% and 100% fractions (Figure 3.1) were tested for LLVY-AMC hydrolysis at 60°C as tested previously (60). Efficacy of b-FA-fmk, an established papain and cathepsin inhibitor, was tested for efficacy of inhibition of the hydrolysis of these trypsin/cathepsin (ARR-AMC), and papain/chymotrypsin (LLVY-AMC) substrates. B-FA-fmk completely inhibited the hydrolysis of LLVY-AMC in both the 80% and 100% (NH₄)₂SO₄ fractions.
Figure 3.16 Visualization of structural alignments of select *H. volcanii* candidate proteins with the crystal structures of human caspases. (A) Alignment overlay of the predicted caspase-4 structure with resolved, inhibitor-bound caspase-6 crystal structure in PDB. (Caspase-6: blue and orange; caspase-4: green and red; orange and red are superimposed regions). This was done first to demonstrate the close coherence of the modeled caspase-4 structure (for which there is no crystal structure) with that of caspase-6 (for which there is a resolved crystal structure). Subsequent alignment overlays with *H. volcanii* candidate proteins were done with the modeled caspase-4 tertiary structure, given caspase-4-like hydrolysis (LEVD-AMC) was found to be highest in *H. volcanii* cell extracts. (B-F) Structural alignment overlays of (B) thermosome subunit 3 cct3, (C) proteasome subunit alpha 1 psmA, (D) cdc48 cell division control protein 48, (E) putative nuclease, and (F) aminopeptidase with the predicted tertiary structure of caspase-4 (caspase-4: blue and orange; *H. volcanii* proteins: green and red; orange and red are superposition areas. Predicted protein structures were determined using web-based Phyre2 server (118). While the superimpositions shown here are analyses between predicted caspase-4, the structural alignments using TopMatch (https://topmatch.services.came.sbg.ac.at/) and Cytoscape (www.cytoscape.org) were done with the resolved caspase6 with VEID_CHO structure (PDB id: 3P4U; see Figure 3.14).
Figure 3.17 Structural alignment comparison of inhibitor-trapped *H. volcanii* proteins with the known modeled, active site tertiary structure of caspases. Tertiary protein structures for inhibitor-trapped *H. volcanii* proteins were generated in TopMatch and parsed in Cytoscape, which generates alignment comparisons on a structure based-phylogenetic tree. All proteins identified by MALDI-TOF-MS were compared to better elucidate their relative structural relationship to 3P4U (caspase-6 bound to VEID-CHO). Each individual protein was treated as a node and the connectivity of this node to 3P4U (via edges that represent pairwise alignment) is quantified based on the length of alignment and the number of identical residues. The identities of the highest-ranking *H. volcanii* proteins are labeled above, along with the position of caspase proteins. The best alignments have higher x,y values that is the proteins with a high alignment length versus maximum number of identical residues. Top aligned proteins are highlighted and presented in Table 3.2 and were further investigated for caspase activity and physiology.
**Figure 3.18:** Gel visualization of caspase-active proteins by incubation with a fluorescently labeled AOMK-based inhibitor. LE22, a Cy5-labeled VEID-AOMK inhibitor, was used to visualize binding with caspase-like proteins in both *H. volcanii* cell extracts (A) and human caspase-6 (B) on precast 8-16% native Tris-HCl gels visualized on a Biorad Chemidoc MP system (A) or Typhoon 9410 variable mode imager (B). *H. volcanii* cell extract was treated with either LE22, DMSO only (LE22 was dissolved in DMSO) or incubated with 20 µM z-VAD-fmk (as a negative control since it does not contain the Cy5-labeled conjugate and cannot be visualized in this manner) prior to PAGE analysis. Cy5 fluorescence was observed LE22-treated *H. volcanii* cell extract and LE22 only lanes, but they co-migrated to very low molecular weights (~10 kDa), indicative that a bulk of LE22 did not incorporate into caspase active proteins. If a small amount of LE22 did actually bind and label to caspase active proteins, it was below the level of detection. For human caspase-6 control samples, the bulk of LE22 signal also migrated to a low MW position (~10-15 kDa), suggestive that it was in excess and did not incorporate into proteins. However, there was detectable fluorescence with an ~20 kDa and ~37 kDa protein bands, which correspond with the respective sizes of the active homodimer and pro-caspase heterodimer respectively.
Section 3.8 REFERENCES


CHAPTER 4

Phenotypic characterization of *Haloferax volcanii* mutants for caspase activity and growth under unfolded protein stress

Mansha Seth-Pasricha¹, Kelly A. Bidle², Kay D. Bidle¹

¹Department of Marine and Coastal Sciences, Rutgers University, 71 Dudley Road, New Brunswick, NJ 08901
²Department of Biology and Behavioral Neuroscience, Rider University, Lawrenceville, NJ
Section 4.1 **ABSTRACT**

Inhibitor-based, *in situ* trapping combined with genome-enabled proteomics and structural alignments identified a suite of diverse proteins that were associated with high, basal caspase activity in *H. volcanii* and putatively connected to stress-related protein complexes, such as those involved in the unfolded protein response (UPR). Using available mutants and well-established genetic disruption protocols, along with chemical treatments with the UPR-inducing chemical canavanine, a handful of these proteins (thermosomes, chaperonins, proteasomes, cell division protein, a putative nuclease, and a putative aminopeptidase) were tested for their direct link to caspase activity and UPR. Mutations in this suite of genes had variable impacts on caspase activity with mutants in a proteosome subunit (*ΔpsmA*; 100% increase) and a cell division protein (*Δcdc48*; 70% decrease) with links to ubiquitin and ER-associated degradation representing opposite ends of the caspase activity spectrum. Taken together, these findings suggest that non-optimal caspase levels (high or low compared to wild-type) critically compromise fitness in *H. volcanii*, as measured by cellular growth in response to canavanine, an arginine analogue that promotes protein unfolding. Our findings suggest that caspase activity is part of an ‘interactase’ or a protein complex and is involved in proper proteasome function.
Section 4.2 *INTRODUCTION*

The presence of very high, basal caspase 4-like activity in the model haloarchaeon, *Haloferax volcanii* has now been well established [Chapter 2; (1, 2)]. Interestingly, in higher eukaryotes, this activity is associated with the unfolded protein response (UPR) (3, 4), a cellular stress response related to the endoplasmic reticulum (ER) and conserved from yeast to mammals. A major player in UPR is a nano-compartmentalized protein degradation complex known as the proteasome, which is found in all domains of life and functions to degrade inaccurately synthesized, misfolded, or damaged proteins. This quality control system of the cell also regulates processes like cell division and DNA repair by either destroying regulatory proteins during key conditions or cleaving precursor proteins to generate their active counterparts (5). In light of high basal caspase 4-like activity in *H. volcanii* that is associated with physiological stress, it is possible that caspase-associated proteins may have emerged in early evolving microbes like Archaea, for essential housekeeping functions by participating in the cellular stress responses like UPR. *H. volcanii* encodes five protein components that make up its proteasome, psmA (α₁ subunit), psmB (β subunit), psmC (α₂ subunit) and the associated Pan proteins (panA and panB) (6-8). The *H. volcanii* proteasome modifies unfolded proteins by ‘sampylation’, which is hypothesized to target them for degradation (8, 9); a process similar to ubiquitination (9, 10).

Inhibitor-based, *in situ* trapping, in combination with genome-enabled proteomics and structural alignments identified a suite of diverse proteins that were associated with high, basal caspase activity in *H. volcanii* and are putatively connected to stress-related protein complexes,
such as those involved UPR (Chapter 3). These included chaperonins, thermosomes, proteasomes, a cell division protein, an ATPase (recently identified as an activator of proteasomal degradation), a putative nuclease, putative aminopeptidase, elongation factor αEF-2, and an ornithine cyclodeaminase. Many comprise a proteostasis network (PN), consisting of diverse cellular components that are involved in maintaining a healthy proteome balance. This PN comprises various interconnected pathways during cellular stress conditions, like the heat shock response (HSR), endoplasmic reticulum-associated UPR, and mitochondrial UPR. Proper folding of newly synthesized proteins and their conformational maintenance are essential aspects for maintaining proteostasis, as are cellular concentration, localization, and regulation in response to various stimuli.

Molecular chaperones play a key role in PN by monitoring the extent of incorrectly folded proteins, and serving as structural stability factors. There are two types of chaperone proteins: Group I (bacterial), and Group II. The Group II chaperonins are the archaeal thermosomes (cct; chaperonin-containing TCP-1), and their eukaryotic homolog, tailless complex polypeptide (TCP-1) ring complex. Components of thermosomes (cct1, cct2, cct3) and proteasomes (e.g., psmA, panA) have a well-established role in proper protein folding and/or protein turnover and degradation (5, 11, 12). Likewise, the elongation factor αEF-2 has been identified in protein sampylation and proteasome function (13), while ornithine cyclodeaminase is involved in amino acid metabolism, catalyzing the conversion of ornithine to proline and in arginine and proline biosynthesis (14). Cdc48, a cell division control 48 protein is a conserved ubiquitin-selective chaperone protein in eukaryotes from yeast to mice that unfolds substrate proteins for degradation (15). Cdc48 was recently identified to associate with the 20S
proteasome and participate in protein unfolding in the thermophilic archaean, *Thermoplasma acidophilum* (16).

Here, we utilize knockouts of a subset of the aforementioned candidate genes, as well as for others who showed active structural homology to known caspase proteins but do not have a documented role in PN, namely a putative nuclease and putative aminopeptidase, to empirically test their relationship to caspase activity and to assess their physiological roles and link to UPR. Mutant strains were tested for both caspase activity and their growth dynamics in response to canavanine, an arginine analog that incorporates into nascent polypeptides during translation and initiates protein unfolding. Notably, MSP1 and Δcct1Δcct3 double mutants displayed a significant loss of caspase activity, and nearly all mutant strains exhibited a severe growth defect in the presence of canavanine and upon initiation of unfolded protein stress. Taken together, these results indicate a strong biochemical and cellular connection between caspase activity and the UPR in *H. volcanii*.

Section 4.3 **MATERIALS AND METHODS**

**Strains, plasmids, and growth conditions.** Strains, oligonucleotide primers, and plasmids used for cloning are listed in Tables 4.1 and 4.2. *H. volcanii* was routinely grown in rich medium containing (per liter) 125 g NaCl, 45 g MgSO$_4$·6H$_2$O, 45 g MgCl$_2$·7H$_2$O, 10 g KCl (all salts from Sigma), 5 g tryptone (Difco), and 3 g yeast extract (Difco). Media was supplemented with 10% (wt/vol) filter sterilized CaCl$_2$ after autoclaving. Solid medium contained 15 g of Bacto Agar per liter. For Casamino acid medium (Hv-Ca), tryptone and yeast extract were replaced by casamino acids (Difco) at a final concentration of 0.5% wt/vol (6). *H. volcanii* minimal media (Hv-Min) was
prepared as described (17); each 330 ml of media contained 110 ml distilled water, 200 ml 20% salt water, 10 ml of 1M Tris HCl pH 7.5, 8.5 ml Hv-Min carbon source (60% lactic acid, succinic acid, glycerol), 4 ml Hv-Min salts (NH₄Cl, CaCl₂, and trace elements: MnCl₂·4H₂O, ZnSO₄·7H₂O, FeSO₄·7H₂O, CuSO₄·5H₂O), 650 μl 0.5 M KPO₄ buffer, and 300 μl thiamine and biotin. When required, uracil was added at a final concentration of 10 or 50 μg/μl (for pop-out selection or medium supplementation, respectively), 5-FOA at 50 μg/μl and canavanine at 11 μM. E. coli Top10 cells were grown in LB media supplemented with 50 μg/ml kanamycin and 100 μg/ml ampicillin, as required. E. coli INV 110 cells (dam⁻) were used for final passage of plasmids prior to transformation in H. volcanii.

**Chromosomal knockouts.** H. volcanii knockout strains were created using established gene disruption mutagenesis (“pop-in/pop-out” method) protocols that rely on transformant selection against a ΔpyrE2 background (18) using constructs made in pTA131, a pBluescript-based suicide plasmid containing the pyrE2 gene as a selectable marker (19). These constructs each contain a portion of the 5’ region of the gene to be mutated fused to a portion of the 3’ end of the gene, creating a deletion within the middle of each gene. To facilitate the creation of these constructs, PCR was used to amplify both the 5’ upstream and 3’ downstream regions of interest, with the resulting PCR products subsequently being cloned into pCR2.1-TOPO (Invitrogen). Positive clones were then digested with the enzyme EcoRI, gel purified, and ligated together. Ligation products were screened using the 5’ upstream gene’s forward primer and 3’ downstream gene’s reverse primer to amplify a sequence that contained the 5’ region concatenated to the 3’ region in the proper orientation (i.e., only correctly ligated products,
those that run 5′ to 3′ with a truncation in the middle, will amplify). Following this, resulting PCR products were cloned into pCR2.1-TOPO, and positively identified clones were digested with the enzymes XhoI and BamHI to facilitate cloning into the suicide vector pTA131.

Simultaneously, each of the full-length genes, along with ~200 bp of flanking DNA on each 5′ and 3′ end, were cloned into the autonomously replicating shuttle vector pTA354 (20). These clones served both to monitor transformation efficiency during knockout strain construction, as well as to complement, in trans, the single gene knockouts in subsequent experiments. All constructs were subjected to DNA sequencing analysis prior to H. volcanii transformations to verify their DNA fidelity.

**H. volcanii transformations.** The H. volcanii uracil auxotrophic strain H26 (ΔpyrE2; (19)) was transformed with each construct using established methods (21). Positive transformants, displaying uracil prototrophy on selective medium, were subsequently patched onto Hv-CA plates (Hv-CA medium contains 0.5% (w/v) casamino acids, replacing the tryptone and yeast extract from the standard H. volcanii growth medium described above) containing 5-fluoroorotic acid (5-FOA), as uracil auxotrophs are unable to convert 5-FOA to its toxic analog, 5-fluorouracil. Thus, transformants that have lost the plasmid through homologous recombination events display 5-FOA resistance. Ura⁻ 5-FOAᵦ colonies were screened via PCR to determine if they harbored the wild type or deleted version of each gene. The resulting knockout strains were confirmed via PCR and DNA sequencing analysis.

**Complementation of gene knockouts.** Following knockout strain construction, full-length gene
sequences were transformed into *H. volcanii* for complementation analyses. For these experiments, only the cdc48 knockout was transformed with the autonomously replicating shuttle vector harboring the wild-type gene (Table 4.1). Following successful transformation, clones were assayed for a restored caspase activity phenotype.

**Caspase activity and protease assays.** Caspase activity assays and protease assays were performed as described in Chapter 3.

**Growth Studies.** H26 and all mutant strains were streaked on fresh agar plates containing rich media with optimal 2.1M salt (1) and incubated at 42°C. Fresh colonies were picked and inoculated into 2.1M media and grown aerobically, with shaking, at 42°C until mid-exponential phase. Optical Density (O.D.) was measured at 600 nm on a NanoDrop spectrophotometer. Starter cultures for growth studies were initiated at equal starting O.D.600. Growth studies were performed in triplicate for each of the different types of growth medium (e.g., minimal media with uracil, minimal media with uracil and canavanine) and all cultures were grown aerobically, with shaking, at 42°C. Growth was monitored over 60 hours (i.e., the point at which wild-type enters late stationary phase/early death in minimal medium).

Section 4.4 **RESULTS AND DISCUSSION**

**Rationale for genes chosen for knockout studies**

A fundamental aspect of the research in this chapter was to either obtain, create, and phenotype strains of *H. volcanii* that had mutations in genes whose proteins were recovered in
previous fmk- and AOMK-based inhibitor trapping pull down experiments (Chapter 3). In order to better guide these studies, the focus was on targeting a promising subset of genes whose proteins have putative roles in cellular stress and/or had structural homology with caspase active sites (as implicated in the Cytoscape structural alignment analyses to inhibitor-bound caspase 6). Using these criteria, the subset of proteins included three thermosome genes [cct1 (HVO_0133), cct2 (HVO_0455), and cct3 (HVO_0778)], genes encoding subunits in the H. volcanii proteasome [psmA (HVO_1091) and panA (HVO_0850)], a cell division control 48 protein [cdc48 (HVO_2700)], a putative aminopeptidase (HVO_0836), and a putative nuclease (HVO_2889). Fortunately, unrelated studies had previously generated mutant strains of thermosomes and proteasomes in H. volcanii (6, 12) and they were available for our use [kindly provided by Peter Lund at the University of Birmingham and Julie Maupin-Furlow at the University of Florida, respectively]. All other gene knockouts were created de novo.

The aforementioned thermosome subunits were consistently among the top 15 proteins pulled down in our analyses, regardless of the method employed (see Chapter 3). Structural superimposition studies using TopMatch indicated alignment of β fold structure in each of the chaperone proteins to structures of known caspases, yet superimposition studies with the annotated caspase-6 structure (pdb ID 3P4U) using the software Cytoscape (22, 23), did not yield a high match in the active site folds. This points to the possible influence of the high, ionic environment for this haloarchaeon on protein folding. Thermosomes are protein chaperonins that are double-ring complexes of multiple subunits that have a central cavity to fold proteins one at a time, away from cytosolic aggregation-promoting complexes. There are two types of chaperonins: Group I and Group II. The Group II chaperonins are the archaeal thermosomes (cct
stands for chaperonin-containing TCP-1), and their eukaryotic homologs, the tailless complex polypeptide (TCP-1) ring complex (24, 25). *H. volcanii* was reported to have three *cct* genes that assist in protein folding (12, 25, 26). Knockout studies revealed that two of three double mutant strains in the *cct* genes were viable, but a Δ*cct1Δcct2* knockout was unviable; it was later demonstrated that the *cct1* subunit is essential and indispensable for growth (27). Temperature and salt stress studies on *H. volcanii* chaperone proteins suggested that each of the Cct proteins were dispensable for growth, with a reduced stress resistance in double mutants; only the presence of only *cct3* (Δ *cct1*, Δ *cct2*) was lethal for the organism. This suggested that archaeal chaperone system was more similar to bacteria than to eukaryotes, as every single chaperone component is essential in eukaryotes but not in archaea and bacteria. Cct complex analysis by electron microscopy and mass spectrometry indicate that individual chaperonin proteins in archaea do not have a specialized function like their eukaryotic counterparts however genetic analysis along with ATP hydrolysis indicate to a more potent function of *cct1* (6, 12).

Proteasome mediated protein degradation has been extensively studied in *H. volcanii* (5-7, 9, 28). Proteasomes are found in all domains of life and function to degrade inaccurately synthesized, misfolded, or damaged proteins. This quality control system of the cell also regulates processes like cell division and DNA repair by destroying regulatory proteins at specific times and places, or by cleaving precursor proteins to generate their active counterparts (5). Proteins undergo modification by a small protein called ubiquitin, with polyubiquitylation targeting a protein for degradation in eukaryotes. A similar mechanism exists in Archaea as *H. volcanii* has been shown to modify proteins by sampylation, which is hypothesized to target them for degradation (8). The 20S core particle is the central component
of the proteasome machinery that harbors the proteolytic sites. *H. volcanii* encodes five components associated with proteasome degradation system: a 600-kDa barrel shaped structure composed of α₁ (37.5 kDa), α₂ (34.5 kDa), and β (30 kDa) proteins encoded by *psmA, psmC*, and *psmB*, respectively. These form two 20S proteasome subtypes of differing subunit composition α₁β and α₁α₂β. Two proteasome activating nucleotidase (PAN) proteins, PanA and PanB assist the core particle in the unfolded protein degradation (6, 7, 28). Gene knockouts of proteasome protein components were created and *psmB* was shown to be essential for proteasome function (6). All the other single gene knockouts in proteosomal components were viable and phenotypically resembled wild type; with only the ΔpanA mutant strain demonstrating reduced growth.

While Cdc48 wasn’t among the best-aligned structures in the Cytoscape analysis, two of its five peptides identified in mass spectrometry were significantly enriched in the b-VAD fraction (Table 3.3, Chapter 3). Further, Cdc48 is included among the proteotoxic stress associated proteins that were abundantly represented in these pull down assays. Cdc48 was first identified as an ATPase required for progression of cell division in yeast; these studies provided the first link associating *cdc48* with ubiquitin in ER-associated degradation, ERAD (15, 16, 29). It was found to associate with ubiquitinated substrates directly to target them for degradation, working as a “gearbox” to control the fate of proteins. Recent work in the archaeon, *Thermoplasma acidophilum*, ascribed Cdc48 as a partner protein for 20S proteasome mediated protein degradation (16), and it was reported that the N-terminal domain of the protein possibly regulates the rate of intracellular protein degradation.
The putative nuclease has been annotated as an archaeal-specific RecJ-like exonuclease containing a DnaJ-type Zn finger domain. Such proteins are recognized in DNA replication, recombination and repair. Its predicted structure was shown to have a strong alignment to caspase-6 bound to its inhibitor, VEID-CHO (3P4U) over the length of the query covered and the number of identical residues in Cytoscape analysis (Table 3.4, Chapter 3), even though the number of spectral counts by LC MS/MS analysis is not high. To date, no studies have been conducted to identify the function of this gene.

The putative aminopeptidase ranked highest in the criteria chosen for selection of best structural alignments (i.e., the length of alignment versus the number of identical residues among the compared proteins; described in Chapter 3). The 1,380 bp gene encodes for a 435-residue protein that has a PA_2 protease-associated domain (cd04819), a subgroup 2 from residues 67-195 and a Zn binding peptidase domain from residues 163-399. The PA domain is an insert domain with unknown function seen in many signal peptidases, such as E3 ubiquitin ligases like the human protein GRAIL, and EDEM3 an ER-degradation-enhancing mannosidase-like-3 protein. In addition to having the strongest structural alignment with 3P4U among H. volcanii proteins, it contained bioinformatically identified domains associated with protein-degradation machinery genes, justifying its inclusions in these studies.

De novo mutant strains were successfully generated for cdc48 (MSP1; Table2) and the putative nuclease (MSP3; Table 2). Unfortunately, the creation of an aminopeptidase knockout was unsuccessful despite screening over 300 clones despite several attempts; it suggests that the mutation might be lethal. Given its strong structural alignment to caspase-6 and the
reasons outlined above, additional attempts and alternative methods are being used to
generate this knockout.

*Phenotyping mutant strains for caspase activity*

The caspase activity profiles for each mutant strain were examined compared to the
wild-type parental strain, H26. Cell extracts were tested for caspase activity with caspase-4
(LEVD-AMC), caspase-6 (VEID-AMC), and caspase-8 (IETD-AMC) substrates, but given its higher
specificity and the fact that the observed patterns were similar, data is reported only for
caspase-4 activity. There were variable impacts on caspase activity for each of the mutant
strains examined (Figure 4.1). None of the single *cct* knockouts demonstrated any loss of
caspase activity. However, there was an ~60% reduction in caspase activity in the Δ*cct1, Δcct3*
double mutant and an ~30% loss in the Δ*cct 2, Δcct3* double mutant, indicating the *H. volcanii*
caspase activity is linked to thermosome function. In the proteasome mutants examined,
there was a reduction in activity in the Δ*panB* and Δ*psmC* strains, 40% and 30%, respectively.
Interestingly, there were ~20% and ~99% increases in activity for Δ*panA* and in Δ*psmA* mutants
(Figure 4.1), suggesting that the *psmA* (α1) subunit of the *H. volcanii* proteasome acts
antagonistically to Cdc48 and the thermosomes or there is some feedback regulation of these
proteins in the protein turnover function. The most significant reduction in caspase activity was
seen in strain MSP1 (Δ*cdc48*), which lost ~70% of activity (Figure 4.1). Taken together, the loss
of activity in Δ*cdc48* in conjunction with the activity seen in the *cct* double mutants indicates
that some of the proteins identified in the pull down assays link the caspase activity in *H.
volcanii* to an overall protein turnover stress in the cell. Given that Cdc48 has been reported to
be a partner protein for 20S proteasome-mediated protein degradation in *Thermoplasma acidophilum* (16), there is likely a feedback regulation of this activity for *cdc48* and *psmA* mutants. Finally, strain MSP3 (Δnuc) demonstrated a modest reduction of 10-15% in caspase activity (Figure 4.1).

Given that strain MSP1 exhibited the greatest loss of caspase activity, we performed a genetic rescue to validate that the Δ*cdc48* gene knockout was the reason for this phenotype. MSP1 was complemented, *in trans*, by introducing a full-length, wild type copy of the *cdc48* gene on a shuttle vector, pTA354 (data not show). Caspase activity was completely recovered in the complemented strain in comparison with H26 indicating that the loss of activity was indeed a function of the Δ*cdc48* (Figure 4.1).

**Physiological connection between caspase activity and the unfolded protein response**

The observations that several mutant strains had notably different levels of caspase activity compared to wild type suggested a direct link between this activity and the protein turnover response machinery in *H. volcanii*. Consequently, we examined the relationship between caspase activity and physiological state of the parental wild-type strain (H26) in different growth media including rich, minimal, and minimal supplemented with canavanine. L-Canavanine is an arginine amino acid analog that induces protein unfolding by incorporating into growing polypeptide chains during translation and induces UPR (30, 31). These experiments provided an initial test of whether UPR induction triggers a response in caspase activity.
As expected, growth of H26 in minimal medium was lower in cell density, as compared with growth in rich medium. When H26 was exposed to canavanine, it barely reached an O.D.\textsubscript{600} of 1 after 60 hours, at which time untreated H26 had already entered death phase of growth (Figure 4.2). Clearly, treatment with 11 μM canavanine induces significant cellular stress in \textit{H. volcanii}, corroborating previous results seen by Zhou et al. (6). Corresponding caspase activity increased ~2.6 times in H26 cells grown in minimal media alone, compared to that observed in rich media, indicating a clear association of this activity with stress. However, upon induction of protein stress with canavanine, caspase activity was reduced by ~45% compared to H26 grown in minimal media (Figure 4.3). These results, combined with the aforementioned negative impact on growth, imply is that 11 μm canavanine may be too high of a concentration and consequently too severe a cellular stress, \textit{for H. volcanii} to maintain viability, ultimately pushing the cell over the edge. This observation corroborates other work in our lab whereby caspase activity plummets in \textit{H. volcanii} cells exposed to very high levels of different stresses (e.g., temperature, oxidative, pH), compared with sub-lethal doses (B. Enalls, M. Seth-Pasricha, K.D. Bidle and K.A. Bidle, in preparation). Use of this target canavanine concentration was initially justified based on previously published work (6), but clearly it will be prudent to test the caspase activity levels for \textit{H. volcanii} mutants incubating in lower, sub-lethal concentrations of canavanine. We hypothesize that a critical threshold of canavanine will initiate an UPR and induce higher levels of caspase activity in \textit{H. volcanii}. Of course, an ability to diagnose the extent of subcellular UPR, perhaps by assessing the sampylation state of proteins (9, 13, 32), will also be essential to mechanistically link it to caspase activity [e.g., discern
whether different levels of UPR induce specific changes (up or down) in activity]. Methods for this are not currently possible in our lab.

Given proteasomes are known participants in UPR and have documented chymotrypsin activity (LLVYase), we also assessed its response during growth in minimal media with and without canavanine to provide additional context. In both cases, cells displayed successive reductions in chymotrypsin activity (~19 and 52% respectively) when grown in minimal media and when exposed to canavanine, compared to rich media (Figure 4.3). This suggests that growth in minimal media and an additional supplement with canavanine, both represent stress conditions for the cell, impacted, the proteasome protein degradation machinery of the cell.

Importantly, the response of LLVYase activity in *H. volcanii* needs to be tested more rigorously, in order to elucidate the role(s) of proteasome function under different stress conditions.

When the various mutant strains generated for these studies were examined under rich medium growth conditions, each resembled the growth of the parental strain (data not shown). Similarly, all mutant strains, with the exception of Δcct1 Δcct3, grew similarly to H26 in minimal growth medium (Figures 4.4, 4.5A, 4.6A). However, when these strains were supplemented with 11 μM canavanine, clear differences could be detected between the parent and mutant strains, with each mutant strain, growing half as well and struggling to reach an O.D.600 of 0.5 after 60 h of growth (Figures 4.4, 4.5B, 4.6B). Growth stress has previously been reported in proteasome mutants, ΔpsmA and ΔpanA when supplemented with 11 μM canavanine (6). A link between proper proteosomal component function and UPR has already been established in *H. volcanii* with strains carrying gene knockouts in the 20S core α1 subunit (ΔpsmA) and a proteasome-activating nucleotidase (ΔpanA) subunit displaying limited growth in the presence
of L-canavanine (6). Zhou et al. report that the ΔpsmA proteasome knockout strain is the most sensitive of all proteasome mutants tested to these same low salt- or high temperature-stresses. Intriguingly, these same knockout strains are the only ones that displayed an appreciable increase (by 26 and ~100%, respectively) in caspase activity (Figure 4.1), reinforcing the point that non-optimal caspase levels (high or low) critically compromise fitness. Interestingly, the thermosome double knockout strains (Δcct1,3 and Δcct2,3), which displayed reduced caspase activity (Figure 4.1), also demonstrated the lowest survival rates when exposed to low salt- or high temperature-stress (12). Taken together, these findings suggest that non-optimal caspase levels (high or low compared to wild-type) critically compromise fitness in H. volcanii, at least as measured by cellular growth in response to canavanine. As noted above, growth of these various mutant strains will need to be examined under conditions of lower canavanine exposure. We would anticipate that growth stress will be slightly alleviated under these conditions, particularly if they are correlated with increased caspase activity levels. Clearly, it will be necessary to examine caspase activity levels in all mutant strains from these studies, as these experiments have yet to be performed.

Our findings that caspase 4 activity (LEVDase) is the dominant proteolytic activity in H. volcanii (34) and is biochemically linked (established both in inhibitor pull down experiments and in targeted gene knockouts) with proteasome protein components and cdc48 strongly suggest that caspase activity is part of an ‘interactase’ and is involved in proper proteasome function. It has been previously suggested that molecular chaperone networks function in de novo folding and conformational maintenance of proteins. Some proteins rely on molecular chaperones throughout their cellular lifetime to maintain their functionally active
conformation. Molecular chaperones are tightly integrated with the protein degradation pathways to remove non-functional, misfolded, or aggregated proteins to maintain proteostasis via the ubiquitin proteasome system (UPS). Chaperone proteins like Hsp70 and Hsp90 cooperate with the proteasome or the E3 ubiquitin ligase to either maintain target proteins from aggregating, or to ubiquitinate them. While these pathways are well characterized in bacteria and eukaryotes, this protein homeostasis system is not well understood in archaea (11, 12). In combination with previous reports on functional importance of individual thermosomes subunits (12), this study implicates a potent role of cct1 in both UPR and the relationship of caspase activity associated with the UPR.

Mutations in the three critical “protein homeostasis” participants raise the possibility that caspase activity evolved as a protein-quality control system in a cell ensuring that individual proteins were properly produced, folded, and degraded. Thus, the highly elevated caspase-4-like activity and its association with the H. volcanii UPR. Indeed, the yeast metacaspase, Mca1 was recently suggested to be involved in preventing protein aggregates in Saccharomyces cerevisiae, and facilitates the removal of unfolded proteins in a protein quality control pathway of the cell (35). The loss or increase in caspase-like activity in the proteasome, thermosome, and the cdc48 mutant strains also seems to impact growth in a protein stress environment, possibly due to enhanced proteotoxicity in the cell. We hypothesize that Cdc48 and proteasomes function redundantly in the cell for protein turnover, and, together with thermosomes, maintain protein homeostasis for proper cell growth and maintenance of cellular fitness.
Section 4.5 **FUTURE DIRECTIONS**

Further work is needed to rigorously examine the mechanistic link between caspase-like activity and the UPR by treating wild-type *H. volcanii* grown in the presence or absence of the caspase inhibitor z-VAD-fmk, and our suite of gene knockouts, with L-canavanine and test their physiological response. Do knockouts and/or z-VAD-fmk-treated cells exhibit an impaired UPR response, manifesting in an enhanced sensitivity to L-canavanine? We hypothesize that knockouts with fundamentally altered caspase activity, as well as wild-type inhibited with z-VAD-fmk, will not only be more sensitive to L-canavanine than wild-type, but that this distinction will be accentuated under stress conditions (when UPR is preferentially needed). We are particularly interested in the response of ΔpsmA and Δcdc48 because they represent opposite ends of the caspase activity spectrum.

We have by no means created a comprehensive set of knockout strains in potential target genes identified by LC-MS/MS data of *in situ* inhibitor trapped proteins. Instead, we focused on a subset of key proteins with affiliation to PN function as an initial attempt to document a link between caspase activity and protein folding stress. Future studies will include the creation of additional strains with knockouts in the following target genes such α-EF2 (HVO_0356), EF-1 (HVO_0359), DnaK chaperone protein (HVO_1590), and ornithine cyclodeaminase (HVO_2879). Each of these genes play additional key roles in UPR and/or protein stress pathways and reside in the top ~2% (for elongation factors) and top 20% (for DnaK chaperone and ornithine cyclodeaminase) of proteins identified in the pull down assays (Chapter 3). They will help solidify the connection between caspase activity and the cellular PN.
**DnaK, a chaperone protein (HVO_1590):** The DnaK chaperone protein in *H. volcanii* has an hsp70 and actin-like super family domain as assessed by BlastP. This superfamily includes the actin family, the HSP70 family of molecular chaperones proteins, actin family proteins, the hexokinase family proteins, cell shape determining protein MreB, and cell cycle proteins like FtsA (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Most interestingly, it has 54% identity and 70% similarity to the yeast BiP protein, an ER chaperone protein that functions to bind IRE1, PERK, and ATF6 that are repressed during normal cell conditions. In ER stress, BiP is activated and releases each of these proteins that trigger a distinct and multiple transcriptional response and activation of target genes associated with, for example, ERAD and apoptosis (36-38).

**The CBS domain pair protein (HVO_2384):** The CBS domain pair protein was among the proteins identified in the pull down assays. In recent work, it has been identified as an interactor with the DnaK chaperone protein to promote recovery of stalled DNA replication forks and avoid harmful recombination in *H. volcanii* (personal communication, thesis dissertation work of Kayleigh Wardell, University of Nottingham). Since both proteins were seen in the pull downs, the CBS domain pair mutant strain was tested for caspase activity and growth physiology (kindly donated to us by Thorsten Allers, University of Nottingham). While Wardell and Allers did not report growth defects in this mutant (unpublished data), our results indicated a reduced growth, correlated with 2-4 fold lower cell numbers, as compared with parental strain H26. Further, the mutant showed significantly reduced activity although we speculate this is likely a result of impaired growth.

**Translation elongation factors α-EF2 and EF1 (HVO_0356 and HVO_0359):** EF-1 has been previously reported to increase in abundance in *H. volcanii* cells when proteasomal
function is inhibited by the inhibitor clasto-lactocystin-\(\beta\)-lactone (cL\(\beta\)L). EF-1 expression also increases in \(H.\ volcanii\) during heat shock, cold shock, and nutritional stress (39), with this latter condition causing SAMPylation (10). A recent study in \(Schizosaccharomyces\ pombie\) revealed the eIF3 translation initiation factor associates with protein synthesis and degradation machinery like proteasomes and elongation factors EF-1 and EF-2, to form a large supercomplex known as the translasome (40). Taken together, these data suggest that the elongation factors may be useful targets in knockout studies to assess caspase activity and growth physiology.

**Ornithine cyclodeaminase (HVO_2879):** In addition to being among the top proteins identified by skyline quantification of MS data (Chapter 3), ornithine cyclodeaminase protein was shown to be highly abundant (~9.7x) in optimal salt conditions as compared with high salt stress in \(H.\ volcanii\) (41), indicating a housekeeping role in the cell under optimal conditions. Ornithine cyclodeaminase participates in protein biogenesis of proline and arginine. Arginine has important roles in removal of ammonia and is also a precursor of NO, a signaling molecule implicated in PCD pathways (14, 42, 43).
Section 4.6 **ACKNOWLEDGEMENTS**

We thank Michael Maniscalco, Piotr Nawrot, and Charlotte Fuller for technical help. This research was supported in part by a grant from The Gordon and Betty Moore Foundation.

The dissertation author was the primary researcher and author, and the co-authors listed in this publication directed and supervised the research which forms the basis for this chapter.

Author Contributions: M.S.P. and K.A.B. designed research, M.S.P. performed research, M.S.P, K.A.B., and K.D.B. analyzed data, M.S.P., K.A.B., and K.D.B. wrote the chapter.
### Table 4.1. Primers used in the creation of mutant strains.

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Sequence</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>PutNuc UpF</td>
<td>5'-GAATTCCGGGAATAGCCC-3'</td>
<td>36 bp downstream of the +1 ATG start site</td>
</tr>
<tr>
<td>PutNuc UpR</td>
<td>5'-TTCTTCACCTGAAACCGAAGACTCG-3'</td>
<td>233 bp downstream form +1ATG site</td>
</tr>
<tr>
<td>PutNuc DnF</td>
<td>5'-GTATCAGGACTGCGAATCTCAT-3'</td>
<td>466 bp upstream of the 3' end of gene</td>
</tr>
<tr>
<td>PutNuc DnR</td>
<td>5'-GGAAGTGAAGCCGAGCTGAATG-3'</td>
<td>266 bp upstream of 3' end of gene</td>
</tr>
<tr>
<td>PutPep UpF</td>
<td>5'-GGCCATGCACCCATCTGATGCC-3'</td>
<td>407 bp upstream of the +1 site</td>
</tr>
<tr>
<td>PutPep UpR</td>
<td>5'-TCCAGGCGACTCCCGTGAACGT-3'</td>
<td>46 bp after the +1 site</td>
</tr>
<tr>
<td>PutPep DnF</td>
<td>5'-CGGCAGACCCATCGAGAACCTGGA-3'</td>
<td>194 bp upstream of the stop site</td>
</tr>
<tr>
<td>PutPep DnR</td>
<td>5'-TACCGAACCGTGACCACTCCCG-3'</td>
<td>156 bps downstream of the stop site</td>
</tr>
<tr>
<td>cdc48UpF</td>
<td>5'-ATGAAACTCAGCTGAACCGCC-3'</td>
<td>ATG +1 start site</td>
</tr>
<tr>
<td>cdc48UpR</td>
<td>5'-GGATGTAGCTGGGAATTT-3'</td>
<td>338 bps downstream of +1</td>
</tr>
<tr>
<td>cdc48DnF</td>
<td>5'-TCCCCCGCGAGATCTTCC-3'</td>
<td>231 bps upstream of the stop codon</td>
</tr>
<tr>
<td>cdc48DnR</td>
<td>5'-TTACTGGAAGGTGCGACCG-3'</td>
<td>The last 22 bps of the gene</td>
</tr>
<tr>
<td>cdc48FullF</td>
<td>5'-TCCGAAAGCAGAATAG-3'</td>
<td>starts 80 bp upstream of the +1 site</td>
</tr>
<tr>
<td>cdc48FullR</td>
<td>5'-ACTGAGTTCGAACTCC-3'</td>
<td>starts 81 bp after the end of the gene</td>
</tr>
</tbody>
</table>
Table 4.2. List of strains and plasmids created and used in this work.

<table>
<thead>
<tr>
<th>Strains</th>
<th>Description</th>
<th>Source or Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>DS70</td>
<td>Wild-type isolate DS2 cured of plasmid pHV2</td>
<td>Wendoloski et al., 2001</td>
</tr>
<tr>
<td>H26</td>
<td>DS70 ΔpyrE2</td>
<td>Allers et al., 2004</td>
</tr>
<tr>
<td>MSP1</td>
<td>H26 (Δcdc48)</td>
<td>This study</td>
</tr>
<tr>
<td>MSP3</td>
<td>H26 (Δnuc)</td>
<td>This study</td>
</tr>
<tr>
<td>Top10</td>
<td>E.coli (F- mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 recA1 araD139 Δ(ara- leu)7697 galU galK rpsL (StrR) endA1 nupG)</td>
<td>Invitrogen</td>
</tr>
<tr>
<td>Inv110</td>
<td>F’ (traΔ36 proAB lacIq lacΔ M15) rpsL (StrR) thr leu endA thi-1 lacY galK galT ara tonA tsx dam dcm supE44 Δ(lac-proAB) Δ(mcrC-mrr)102::Tn10 (TetR)</td>
<td>Invitrogen</td>
</tr>
<tr>
<td>cct1</td>
<td>H26 (Δcct1)</td>
<td>Kapatai et al., 2006</td>
</tr>
<tr>
<td>cct2</td>
<td>H26 (Δcct2)</td>
<td>Kapatai et al., 2006</td>
</tr>
<tr>
<td>cct3</td>
<td>H26 (Δcct3)</td>
<td>Kapatai et al., 2006</td>
</tr>
<tr>
<td>cct1,cct3</td>
<td>H26 (Δcct1,Δcct3)</td>
<td>Kapatai et al., 2006</td>
</tr>
<tr>
<td>cct2,cct3</td>
<td>H26 (Δcct2,Δcct3)</td>
<td>Kapatai et al., 2006</td>
</tr>
<tr>
<td>GZ108</td>
<td>H26 panB (ΔpanB)</td>
<td>Zhou et al., 2008</td>
</tr>
<tr>
<td>GZ109</td>
<td>H26 panA (ΔpanA)</td>
<td>Zhou et al., 2008</td>
</tr>
<tr>
<td>GZ114</td>
<td>H26 psmC (Δ a2)</td>
<td>Zhou et al., 2008</td>
</tr>
<tr>
<td>GZ130</td>
<td>H26 psmA (Δ α1)</td>
<td>Zhou et al., 2008</td>
</tr>
</tbody>
</table>

Plasmids

<table>
<thead>
<tr>
<th>Plasmids</th>
<th>Description</th>
<th>Source or Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>pTA131</td>
<td>Amp'; pBluescript II containing P_{ps-psy2}</td>
<td>Allers et al., 2004</td>
</tr>
<tr>
<td>pTA354</td>
<td>E. coli/H. volcanii pTA131-based shuttle vector, with pyrE2 marker &amp; pHV1/4 replication origin</td>
<td>Norais et al., 2007</td>
</tr>
<tr>
<td>pMSP1</td>
<td>pTA131 with cdc48 fragment containing a 1691 bp deletion within its coding region</td>
<td>This study</td>
</tr>
<tr>
<td>pMSP2</td>
<td>pTA354 with ~2.2 kb full-length cdc48 gene and 80bp each from 5' and 3' flanking regions</td>
<td>This study</td>
</tr>
<tr>
<td>pMSP3</td>
<td>pTA131 with nuc fragment containing a 1491 bp deletion within its coding region</td>
<td>This study</td>
</tr>
<tr>
<td>pMSP4</td>
<td>pTA131 with pep fragment containing a 1068 bp deletion within its coding region</td>
<td>This study</td>
</tr>
<tr>
<td>pMSP5</td>
<td>pTA354 with ~1.3 kb full-length pep gene and 5' and 3' flanking regions</td>
<td>This study</td>
</tr>
</tbody>
</table>
**Figure 4.1. Examination of caspase activity in all mutant strains.** Caspase activity was monitored in mutant strains as compared with activity in the parental strain, H26. Activity was recorded as a percentage of wild-type activity. The *cdc48* mutant strain, complemented, *in trans*, with pMSP2, demonstrates full recovery of activity. All experiments were done in technical triplicates. Data is indicative of biological duplicates.
**Figure 4.2. Growth of H26 in Hv minimal and canavanine stress.** Growth of H26 was monitored at 45°C in rich, minimal, and minimal medium supplemented with 11 µm canavanine. Canavanine supplemented medium exhibited a drastic reduction in growth for strain H26.
Figure 4.3. Impact of growth conditions on caspase and chymotrypsin activity in *H. volcanii*. Caspase (top panel) and chymotrypsin (bottom panel) activity (expressed as a hydrolysis rate of canavanine LEVD-AMC and substrate LLVY-AMC, respectively) was examined for cells growing under different media conditions and in the presence of canavanine. The change in activity between rich media and minimal media, as well as between minimal media in the presence and absence of canavanine are indicated by arrows and %. Both activities are differently impacted by stress conditions. Caspase activity increases when grown under the low level stress of minimal media, with activity dropping significantly under the more intense stress of canavanine treatment. Chymotrypsin activity drops considerably under both stress conditions. Hydrolysis rates for LEVD- and LLVY-AMC were measured for cell extracts at 42°C as previously reported (60).
**Figure 4.4. Impact of growth conditions in strain MSP1.** Growth of MSP1 compared to parental strain H26 was monitored at 45°C in minimal medium and minimal medium supplemented with 11 μM canavanine over a period of 60 h.
Figure 4.5. Impact of growth conditions on the thermosome mutant strains. Growth of the thermosome mutants compared to parental strain H26 at 45°C in (A) minimal media and (B) minimal medium supplemented with 11 μm canavanine over a period of 60 h.
Figure 4.6. Impact of growth conditions on the proteasome mutants. Growth of the proteasome mutants as compared with parental strain H26 at 45°C in (A) minimal medium and (B) minimal medium supplemented with 11 µm canavanine over a period of 60 hours.
Section 4.8 REFERENCES


34. **Seth-Pasricha M, Bidle KA, Bidle KD.** 2013. Specificity of archaeal caspase activity in the extreme halophile Haloferax volcanii. Environ Microbiol Reports **5:**263-271.


