Hydrophobic carboxy-terminal residues dramatically reduce protein levels in the haloarchaeon *Haloferax volcanii*

Christopher J. Reuter, Sivakumar Uthandi, Jose A. Puentes and Julie A. Maupin-Furlow

University of Florida, Department of Microbiology and Cell Science, Gainesville, FL 32611-0700, USA

Correspondence

Julie A. Maupin-Furlow
jmaupin@ufl.edu

Received 22 July 2009
Revised 29 September 2009
Accepted 18 October 2009

Proteolysis is important not only to cell physiology but also to the successful development of biocatalysts. While a wide-variety of signals are known to trigger protein degradation in bacteria and eukaryotes, these mechanisms are poorly understood in archaea, known for their ability to withstand harsh conditions. Here we present a systematic study in which single C-terminal amino acid residues were added to a reporter protein and shown to influence its levels in an archaeal cell. All 20 amino acid residues were examined for their impact on protein levels, using the reporter protein soluble modified red-shifted GFP (smRS-GFP) expressed in the haloarchaeon *Haloferax volcanii* as a model system. Addition of hydrophobic residues, including Leu, Cys, Met, Phe, Ala, Tyr, Ile and Val, gave the most pronounced reduction in smRS-GFP levels compared with the addition of either neutral or charged hydrophilic residues. In contrast to the altered protein levels, the C-terminal alterations had no influence on smRS-GFP-specific transcript levels, thus revealing that the effect is post-transcriptional.

INTRODUCTION

Post-translational regulation of protein levels is critical to the viability of all living cells and is controlled to a large extent by proteolysis. Protein turnover removes damaged and extraneous proteins, resulting in free amino acids, which can be recycled (Gottesman, 2003; Wolf & Hilt, 2004). Energy-dependent proteolysis, carried out by such proteases as ClpP, FtsH, Lon and proteasomes, is responsible for the degradation of the majority of these proteins (Gottesman, 2003). Substrate differentiation is determined in part by protease-associated ATPases of the AAA⁺ superfamily (Ogura & Wilkinson, 2001), whose additional responsibilities include substrate unfolding, protease gate opening and substrate translocation (Lupas & Martin, 2002; Smith *et al.*, 2005).

Selective removal of proteins by proteolysis is based on at least one of a variety of factors, including exposed

hydrophobic groups, post-translational protein modifications (e.g. phosphorylation and glycosylation), specific proteolytic recognition sequences (e.g. SsrA and ubiquitin) and the identity of amino acid residues at the extreme termini of proteins (e.g. N-end rule) (Bohley, 1996; Varshavsky *et al.*, 2000; Withey & Friedman, 2003; Devoy *et al.*, 2005). The small stable RNA A (SsrA) protein-tagging system is found in all bacterial genomes (Keiler *et al.*, 2000) and employs a transfer-messenger RNA (tmRNA) to add an 11-residue sequence (-AANDE-NYALAA) to the end of stalled translation products, thus conveying a proteolytic targeting signal for Clp proteases (Gottesman *et al.*, 1998; Withey & Friedman, 2003). The N-end rule, described in bacteria and eukaryotes, predicts the half-life of a protein based on its extreme N-terminal residue (N-degron) (Varshavsky *et al.*, 2000). In eukaryotes, short-lived N-end rule proteins are recognized by ubiquitin-protein ligases, which in conjunction with other enzymes of the ubiquitination pathway, target these substrates for degradation by the addition of poly-ubiquitin chains (Ciechanover, 1998). Poly-ubiquitination precedes and stimulates substrate targeting of many proteins for degradation by 26S proteasomes.

Although many proteolytic targeting systems have been described, relatively little is known regarding these events in archaea. Archaea do synthesize a variety of proteases, including 20S proteasome core particles (CPs) and

Abbreviations: CP, core particle; α -GFP, anti-GFP antibody; E-GFP, enhanced GFP; qRT-PCR, quantitative real-time PCR; smRS-GFP, soluble modified red-shifted GFP.

A supplementary table, showing strains and plasmids used in this study, and two supplementary figures, showing growth and fluorescence of *H. volcanii* strains expressing smRS-GFP and C-terminal variants, and quantitative immunoblotting of *H. volcanii* strains expressing smRS-GFP wild-type and C-terminal variants, are available with the online version of this paper.

regulatory particle triple-A ATPases (Rpt proteins), that together function as energy-dependent proteases (Zwickl *et al.*, 1999; Wilson *et al.*, 2000). However, the signals that stimulate protein degradation by these types of proteases are poorly understood.

SsrA-tagging systems are not predicted in archaea. However, SsrA can serve as a hydrophobic signal to stimulate proteasome-mediated protein degradation when fused to the C terminus of reporter proteins such as green fluorescent protein (GFP–SsrA) *in vitro* (Navon & Goldberg, 2001) and soluble modified red-shifted GFP (smRS–GFP–SsrA) *in vivo* (Reuter & Maupin-Furlow, 2004). Thus, SsrA fusions have served as convenient reporters to study energy-dependent proteolysis in archaea. Interestingly, addition of a single Ala residue to the C terminus of smRS–GFP also reduces the levels of smRS–GFP by approximately 90% in the halophilic archaeon *Haloflex volcanii* (Reuter & Maupin-Furlow, 2004). Based on these results, it is unclear whether the addition of other single C-terminal amino acids would similarly influence reporter protein levels in this archaeal species.

Here we demonstrate that the levels of smRS–GFP are substantially reduced when hydrophobic amino acid residues are singly added to its C terminus in *H. volcanii* cells. The addition of the hydrophobic residues Leu, Cys, Met, Phe, Ala, Tyr, Ile and Val gave the most pronounced reduction in smRS–GFP levels compared with the addition of either neutral or charged hydrophilic residues. The addition of these C-terminal residues had little if any influence on smRS–GFP-specific transcript levels, thus demonstrating that the effect is post-transcriptional.

METHODS

Materials. Organic and inorganic analytical grade chemicals were from Fisher Scientific and Sigma Chemical, unless otherwise indicated. Desalted oligonucleotides were from Integrated DNA Technologies. Restriction endonucleases and DNA-modifying enzymes were from New England BioLabs.

Strains, media and plasmids. Strains, oligonucleotide primers used for cloning, and plasmids are listed in Supplementary Table S1. *Escherichia coli* DH5 α was used for routine recombinant DNA experiments. *E. coli* strains were grown at 37 °C (200 r.p.m.) in Luria–Bertani medium. *H. volcanii* strains were grown at 37 °C (200 r.p.m.) in ATCC974 medium, unless otherwise indicated. Media were supplemented with 100 mg ampicillin l⁻¹ or 0.1 mg novobiocin l⁻¹ as needed. The fidelity of all PCR-amplified products was confirmed by DNA sequencing using the dideoxy termination method (Sanger *et al.*, 1977) with an Amersham Pharmacia Biotech MegaBACE 1000 DNA sequencer at the DNA Sequencing Facilities, Interdisciplinary Center for Biotechnology Research, University of Florida.

Construction of *H. volcanii* expression plasmids. Expression plasmids were prepared with oligonucleotides according to Supplementary Table S1. PCR mutagenesis was used to add codons to the 3' end of the gene encoding smRS–GFP (*smrs-gfp*) directly preceding its stop codon. The added codons corresponded to each of the 20 standard amino acids and were optimized for expression in *H.*

volcanii according to the Codon Usage Database (<http://www.kazusa.or.jp/codon/>). Restriction enzyme sites were included in all constructs for directional cloning into pJAM202, an *E. coli*–*H. volcanii* shuttle plasmid. Plasmids contained a *Halobacterium cutirubrum* rRNA P2 promoter and T7 transcriptional terminator to regulate transcription of all cloned genes. Plasmid pJAM202c, identical to the expression plasmids described above except for the absence of the *smrs-gfp* gene, was included as a control.

DNA isolation and transformation. Plasmids were isolated using a Qiagen Miniprep kit (Qiagen). DNA fragments were eluted from 0.8% (w/v) SeaKem GTG agarose (FMC Bioproducts) gels with 1 × TAE buffer (40 mM Tris-acetate, 2 mM EDTA, pH 8.5) using a QIAquick Gel Extraction kit (Qiagen). *H. volcanii* DS70 cells were transformed with plasmid DNA isolated from *E. coli* GM2163 as described by Cline *et al.* (1995).

Growth and fluorescence of cell cultures. The OD₆₀₀ of liquid cultures was measured using a Bio-Rad SmartSpec 3000 spectrophotometer (Bio-Rad). Fluorescence (excitation 488 nm, emission 509 nm) of liquid cultures was monitored using an Aminco-Bowman series 2 luminescence spectrometer (Spectronic Instruments). The innate fluorescence of control cultures, *H. volcanii* DS70 (pJAM202c), was subtracted from the fluorescence of *H. volcanii* DS70 cultures expressing smRS–GFP and its C-terminal variants. The growth and fluorescence of selected smRS–GFP variants were also monitored in *H. volcanii* GZ134 (Δ panA Δ psmA) under similar conditions. One unit was defined as the fluorescence intensity of fluorescein at 0.025 nM, as previously described (Reuter & Maupin-Furlow, 2004).

Quantification of fluorescence and protein levels of smRS–GFP and its variants in *H. volcanii*. *H. volcanii* DS70 control (pJAM202c), and DS70 and GZ134 strains individually expressing smRS–GFP or its C-terminal variants from appropriate plasmids (Supplementary Table S1), were streaked from freezer stocks [stored at –70 °C in 20% (v/v) glycerol] onto solid medium and grown for 3 days. Isolated colonies were inoculated into liquid medium (1 ml in 13 × 100 mm tubes) and rotated at a 70° slant at 56 r.p.m. for about 24 h until cells reached OD₆₀₀ 0.4–0.8. These cultures were used (50–100 μ l) to inoculate fresh medium (1 ml) and grown as above. This final step was repeated two more times, with the exception that the third set of cultures were grown to OD₆₀₀ 0.2–0.3. These cultures were used (40–50 μ l) to inoculate fresh medium (5 ml) in triplicate and were grown in tubes (as described above) for 7 days to OD₆₀₀ ~3.0. Cell fluorescence was measured as described above. Cells (4 ml) were harvested (10 000 g for 10 min at 4 °C) and resuspended in 0.5 ml deionized water. Cells were further lysed by sonication, and the insoluble fraction was removed by centrifugation (10 000 g for 10 min at 4 °C). The protein concentration of each clarified cell lysate was determined using the Bradford assay according to the supplier's recommendations (Bio-Rad). Protein samples (10 μ g) were separated by reducing SDS-PAGE (12%) in Criterion cassettes (Bio-Rad) with 26-well combs for 70 min at 200 mV and analysed by quantitative immunoblotting. Each immunoblot included enhanced GFP (E–GFP) (Zhang *et al.*, 1996) purified from recombinant *E. coli* as a standard (37, 55.5, 74, 92.5, 111, 148 and 185 ng) and was performed with Living Colours *Aequorea victoria* peptide rabbit immunoglobulin anti-GFP antibody (α -GFP) as the primary antibody (Clontech Laboratories) and horseradish peroxidase-conjugated anti-rabbit immunoglobulin (H + L) antibody raised in goats (Southern Biotechnology Associates) as the secondary antibody, as previously described (Reuter & Maupin-Furlow, 2004). Cultures (25 ml) were also grown (37 °C, 200 r.p.m.) in 250 ml Erlenmeyer flasks inoculated in triplicate with 50–100 μ l of three times subcultured exponential-phase cells. Fluorescence and growth of these cell cultures

were monitored at various time points over the course of at least 110 h (late stationary phase).

Quantitative real-time PCR (qRT-PCR). Total RNA was isolated from 1.5 ml of cells (OD₆₀₀ 0.7) using the RNeasy Mini kit for bacteria (Qiagen) as previously reported (Reuter *et al.*, 2004), with modifications as previously described (Sherwood *et al.*, 2009). Total RNA concentration was measured by A₂₆₀ and quality was determined by 0.8% (w/v) agarose gel electrophoresis in 1 × TAE buffer (40 mM Tris-acetate, 2 mM EDTA, pH 8.0) after heating the RNA samples to 50 °C for 15 min. qRT-PCR was performed on total RNA preparations as previously described (Sherwood *et al.*, 2009) using the iQ SYBR Green Supermix and iCycler MyiQ real-time PCR detection system (Bio-Rad) with primers specific for *smrs-gfp* and *ribL10* (ribosomal protein). The *ribL10* primers (RibL-RT_f and RibL-RT_r) were as reported elsewhere and served as a normalization control (Brenneis *et al.*, 2007). The *smrs-gfp* primers were Fw, 5'-ACGACTTCTTCAAGAGCG-3', and Rev, 5'-CTGCCGTGATGT-ATACGT-3'.

Genome sequence. All analysis of the theoretical proteome of *H. volcanii* was based on the April 2007 genome sequence generated by J. Eisen and colleagues at the Institute for Genomic Research (available at <http://archaea.ucsc.edu/>).

RESULTS

Generation of *H. volcanii* strains expressing smRS-GFP C-terminal variants

Previously, it was shown that addition of an SsrA motif (-AANDENYALAA) to the C terminus of smRS-GFP rendered the chimeric protein undetectable by immunoblotting when synthesized in *H. volcanii* cells (Reuter & Maupin-Furlow, 2004). Alterations and truncations of this motif had variable effects on the levels of smRS-GFP present in the recombinant *H. volcanii* cells. Surprisingly, a 10-residue truncation of the SsrA motif, which left only a single Ala residue fused to the C terminus, diminished the cellular levels of smRS-GFP to as low as only 10% of the unmodified form. To further investigate these findings and systematically determine which C-terminal residue addition most influences smRS-GFP levels in *H. volcanii* cells, 20 *smrs-gfp* reporter genes were constructed such that each included an extra codon at the 3' end of the gene that was optimized for expression in *H. volcanii* and corresponded to one of the 20 standard amino acids. These *smrs-gfp* gene variants, in addition to the unmodified *smrs-gfp* gene, were similarly ligated into the *E. coli*-*H. volcanii* shuttle plasmid pJAM202 (Supplementary Table S1) (Kaczowka & Maupin-Furlow, 2003). Fidelity was confirmed by DNA sequencing of the *smrs-gfp* genes using the dideoxy termination method (Sanger *et al.*, 1977). With the exception of the codons added at the 3' end of the reporter gene, the plasmids were identical and included a haloarchaeal rRNA P2 promoter and T7 transcriptional terminator for robust expression in *H. volcanii*. The plasmid vector used for these constructs (pJAM202) has already been used for the high-level production of a wide variety of proteins in this archaeon, and thus has proven

quite versatile (Kaczowka & Maupin-Furlow, 2003; Reuter & Maupin-Furlow, 2004; Zhou *et al.*, 2008; Humbard *et al.*, 2009).

C-terminal variants reduce smRS-GFP-mediated whole-cell fluorescence and reporter protein levels in *H. volcanii*

The whole-cell fluorescence of the *H. volcanii* strains carrying the various site-directed mutations of the *smrs-gfp* gene was expected to closely parallel the intracellular levels of the encoded fluorescent protein reporter. This is based on the compact structure of GFP as well as that of its red-shifted variant, both of which are composed of a single-domain β -can with a centrally located chromophore (Ormö *et al.*, 1996; Yang *et al.*, 1996). Although deletion of more than the N-terminal methionine or seven amino acids from the C terminus of GFP abolishes its fluorescence and presumably the β -can structure (Yang *et al.*, 1996), the protein is highly resistant to denaturation (Yang *et al.*, 1996). Furthermore, the rather flexible N and C termini on the surface of GFP enable this protein and its variants to be fused to other proteins or amino acid residues without structural distortion or alteration of fluorescence (Chalfie, 1995; Gottesman *et al.*, 1998).

To determine whether the levels of the smRS-GFP variant proteins could be monitored by whole-cell fluorescence, *H. volcanii* cells were grown to stationary phase (OD₆₀₀ ~3.0). The fluorescence of each recombinant strain was determined and corrected for *H. volcanii* autofluorescence using the strain carrying the vector alone (pJAM202c) as a negative control (Fig. 1 and Supplementary Fig. S1). In addition, total protein was prepared from each strain and the concentration of smRS-GFP protein was determined by quantitative immunoblotting using α -GFP. Total protein of the strain expressing unmodified smRS-GFP (wild-type) and serial dilutions of purified E-GFP protein were included as standards on each immunoblot for reproducibility and to verify that samples were within the detection limit of the assay. In addition, the *H. volcanii* strain with plasmid vector alone (pJAM202c) was included on the immunoblot as a negative control. Using this approach, whole-cell fluorescence was found to closely parallel the protein levels of smRS-GFP and its C-terminal variants in whole cells (Fig. 1). The results also revealed that the addition of C-terminal residues to smRS-GFP reduced the levels of this protein and its corresponding fluorescence in whole cells. In particular, the addition of hydrophobic residues gave a more pronounced reduction in the levels of smRS-GFP than the addition of hydrophilic residues to its C terminus.

Reduction in smRS-GFP levels by addition of C-terminal residues is post-transcriptional

To further investigate whether the reduction in smRS-GFP levels and whole-cell fluorescence observed after the

Fig. 1. Whole-cell fluorescence correlates with protein levels of smRS-GFP wild-type and C-terminal variants expressed in *H. volcanii*. (a) Representative α -GFP immunoblot of *H. volcanii* strains expressing unmodified (wild-type, wt) and variant smRS-GFP proteins with Pro, Leu and Tyr residues added to the C terminus (smRS-GFP-P, -L and -Y, respectively). Total protein (10 µg per lane) from biological triplicate cultures was separated by 12 % SDS-PAGE and analysed by α -GFP immunoblotting as described in Methods. Purified E-GFP standards and total protein of *H. volcanii* expressing unmodified smRS-GFP (wt) were included on each blot as indicated. (b) Comparison of relative protein levels (white bars) and whole-cell fluorescence (black bars) of smRS-GFP and its C-terminal variants expressed in *H. volcanii*. Relative protein levels were determined by quantitative immunoblotting for unmodified (wt) and C-terminal variant smRS-GFP proteins, as indicated. A 100 % relative protein level represents 2 % of total soluble protein. A 100 % relative fluorescence represents 7 U (as defined in Methods). All experiments were performed in biological triplicate, and the means $\pm$ SD were calculated.

addition of C-terminal amino acid residues was related to altered transcript levels, qRT-PCR analysis was performed. Three smRS-GFP variants were chosen for this analysis, including those demonstrating a dramatic, moderate and slight reduction in reporter protein levels (C-terminal-added Val, Ser and Asp residues, respectively). Primer pairs were designed to detect the levels of *smrs-gfp*-specific transcript in total RNA preparations. In addition, the *ribL10*-specific transcript levels (encoding a ribosomal protein) were determined and served as an unregulated transcript control (Brenneis *et al.*, 2007). The qRT-PCR analysis revealed no dramatic difference in the levels of *smrs-gfp*-specific transcript among the various strains (i.e. cells producing wild-type smRS-GFP compared with strains producing the reporter protein with Val, Ser or Asp residues added to the C terminus) (Table 1). These results suggest that the reduced levels of smRS-GFP

detected in the variant strains were due to a post-transcriptional event such as reduced protein half-life or translation efficiency. Although influences on translation cannot be ruled out, the modest changes in nucleotide sequence (addition of only a single 'optimized' codon to the 3' end of the gene) favour the likelihood that the half-life of the fluorescent protein reporter was reduced in the C-terminal variants.

Levels of C-terminal smRS-GFP variants are rapidly reduced during exponential-phase growth

To more closely examine the finding that the addition of C-terminal residues to smRS-GFP reduces the levels of this protein in *H. volcanii* cells, the fluorescence and growth of the various strains were monitored over time (lag to stationary phase) (Fig. 2 and Supplementary Fig. S2). From

Table 1. C-terminal variants do not influence *smRS-gfp*-specific transcript levels in *H. volcanii* cells

qRT-PCR analysis of transcripts encoding smRS-GFP and its variants was performed on total RNA isolated from *H. volcanii* DS70 cells expressing genes encoding unmodified (wt) and C-terminal variant (D, S and V) smRS-GFP proteins *in trans*, as indicated. *ribL10* primers were included as a normalization control (see Methods for details). N-fold change is expressed as the levels of *smrs-gfp*-specific transcripts in cells encoding the C-terminal variant compared with those of wild-type smRS-GFP. All experiments were performed in biological triplicate, and the means $\pm$ SD were calculated.

C-terminal variant	N-fold change
smRS-GFP-wt	1.0
smRS-GFP-D	0.87 ± 0.3
smRS-GFP-S	0.93 ± 0.3
smRS-GFP-V	1.08 ± 0.2

this analysis, it was found that expression of neither the fluorescent reporter protein nor its variants had an impact on growth. Growth of the *H. volcanii* strains expressing either smRS-GFP or its variants was similar to that of the negative control, *H. volcanii* with vector alone (plasmid pJAM202c). The whole-cell fluorescence of each strain, however, did vary and depended on the nature of the amino acid added to the C terminus of the reporter protein (Figs 2 and 3, and Supplementary Fig. S2). Cells expressing the unaltered *smrs-gfp* gene rapidly (0.54 U h^{-1}) reached a near-maximum fluorescence of 6.4 U during exponential phase, which remained at a plateau for about 1.5 days. After this period, fluorescence steadily increased from 6.4 to 8.1 U at a much lower rate (0.03 U h^{-1}) throughout stationary phase (Fig. 2). In contrast to the wild-type smRS-GFP, addition of any of the 20 standard amino acids to the C terminus of smRS-GFP reduced whole-cell fluorescence (Fig. 3). The extent of this reduction was highly dependent on the type of amino acid, with hydrophobic residues, including Leu, Cys, Met, Phe, Ala, Tyr, Ile and Val, giving the most pronounced reduction in cellular fluorescence (Fig. 3).

Interestingly, all of the *H. volcanii* DS70 cultures expressing the smRS-GFP variants displayed a notable increase in whole-cell fluorescence (up to 1–4 U) during the early stages of growth that rapidly declined or reached a low-level plateau, often during the transition to stationary phase (Fig. 2 and Supplementary Fig. S2). In addition to this exponential-phase ‘spike’ and ‘decline’ in fluorescence, many of the variant strains exhibited a second increase in smRS-GFP fluorescence during the later stationary phase. For example, the strain synthesizing smRS-GFP-K displayed an early and rapid increase in fluorescence at a rate of 0.2 U h^{-1} that was closely linked to exponential-phase growth (from 25 to 41 h). Following this increase, the strain displayed a subsequent and relatively rapid (0.044 U h^{-1}) drop in fluorescence during the transition to

Fig. 2. C-terminal variants of smRS-GFP influence the fluorescence of *H. volcanii*. Growth of *H. volcanii* DS70 strains with smRS-GFP reporter protein expression plasmids (Δ) and control plasmid (pJAM202c) ($\blacktriangle$) as measured by OD₆₀₀. The fluorescence of *H. volcanii* DS70 (pJAM202c) cultures was subtracted from the fluorescence of *H. volcanii* DS70 cultures with the smRS-GFP reporter protein expression plasmids to generate a net fluorescence curve ($\circ$) for each strain, as indicated. All experiments were performed in biological triplicate, and the means $\pm$ SD were calculated.

stationary phase. Once the cells reached stationary phase (at 63 h), the fluorescence steadily increased to 6.2 U (0.039 U h^{-1}) for the duration of the time-course. Based on these results, it appears that a post-transcriptional event predominant in the exponential phase is responsible for the decreases observed in the fluorescence and protein levels of the C-terminal smRS-GFP variants.

Whether the proteasome system is involved in the reduction of the C-terminal smRS-GFP protein levels in *H. volcanii* cells remains to be determined. Knockout of the *panA* and *psmA* genes (encoding the Rpt-like and α -type

Fig. 3. Addition of C-terminal hydrophobic residues to smRS-GFP reduces the fluorescence to an extent that correlates to levels of this reporter protein in *H. volcanii* cells. Maximum fluorescence (y axis) was determined by monitoring the fluorescence of whole-cell cultures over time, as shown in Fig. 2 and Supplementary Fig. S2. C-terminal amino acid residue additions to smRS-GFP (wt) are on the x axis, with those having a positive hydrophathy index based on the Eisenberg consensus scale (open circles) (Eisenberg *et al.*, 1984) or charged side chain (open square, positive; open triangle, negative) indicated. All experiments were performed in biological triplicate, and the means $\pm$ SD were calculated.

proteasomal components PAN-A and $\alpha 1$, respectively) had little if any influence on the fluorescence of *H. volcanii* cells expressing smRS-GFP C-terminal variants, including those with added Asp, Lys, Ser and Val residues (data not shown). Whether full depletion of the proteasomal 20S CPs would have a similar impact is unclear. Our recent work demonstrates that knockout of the sole β -type CP subunit or double knockout of the two α -type CP subunits ($\alpha 1$ and $\alpha 2$) is lethal to *H. volcanii* cells, thus complicating our ability to analyse smRS-GFP levels in whole cells (Zhou *et al.*, 2008).

DISCUSSION

In this study, the fluorescent reporter protein smRS-GFP was modified by the addition of each of the 20 standard amino acids singly to its C terminus and synthesized in recombinant *H. volcanii*. The fluorescence of these *H. volcanii* strains was monitored and found to correlate with smRS-GFP protein levels. Compared to the unaltered smRS-GFP, addition of any single amino acid to its C terminus had a negative impact on reporter protein levels and corresponding fluorescence in *H. volcanii* cells. The levels of smRS-GFP-specific transcript, in contrast, were not influenced by these site-directed modifications. Thus, although alterations in translation cannot be ruled out, the modest changes in nucleotide sequence (addition of only a single ‘optimized’ codon to the 3’ end of the gene) strongly suggest that the half-life of the fluorescent protein reporter

was reduced by the addition of these C-terminal residues. The reason for this reduction is still unclear. The unstructured nine-residue tail of wild-type GFP, which includes a C-terminal Lys residue (-THGMDELYK), has been implicated as a potential recognition factor of VAT, an archaeal AAA-type ATPase homologue of VCP/Cdc48/p97 that is a putative regulatory particle of proteasomes (Gerega *et al.*, 2005). In our case, extension of this unstructured region of the smRS-GFP by addition of single amino acid residues may have increased its vulnerability as a substrate of this system.

The type of amino acid added to the C terminus of smRS-GFP had a great impact on the levels of this reporter protein and corresponding fluorescence in *H. volcanii* cells. Neutral non-polar residues, Leu, Cys, Met, Phe, Ala, Tyr, Ile and Val, gave the most pronounced reduction in cellular fluorescence, whereas the addition of hydrophilic amino acids was less perturbing (Fig. 3). Of the hydrophilic residues, the acidic amino acids Asp and Glu as well as Lys and Asn were the least disruptive. Proline was also found to have a minimal impact on smRS-GFP levels. These results suggest that addition of hydrophilic acidic amino acids to the surface C termini of proteins of compact structure does not substantially influence their susceptibility to proteolysis in *H. volcanii* cells. Consistent with this, acidic amino acids constitute about 20 % of the residues in the theoretical *H. volcanii* proteome and are often found on the outer surface of these salt-loving proteins (Mevarech *et al.*, 2000). As a result, most *H. volcanii* proteins require salt for activity and

possess a highly charged (acidic) surface that serves as a hydration shell, enabling catalysis to occur under high-salt conditions and preventing protein aggregation (Mevarech *et al.*, 2000). Although basic amino acids are found in a relatively low abundance in halophilic proteins, Arg is abundant on the C termini of *H. volcanii* proteins and constitutes about 13 % of the proteome deduced from the genome sequence. It should also be noted that the last amino acid of the wild-type smRS-GFP is Lys, a basic residue, which is highly stable in *H. volcanii* cells.

Proteolysis based on recognition of substrates with exposed hydrophobic amino acid residues has been demonstrated in other organisms. These types of residues are typically found in the interior of native protein structures and their exposure occurs most often after protein damage and unfolding (Bohley, 1996; Grune *et al.*, 2003). Typically these proteins are removed by proteolysis to prevent their aggregation and to maintain protein quality control within the cell. In addition to improperly folded proteins, hydrophobic residues are also found on the C termini of some proteins that require controlled and timed turnover by proteases. For example, the FliF motor protein of *Caulobacter crescentus* requires a hydrophobic patch at its C terminus for cell cycle- and ClpA-dependent degradation (Jenal & Shapiro, 1996). Likewise, hydrophobic residues of SsrA, a tag added to the C terminus of incomplete proteins in bacteria, are important for recognition and degradation by ClpXP. These include the most distal Ala-10 and Ala-11, which are critical determinants for recognition by ClpX (Flynn *et al.*, 2001).

Conclusions

To further understand the production and stability of proteins in archaea, it is important to determine how the biochemical properties of these proteins can alter their levels within the cell. Archaeal organisms are increasingly being relied upon for use as model systems for eukaryotes and are considered by many to be attractive hosts in biotechnological applications due to their ability to survive extreme conditions along with other unique traits and enzymes (Margesin & Schinner, 2001; Schiraldi & De Rosa, 2002; Kaczowka *et al.*, 2005; Egorova & Antranikian, 2005; Soppa, 2006). Advances in these areas of research will be achieved through the comprehension of post-transcriptional regulation and optimization of protein expression in archaea.

ACKNOWLEDGEMENTS

We thank Savita Shanker at the Genomics Core of the University of Florida Interdisciplinary Center for Biotechnology Research for DNA sequencing. This work was funded in part by NIH R01 GM057498 and DOE DE-FG02-05ER15650.

REFERENCES

- Bohley, P. (1996). Surface hydrophobicity and intracellular degradation of proteins. *Biol Chem* **377**, 425–435.
- Brenneis, M., Hering, O., Lange, C. & Soppa, J. (2007). Experimental characterization of *cis*-acting elements important for translation and transcription in halophilic archaea. *PLoS Genet* **3**, e229.
- Chalfie, M. (1995). Green fluorescent protein. *Photochem Photobiol* **62**, 651–656.
- Ciechanover, A. (1998). The ubiquitin-proteasome pathway: on protein death and cell life. *EMBO J* **17**, 7151–7160.
- Cline, S. W., Pfeifer, F. & Doolittle, W. F. (1995). Transformation of halophilic archaea. In *Archaea: a Laboratory Manual*, pp. 197–204. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Devoy, A., Soane, T., Welchman, R. & Mayer, R. J. (2005). The ubiquitin-proteasome system and cancer. *Essays Biochem* **41**, 187–203.
- Egorova, K. & Antranikian, G. (2005). Industrial relevance of thermophilic Archaea. *Curr Opin Microbiol* **8**, 649–655.
- Eisenberg, D., Weiss, R. M. & Terwilliger, T. C. (1984). The hydrophobic moment detects periodicity in protein hydrophobicity. *Proc Natl Acad Sci U S A* **81**, 140–144.
- Flynn, J. M., Levchenko, I., Seidel, M., Wickner, S. H., Sauer, R. T. & Baker, T. A. (2001). Overlapping recognition determinants within the ssrA degradation tag allow modulation of proteolysis. *Proc Natl Acad Sci U S A* **98**, 10584–10589.
- Gerega, A., Rockel, B., Peters, J., Tamura, T., Baumeister, W. & Zwickl, P. (2005). VAT, the *Thermoplasma* homolog of mammalian p97/VCP, is an N domain regulated protein unfoldase. *J Biol Chem* **280**, 42856–42862.
- Gottesman, S. (2003). Proteolysis in bacterial regulatory circuits. *Annu Rev Cell Dev Biol* **19**, 565–587.
- Gottesman, S., Roche, E., Zhou, Y. & Sauer, R. T. (1998). The ClpXP and ClpAP proteases degrade proteins with carboxy-terminal peptide tails added by the SsrA-tagging system. *Genes Dev* **12**, 1338–1347.
- Grune, T., Merker, K., Sandig, G. & Davies, K. J. (2003). Selective degradation of oxidatively modified protein substrates by the proteasome. *Biochem Biophys Res Commun* **305**, 709–718.
- Humbard, M. A., Zhou, G. & Maupin-Furlow, J. A. (2009). The N-terminal penultimate residue of 20S proteasome α 1 influences its N α -acetylation and protein levels as well as growth rate and stress responses of *Haloferax volcanii*. *J Bacteriol* **191**, 3794–3803.
- Jenal, U. & Shapiro, L. (1996). Cell cycle-controlled proteolysis of a flagellar motor protein that is asymmetrically distributed in the *Caulobacter* predivisional cell. *EMBO J* **15**, 2393–2406.
- Kaczowka, S. J. & Maupin-Furlow, J. A. (2003). Subunit topology of two 20S proteasomes from *Haloferax volcanii*. *J Bacteriol* **185**, 165–174.
- Kaczowka, S. J., Reuter, C. J., Talarico, L. A. & Maupin-Furlow, J. A. (2005). Recombinant production of *Zymomonas mobilis* pyruvate decarboxylase in a recombinant haloarchaeon, *Haloferax volcanii*. *Archaea* **1**, 327–334.
- Keiler, K. C., Shapiro, L. & Williams, K. P. (2000). tmRNAs that encode proteolysis-inducing tags are found in all known bacterial genomes: a two-piece tmRNA functions in *Caulobacter*. *Proc Natl Acad Sci U S A* **97**, 7778–7783.
- Lupas, A. N. & Martin, J. (2002). AAA proteins. *Curr Opin Struct Biol* **12**, 746–753.
- Margesin, R. & Schinner, F. (2001). Potential of halotolerant and halophilic microorganisms for biotechnology. *Extremophiles* **5**, 73–83.
- Mevarech, M., Frolov, F. & Gloss, L. M. (2000). Halophilic enzymes: proteins with a grain of salt. *Biophys Chem* **86**, 155–164.
- Navon, A. & Goldberg, A. L. (2001). Proteins are unfolded on the surface of the ATPase ring before transport into the proteasome. *Mol Cell* **8**, 1339–1349.

- Ogura, T. & Wilkinson, A. J. (2001). AAA⁺ superfamily ATPases: common structure–diverse function. *Genes Cells* **6**, 575–597.
- Ormö, M., Cubitt, A. B., Kallio, K., Gross, L. A., Tsien, R. Y. & Remington, S. J. (1996). Crystal structure of the *Aequorea victoria* green fluorescent protein. *Science* **273**, 1392–1395.
- Reuter, C. J. & Maupin-Furlow, J. A. (2004). Analysis of proteasome-dependent proteolysis in *Haloferax volcanii* cells using short-lived green fluorescent proteins. *Appl Environ Microbiol* **70**, 7530–7538.
- Reuter, C. J., Kaczowka, S. J. & Maupin-Furlow, J. A. (2004). Differential regulation of the PanA and PanB proteasome-activating nucleotidase and 20S proteasomal proteins of the haloarchaeon *Haloferax volcanii*. *J Bacteriol* **186**, 7763–7772.
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* **74**, 5463–5467.
- Schiraldi, C. & De Rosa, M. (2002). The production of biocatalysts and biomolecules from extremophiles. *Trends Biotechnol* **20**, 515–521.
- Sherwood, K. E., Cano, D. J. & Maupin-Furlow, J. A. (2009). Glycerol-mediated repression of glucose metabolism and glycerol kinase as the sole route of glycerol catabolism in the haloarchaeon *Haloferax volcanii*. *J Bacteriol* **191**, 4307–4315.
- Smith, D. M., Kafri, G., Cheng, Y., Ng, D., Walz, T. & Goldberg, A. L. (2005). ATP binding to PAN or the 26S ATPases causes association with the 20S proteasome, gate opening, and translocation of unfolded proteins. *Mol Cell* **20**, 687–698.
- Soppa, J. (2006). From genomes to function: haloarchaea as model organisms. *Microbiology* **152**, 585–590.
- Varshavsky, A., Turner, G., Du, F. & Xie, Y. (2000). The ubiquitin system and the N-end rule pathway. *Biol Chem* **381**, 779–789.
- Wilson, H. L., Ou, M. S., Aldrich, H. C. & Maupin-Furlow, J. A. (2000). Biochemical and physical properties of the *Methanococcus jannaschii* 20S proteasome and PAN, a homolog of the ATPase (Rpt) subunits of the eucaryal 26S proteasome. *J Bacteriol* **182**, 1680–1692.
- Withey, J. H. & Friedman, D. I. (2003). A salvage pathway for protein synthesis: tmRNA and trans-translation. *Annu Rev Microbiol* **57**, 101–123.
- Wolf, D. H. & Hilt, W. (2004). The proteasome: a proteolytic nanomachine of cell regulation and waste disposal. *Biochim Biophys Acta* **1695**, 19–31.
- Yang, F., Moss, L. G. & Phillips, G. N., Jr (1996). The molecular structure of green fluorescent protein. *Nat Biotechnol* **14**, 1246–1251.
- Zhang, G., Gurtu, V. & Kain, S. R. (1996). An enhanced green fluorescent protein allows sensitive detection of gene transfer in mammalian cells. *Biochem Biophys Res Commun* **227**, 707–711.
- Zhou, G. Y., Kowalczyk, D., Humbard, M. A., Rohatgi, S. & Maupin-Furlow, J. A. (2008). Proteasomal components required for cell growth and stress responses in the haloarchaeon *Haloferax volcanii*. *J Bacteriol* **190**, 8096–8105.
- Zwickl, P., Ng, D., Woo, K. M., Klenk, H.-P. & Goldberg, A. L. (1999). An archaeobacterial ATPase, homologous to ATPases in the eukaryotic 26S proteasome, activates protein breakdown by 20S proteasomes. *J Biol Chem* **274**, 26008–26014.

Edited by: H.-P. Klenk