Stationary Phase and Nutrient Levels Trigger Transcription of a Genomic Locus Containing a Novel Peptide (TM1316) in the Hyperthermophilic Bacterium *Thermotoga maritima*

Andrew D. Frock,* Clemente I. Montero,* Sara E. Blumer-Schuetz, Robert M. Kelly

Department of Chemical and Biomolecular Engineering, North Carolina State University, Raleigh, North Carolina, USA

The genome of the hyperthermophilic bacterium *Thermotoga maritima* encodes numerous putative peptides/proteins of 100 amino acids or less. While most of these open reading frames (ORFs) are transcribed during growth, their corresponding physiological roles are largely unknown. The onset of stationary phase in *T. maritima* was accompanied by significant morphological changes and upregulation of several ORFs located in the TM1298-TM1336 genome locus. This region contains putative HicAB toxin-antitoxin pairs, hypothetical proteins, radical S-adenosylmethionine (SAM) enzymes, and ABC transporters. Of particular note was the TM1315-TM1319 operon, which includes a putative 31-amino-acid peptide (TM1316) that was the most highly transcribed gene in the transcriptome during stationary phase. Antibodies directed against a synthetic version of TM1316 were used to track its production, which correlated closely with transcriptomic data. Immunofluorescence microscopy revealed that TM1316 was localized to the cell envelope and prominent in cell aggregates formed during stationary phase. The only functionally characterized locus with an organization similar to that of TM1315-TM1319 is in *Bacillus subtilis*, which contains subtilisin A, a cyclic peptide with Cys-to- α -carbon linkages that functions as an antilisterial bacteriocin. While the organization of TM1316 resembled that of the *Bacillus* peptide (e.g., in its number of amino acids and spacing of Cys residues), preparations containing high levels of TM1316 affected the growth of neither *Thermotoga* species nor *Pyrococcus furiosus*, a hyperthermophilic archaeon isolated from the same locale as *T. maritima*. Several other putative Cys-rich peptides could be identified in the TM1298-TM1336 locus, and while their roles are also unclear, they merit examination as potential antimicrobial agents in hyperthermophilic biotopes.

The increasing availability of genome sequence information has driven the need for, and development of, improved tools for gene annotation that predict the functions of encoded biomolecules. As a result, the number of open reading frames (ORFs) annotated as “hypothetical proteins” continues to shrink. However, smaller ORFs (i.e., corresponding to putative proteins and peptides of fewer than 100 amino acids [aa]) remain annotated to a disproportionate extent as hypothetical proteins (Table 1). These await closer bioinformatic inspection and complementary biochemical evaluation. The problem, however, is that small proteins and peptides are “needles in a haystack” of small ORFs that are present in genomes (1). Any random stretch of 50 codons has a 9% chance of lacking a stop codon, while a stretch of 100 codons has a 0.8% chance of lacking a stop codon. Therefore, even a small bacterial genome may contain hundreds of small ORFs with physiological and ecological significance.

For many reasons, small ORFs are often subject to inaccurate gene calls. The bioinformatic difficulties in accurately predicting small protein- and peptide-coding genes can be addressed by experimental techniques, such as tiling arrays and deep sequencing of RNA, but these methods are not yet routinely employed as a part of genome annotation. However, even after these small ORFs are accurately identified, there are other challenges in annotating and characterizing the corresponding small proteins and peptides. Many could have physiological roles that depend on posttranslational processing, localization to the membrane, or association with other proteins, such that overproduction in a heterologous host (a workhorse method for characterization of larger proteins) may not be useful. The small size also complicates one- or two-dimensional PAGE and the generation of nonpolar disruption mutants.

Finally, some small ORFs may be located in “silent” genomic regions that are not transcribed, at least under standard laboratory conditions.

Although small proteins and peptides may not be biocatalytic, they can function in a variety of roles as toxins, bacteriocins, extracellular signaling peptides, subunits of larger protein complexes, regulators of other proteins/enzymes, and chaperones of metals and nucleic acids (2, 3). There are many such examples. Increased production of the MntS protein (42 aa) from *Escherichia coli* was observed under manganese-limited conditions, but its overproduction renders cells sensitive to manganese, suggesting that it plays a role in increasing the intracellular bioavailability of the metal (4). SafA (65 aa) from *Escherichia coli* is induced by the EvgS/EvgA two-component system (TCS), and it activates the PhoQ/PhoP TCS by directly interacting with PhoQ (5). KdpF (29 aa) is a stabilizing component of the K⁺-translocating KdpFABC complex in *E. coli* (6). In *Bacillus subtilis*, the ComX peptide (55 aa) regulates competence and surfactin synthesis when the se-

Received 17 May 2013 Accepted 16 August 2013

Published ahead of print 23 August 2013

Address correspondence to Robert M. Kelly, rmkelly@ncsu.edu.

* Present address: Andrew D. Frock, Novozymes, Franklinton, North Carolina, USA; Clemente I. Montero, Albany Medical Center, Albany, New York, USA.

Supplemental material for this article may be found at <http://dx.doi.org/10.1128/AEM.01627-13>.

Copyright © 2013, American Society for Microbiology. All Rights Reserved.

doi:10.1128/AEM.01627-13

TABLE 1 Distribution of ORFs annotated as hypothetical proteins in selected microbial genomes as a function of the numbers of amino acids that they encode^a

Microbe	Total no. of predicted ORFs encoding the indicated no. of amino acids and no. of those annotated as Hyp													
	1–100		101–200		201–300		301–400		401–500		501–600		>600	
	Total	Hyp	Total	Hyp	Total	Hyp	Total	Hyp	Total	Hyp	Total	Hyp	Total	Hyp
<i>T. maritima</i> MSB8	167	113	382	209	453	178	383	108	237	67	87	23	149	35
<i>P. furiosus</i> DSM 3638	286	210	545	330	502	228	395	140	225	67	64	18	108	24
<i>C. saccharolyticus</i> DSM 8903	233	143	631	289	588	142	512	125	341	59	159	17	215	34
<i>B. anthracis</i> A0248	860	692	1,239	570	1,159	341	817	163	510	58	217	19	239	26
<i>H. pylori</i> J99	123	77	331	165	349	151	272	79	189	50	77	14	147	24
<i>E. coli</i> K-12 MG1655	423	20	887	1	949	0	785	0	539	0	224	0	338	0
<i>P. aeruginosa</i> PAO1	331	237	1,259	788	1,357	534	1,073	346	756	182	294	73	501	129
Total	2,423	1,492	5,274	2,352	5,357	1,574	4,237	961	2,797	483	1,122	164	1,697	272
% Hyp	61.6		44.6		29.4		22.7		17.3		14.6		16.0	

^a Based on NCBI annotations. Hyp, hypothetical.

creted form activates the histidine kinase ComP, which phosphorylates the transcription factor ComA (7). In the hyperthermophilic bacterium *Thermotoga maritima*, the transfer-messenger RNA (tmRNA) gene is colocated with a gene on the opposite strand (TM0504) that encodes a 42-aa peptide involved in quorum sensing-like induction of exopolysaccharide production (8, 9). Finally, a number of class I and class II bacteriocins, including subtilisin A (43 aa), and lantibiotics, such as nisin (57 aa), are encoded in small ORFs (10, 11).

To examine the significance of small ORFs in microbial genomes, conditions under which their transcription and production are triggered must be identified. Here, the transcriptional response of *T. maritima* during growth phase transitions and subjected to various nutrient levels was examined with an eye toward ORFs corresponding to fewer than 100 aa. These experiments revealed that transcripts for most small ORFs were generated under these growth conditions. In particular, this led to the identification of a genomic region that contained several putative peptides, one of which (TM1316) corresponded to the most highly expressed ORF in the *T. maritima* transcriptome during stationary phase.

MATERIALS AND METHODS

Growth of *Thermotoga maritima*. *T. maritima* DSM-3109 was obtained from the Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ). *T. maritima* was grown anaerobically in an oil bath at 80°C with shaking at 100 rpm. Culture bottles were made anaerobic by sealing them, sparging them with nitrogen, and adding Na₂S·9H₂O at a final concentration of 0.05% (wt/vol). For preparation of freezer stocks, 970 μl of cooled log-phase culture was added to 30 μl of dimethyl sulfoxide and stored at –80°C. Freezer stocks were thawed and used to inoculate 1% (vol/vol) artificial seawater (ASW) media (12) with 2.5 g/liter maltose. After ~12 h of growth, this culture was used to inoculate medium with various concentrations of yeast extract (YE). This medium was based on a defined medium developed from ASW. The base medium consisted of (per liter, final concentration) 15 g NaCl, 2 g Na₂SO₄, 2 g MgCl₂·6H₂O, 0.25 g NaHCO₃, 20 mg KBr, 20 mg H₃BO₃, 20 mg KI, 3 mg Na₂WO₄, and 2 mg NiCl₂·6H₂O. The base medium was typically prepared at a 2× concentration, adjusted to pH 7 with H₂SO₄, and sterilized by autoclaving. After it was autoclaved, 1.2 g KH₂PO₄, 2.5 g maltose monohydrate, 5 g NH₄Cl, 5.4 mg FeCl₃·6H₂O, 1× trace elements (see below), and 1× vitamins (see below) were added (per liter, final concentration). All of these components were prepared as concentrated, sterile-filtered stock solutions. YE concentration was varied by addition of a 25-g/liter stock solu-

tion that had been sterilized by filtration. Vitamin solution (1,000×) consisted of (per liter) 10 mg pyridoxine-HCl, 5 mg calcium pantothenate, 5 mg nicotinic acid, 5 mg riboflavin, 5 mg thiamine-HCl, 5 mg thiocctic acid, 2 mg biotin, 2 mg folic acid, and 0.1 mg cyanocobalamin. Trace element solution (1,000×) consisted of (per liter) 1.5 g nitrilotriacetic acid (NTA), 0.5 g MnSO₄·H₂O, 0.2 g NiCl₂·6H₂O, 0.1 g ZnSO₄·7H₂O, 0.01 g CuSO₄·5H₂O, and 0.01 g Na₂MoO₄·2H₂O. NTA was dissolved by adding concentrated KOH to pH 6.5. For cell enumeration, cells were fixed with 0.25% glutaraldehyde, stained with acridine orange, and viewed under an epifluorescence microscope (13).

Transcriptional-response analysis via cDNA microarray. Cultures (300 ml) were harvested at different time points as indicated in Fig. 1 and rapidly cooled in a dry ice-ethanol bath, following which cells were collected by centrifugation at 7,500 × g for 10 min. RNA extraction, reverse transcription, labeling, hybridization, and slide scanning were performed with modifications of previously described protocols (8, 14). Loess normalization, analysis of variance (ANOVA) normalization, and ANOVA row-by-row modeling were performed using JMP Genomics software (SAS Institute, Cary, NC). Fold changes were at or above 2-fold, with significance based on values at or above the Bonferroni correction (typically equivalent to a *P* value of approximately 10^{–5}). These experiments were performed with a multispecies microarray (15), which included probes for *T. maritima* and other *Thermotoga* species (*T. neapolitana*, *T. trophila*, and *Thermotoga* sp. strain RQ2).

Immunocytochemistry of TM1316. Polyclonal rabbit antibodies (GeneTel Laboratories LLC, Madison, WI) were raised against a chemically synthesized version of the TM1316 peptide conjugated to carrier protein (keyhole limpet hemocyanin). For dot blots, 2 μl of sample was manually spotted onto a nitrocellulose membrane and allowed to dry. The membrane was blocked with 2.5% nonfat dry milk in Tris-buffered saline (TBS) for 30 min and then incubated for 30 min with anti-TM1316 rabbit serum diluted 5,000× in 2.5% nonfat dry milk in TBS plus 0.1% Tween 20 (TBS-T). The membrane was washed twice briefly with TBS-T and then washed four times for 5 min each before a 30-min incubation with goat anti-rabbit IgG-horseradish peroxidase (HRP) conjugate (Invitrogen) diluted 100,000× in 2.5% nonfat dry milk in TBS-T. The membrane was again washed as before; HRP was detected with SuperSignal West Femto chemiluminescent substrate (Pierce), and the membrane was imaged using a Carestream Gel Logic 2200 Pro system. For Western blots, NuPAGE 4 to 12% Bis-Tris gels (Invitrogen) were run according to the manufacturer's instructions, and proteins were transferred to a nitrocellulose membrane at 1 mA per cm² of membrane for 40 min. For chemiluminescent-ladder detection, a biotinylated protein ladder was loaded into at least one lane of the gel, and 10,000×-diluted goat anti-biotin IgG-HRP conjugate (Cell Signaling Technology, Danvers, MA) was added to the secondary-antibody incubation.

FIG 1 Growth of *T. maritima* and production of a small peptide encoded in TM1316 varies with the levels of yeast extract in culture medium. The cell densities of *T. maritima* when grown on 1 g/liter (gray) versus 5 g/liter (black) yeast extract are denoted by solid lines. TM1316's abundance, estimated by determining immunoblot intensity, is denoted by dotted lines for 1 g/liter (gray) and 5 g/liter (black) yeast extract. For each growth condition, cultures were sampled for transcriptomic analysis during exponential phase (triangles) and stationary phase (squares) at the times indicated by the black arrows. Note that TM1316 levels were significantly higher in stationary phase when *T. maritima* was grown on the 1 g/liter yeast extract medium than when it was grown on 5 g/liter yeast extract.

Nucleotide sequence accession numbers. Full microarray results are available through the GEO platform accession no. GPL13655 and GEO series accession no. GSE47072.

RESULTS

Growth and morphology of *T. maritima* during phase transitions. Growth rates for *T. maritima* on either 1 or 5 g/liter yeast extract (YE) in ASW-based medium with 2.5 g/liter maltose were not significantly different (doubling times of ~80 to 90 min), although final cell densities on 5 g/liter YE were somewhat higher (4×10^8 to 5×10^8 cells/ml versus 1.5×10^8 to 2×10^8 cells/ml) (Fig. 1). During exponential phase of both cultures, cells exhibited the rod-like shape that is generally observed for *T. maritima* (Fig. 2A and C) (16). However, during stationary phase of the 1-g/liter YE culture, a significant proportion of cells (roughly 20 to 50%) became bloated and coccoidal, as shown in Fig. 2B. This was not observed in cultures grown on 5 g/liter YE (Fig. 2D).

Exponential- and stationary-phase transcriptomes for growth on rich medium. When *T. maritima* was grown on medium containing 5 g/liter YE, the genes differentially regulated between exponential and stationary phase reflected changes in cellular functions upon entry into stationary phase. Of the 187 ORFs significantly upregulated in log phase compared to stationary phase (Tables 2 and 3, and see Table S1 in the supplemental material), 55 were involved in carbohydrate transport and metabolism, including glucose-inducible components of transporters for cellobiose (17), glucose (18), and trehalose (18). Two maltose

transporters (19) were transcribed at slightly higher levels during exponential phase, but these differences generally did not meet the criteria for statistical significance. Cell division genes (*minC*, *minD*, and *ftsI*) were more highly transcribed during log phase, along with genes involved in fatty acid biosynthesis (2), nucleic acid replication and repair (11), translation (35), and protein processing and secretion (10).

The stationary-phase transcriptome for the rich medium (5 g/liter YE) suggested that various vitamins and amino acids were depleted during exponential growth. Assimilation of ammonium as a nitrogen source apparently increased during stationary phase, as an ammonium transporter (TM0402) and the nitrogen-regulatory protein P-II (TM0403) were upregulated. Also upregulated were 12 flagellum-related genes and genes involved in the biosynthesis of tyrosine (TM0344), leucine (TM0548-TM0555), cysteine (TM0882), and lysine (TM1519-TM1520), among others. Genes involved in the biosynthesis of the vitamins riboflavin (TM1826-TM1827), thiamine (TM0787-TM0788, TM0888, TM1244-TM1251), and pantothenate (TM1077) were also more highly transcribed during stationary phase.

The transcriptome provided evidence that balancing redox reactions was a key issue for *T. maritima* during growth, although genes implicated in the oxidative-stress response (21) were not differentially regulated. The Rex transcriptional regulator (TM0169) responds to the cellular NADH/NAD⁺ ratio and controls expression of 39 genes in 12 operons (22, 23). Here, 10 of

FIG 2 Fluorescence microscopy of *T. maritima* grown on media containing 1 g/liter and 5 g/liter yeast extract (YE) during log phase and stationary phase. Cell morphology of *T. maritima* cultures grown on medium containing 1 g/liter yeast extract in exponential phase (A) and stationary phase (B) or of cultures grown on medium containing 5 g/liter yeast extract in exponential phase (C) and stationary phase (D). All images used acridine orange as the fluorescent dye, and the scale bars indicate 5 μ m. The background apparent in panels C and D was attributed to the higher concentrations of yeast extract.

these genes were upregulated in stationary phase compared to exponential phase, and half of those use NAD(P)H as a cofactor. These Rex-regulated genes include rubrerythrin (TM0009), an NADP-reducing hydrogenase (TM0011 and TM0012), and a serine-pyruvate aminotransferase and hydroxypyruvate reductase (TM1400 and TM1401) (24), as well as glycerol dehydrogenase (TM0423) and glyceraldehyde-3-phosphate dehydrogenase (TM0688).

Exponential- and stationary-phase transcriptomes for growth on lean medium. The exponential- to stationary-phase transition on 1 g/liter YE shared many features with that on 5 g/liter YE. Of the 187 probes that were expressed at least 2-fold more in exponential phase than in stationary phase on 5 g/liter, 80 met the same criterion for the same comparison on 1 g/liter YE. In general, the transition from exponential phase to stationary phase

on 1 g/liter YE was less clear from the transcriptome. For example, only 2 flagellum genes were upregulated for this transition on 1 g/liter YE (Table 3). The transcription of many growth-associated genes (including ribosomal proteins and the cellobiose, glucose, and trehalose transporters) decreased significantly upon entering stationary phase on 5 g/liter YE, while the transcription of these genes decreased to a much lesser extent (or not at all) on 1 g/liter YE.

Comparison of transcriptomes on 1 g/liter and 5 g/liter yeast extract. The numbers of genes that responded to the change in medium composition (right side of Table 2) were clearly much smaller than the numbers that responded to phase transition (left side of Table 2). During exponential phase, two glucose-inducible sugar-binding transporter components (TM0031 and *xylE2*) were upregulated on 1 g/liter YE compared to 5 g/liter YE. This suggests that *T. maritima* was more dependent upon utilization of maltose during growth on the leaner medium, presumably because the increased levels of YE in the richer medium provide additional alternative carbon sources. Clustered regularly interspaced short palindromic repeat (CRISPR)-associated adaptive immunity proteins (TM1792 to TM1802) (25) and a glycerol uptake operon (TM1429 to TM1431) were also upregulated in 1 g/liter versus 5 g/liter YE during exponential phase. Only four genes were upregulated in 5 g/liter versus 1 g/liter YE during exponential phase: those for two hypothetical proteins (TM1266 and TM1268) in the same operon as putative thiamine and biotin synthesis proteins (TM1267 and TM1269, respectively), a prismane protein (TM1172), and a ferrous iron transport protein (TM0051). The majority of iron in the media came from the supplemented ferric chloride (1.1 mg/liter iron) instead of Bacto yeast extract (~0.06 mg/liter iron for 1 g/liter YE and 0.3 mg/liter iron for 5 g/liter YE). A component of a ferric iron transporter (TM0190) was upregulated in 5 g/liter YE versus 1 g/liter YE during stationary phase. Other ORFs upregulated in 5 g/liter YE versus 1 g/liter YE during stationary phase include flagellum-related genes, the primary sigma factor RpoD (TM1451), and the heat shock-inducible sigma factor RpoE (TM1598) (26). The rod shape-determining MreB protein was also upregulated in 5 g/liter YE in stationary phase, which may be related to the coccoidal morphology observed in 1 g/liter YE at stationary phase (Fig. 2B).

As discussed above, the transcription of a number of growth-associated genes did not decrease during the 1-g/liter-YE-phase transition, as was the case for the 5-g/liter-YE-phase transition. Other genes upregulated during stationary phase on the leaner medium included a putative sulfur relay operon (TM0979 to TM0983). TM0981, TM0980, and TM0979 are homologous to DsrEFH and TusBCD, which are involved in sulfur relay for sul-

TABLE 2 Total numbers of *Thermotoga maritima* ORFs differentially transcribed on the basis of growth phase transition and medium composition

Medium or growth phase(s)	Total no. of ORFs upregulated			
	In exponential vs stationary phase	In stationary vs exponential phase	With 1 g/liter vs 5 g/liter YE	With 5 g/liter vs 1 g/liter YE
1 g/liter YE	173	81		
5 g/liter YE	187	172		
Both 1 g/liter and 5 g/liter YE	80	58		
Exponential phase			18	4
Stationary phase			70	45
Both exponential and stationary phase			3	0

TABLE 3 Selected ORFs in *T. maritima* responding to 1 and 5 g/liter YE during exponential and stationary phases

Locus tags or genes	Function/product	Avg fold change ^a			
		1 E vs 5 E	1 S vs 5 S	1 S vs 1 E	5 S vs 5 E
TM0009–TM0012	NADP-reducing hydrogenase			2.1	2.9
TM0024–TM0032	Beta-glucoside utilization		2.7		–3.1
<i>treEFG</i>	Trehalose transporter		3.4		–3.5
<i>xylE2F2K2</i>	Glucose/xylose ABC transporter		3.1		–4.3
TM0063–TM0069	Uronate utilization			–3.5	–4.5
TM0081–TM0086	Flagellum protein			2.0	3.4
TM0087–TM0092	Unknown			–2.7	–2.3
TM0101–TM0105	Sugar ABC transporter				–2.4
TM0132–TM0134	Flagellum protein				2.4
TM0272–TM0275	Central carbon metabolism				–3.0
TM0394–TM0398	Amino acid metabolism			2.1	2.8
TM0402–TM0403	Ammonium assimilation			6.8	9.3
TM0430–TM0432	Digalacturonate ABC transporter			3.7	5.2
TM0441–TM0443	Mannonate-to-gluconate-6-P pathway			2.1	2.7
TM0548–TM0556	Leucine biosynthesis				2.9
TM0729–TM0732	Includes (P)PPGPP synthetase				2.1
TM0733–TM0736	Mannose-6-phosphate isomerase			–2.6	
TM0785–TM0786	Ferritin and encapsulin				–3.8
TM0787–TM0788	Thiamine biosynthesis				2.6
TM0816–TM0817	Multidrug efflux		5.1	3.8	
TM0823–TM0825	Tet ^r family, radical SAM		–2.3		3.4
TM0849–TM0851	Heat shock response				2.6
TM0868–TM0870	Cell division and redox			–2.8	–2.5
TM0979–TM0983	Intracellular sulfur oxidation		2.2		
TM1120–TM1122	Glycerol-3-P ABC transporter				3.7
TM1200–TM1204	Maltose ABC transporter				–2.1
TM1243–TM1253	Thiamine biosynthesis				2.4
TM1288–TM1292	MinD superfamily, radical SAM	2.1			
TM1400–TM1401	Serine metabolism				2.1
TM1429–TM1431	Glycerol utilization	5.9			6.5
TM1468–TM1505	Ribosomal proteins				–3.2
TM1538–TM1546	Flagellum protein		–2.1		2.1
TM1609–TM1617	ATP synthase				2.6
TM1659–TM1666	Sure, peptide deformylase			–2.2	–2.1
TM1695–TM1699	Protein synthesis/processing			–2.3	–2.3
TM1792–TM1802	CRISPR-associated proteins	2.1		–2.2	

^a E, exponential phase; S, stationary phase; 1, 1 g/liter YE; 5, 5 g/liter YE. Log₂ fold changes were averaged for the indicated ORFs and then transformed into fold changes. Blank spaces indicate that the fold change was not above 2-fold. Positive values are shaded.

fide oxidation in *Allochromatium vinosum* (27) and tRNA thiolation in *Escherichia coli* (28), respectively. TM0982 appears to be a sulfur transport membrane protein (29). TM0816 (a MarR [multiple antibiotic resistance regulator] homolog) and TM0817 (an AcrB homolog) were also upregulated in the leaner medium during stationary phase. In *E. coli*, MarR is a transcriptional repressor of the *acrAB* locus, among other genes. AcrAB is an efflux pump involved in resistance to many drugs, including tetracycline, chloramphenicol, and ampicillin (30). In this case, it is unclear whether TM0816 and TM0817 were induced in response to a general stress or the presence of a particular compound.

Response of small ORFs annotated as hypothetical proteins. As noted in Table 1, there are 113 ORFs corresponding to 100 or fewer amino acids that are annotated as hypothetical proteins in the *T. maritima* genome. Table 4 shows that many of these putative small ORFs generated transcripts for the growth conditions studied here, the levels of which varied significantly. Thirteen of these ORFs were differentially transcribed during phase transition on 1 or 5 g/liter YE. Only one (TM1275) was more highly tran-

scribed during exponential phase, while the other 12 were more highly transcribed during stationary phase. Three of these 12 (TM1309, TM1315, and TM1316) occur in a genomic locus (TM1298 to TM1336), which was chosen for further analysis.

Transcription of the TM1300s region. During growth on 1 g/liter YE, as mentioned above, the transition to stationary phase was associated with a large fraction of the cell population becoming bloated and coccoidal (Fig. 2B); this was not observed during growth on 5 g/liter YE (Fig. 2D). Only 12 ORFs were upregulated both in 1 g/liter YE at stationary phase versus 5 g/liter at stationary phase and in 1 g/liter at stationary phase versus 1 g/liter at exponential phase. Three of these ORFs were the xylobiose-binding protein of an ABC transporter (TM0071) (17), TM0816, and TM0817 (discussed above). The remaining nine are in a particular genomic locus (TM1298 to TM1336 or TM1300s) (Table 5) that was previously found to be responsive when *T. maritima* was grown on spent medium from another hyperthermophile, *Pyrococcus furiosus* (31). All nine TM1300s ORFs that were induced are found in four predicted operons, all of which have similar organi-

TABLE 4 Transcription of hypothetical ORFs with ≤100 aa in the *T. maritima* genome^a

ID	No. of aa	SP	LSM				ID	No. of aa	SP	LSM			
			1 E	1 S	5 E	5 S				1 E	1 S	5 E	5 S
TM0001	41	N	1.3	1.2	1.3	1.2	TM0999	32	N	-0.1	0.0	-0.2	-0.1
TM0002	41	N	0.6	1.0	0.5	1.0	TM1025	32	N	0.0	0.6	-0.2	0.3
TM0003	40	N	-0.3	0.0	-0.3	-0.1	TM1107	80	Y	-0.7	-0.1	-0.7	-0.2
TM0004	31	N	-3.2	-2.1	-3.1	-1.5	TM1108	78	N	-0.6	-0.4	-0.3	0.2
TM0026	68	N	2.3	1.8	1.7	1.1	TM1112	89	N	-0.7	2.3	-0.4	2.3
TM0045	74	N	1.9	2.1	1.7	1.8	TM1186	68	N	2.0	2.6	1.9	2.5
TM0046	39	N	0.3	0.8	0.3	1.1	TM1206	82	N	-1.3	-1.1	-1.6	-1.2
TM0154	76	N	2.7	3.2	2.7	3.0	TM1208	70	N	2.7	3.0	2.7	3.2
TM0170	73	N	0.6	0.3	0.7	0.0	TM1236	62	N	-0.9	-0.3	-0.8	0.0
TM0179	92	N	1.4	2.2	1.2	2.3	TM1242	30	N	2.3	2.2	2.1	1.8
TM0210	69	N	1.5	1.6	1.6	1.2	TM1244	82	N	2.4	3.6	1.6	4.0
TM0223	77	N	-0.4	-0.2	-0.5	-0.3	TM1266	82	N	0.4	1.4	1.7	2.2
TM0242	40	N	1.6	1.1	1.4	1.2	TM1268	81	N	-0.5	-0.1	0.6	0.9
TM0314	59	N	-0.4	-0.1	-0.2	-0.1	TM1275	87	N	1.9	0.9	1.4	1.0
TM0315	83	N	-0.4	0.1	-0.3	0.1	TM1285	41	N	-0.2	0.1	-0.4	0.1
TM0319	64	N	-5.8	-5.2	-5.7	-5.7	TM1298	89	N	0.8	0.9	0.4	0.6
TM0338	35	N	-0.5	0.6	-0.6	0.2	TM1299	37	N	0.5	0.7	0.4	0.8
TM0340	94	N	0.5	1.0	0.4	0.9	TM1300	79	Y	0.8	1.5	0.9	0.7
TM0341	58	N	-0.3	0.4	-0.4	0.0	TM1305	43	N	-0.4	0.3	-0.4	-0.1
TM0357	68	N	-0.3	0.2	-0.4	0.1	TM1308	31	N	0.0	0.3	-0.1	0.0
TM0377	32	N	0.3	0.6	0.2	0.4	TM1309	68	N	-1.2	-0.1	-1.3	-0.7
TM0391	31	N	0.3	0.8	0.4	0.8	TM1313	75	N	0.3	0.7	0.3	0.5
TM0450	100	N	-0.4	-0.1	-0.9	-0.6	TM1315	57	Y	-0.1	2.1	-0.4	0.3
TM0486	94	N	0.0	0.6	0.0	1.1	TM1316	31	N	0.5	5.7	-2.1	0.9
TM0504	42	N	1.6	1.5	1.2	0.8	TM1333	60	Y	1.5	2.5	1.3	2.3
TM0518	64	N	-0.7	0.1	-0.7	-0.1	TM1338	44	N	-0.2	0.3	-0.3	0.1
TM0566	64	N	1.7	1.9	1.4	2.0	TM1386	69	N	1.5	1.3	1.4	1.1
TM0611	56	N	-0.7	-0.5	-0.7	-0.4	TM1392	98	N	1.6	2.0	1.4	2.1
TM0612	39	N	-0.3	0.4	-0.3	0.2	TM1412	47	N	1.3	1.3	1.4	1.6
TM0617	30	N	0.1	0.3	0.0	0.4	TM1413	35	N	0.3	0.4	0.0	0.2
TM0623	79	N	0.6	1.3	0.5	1.3	TM1420	77	N	-0.1	0.3	-0.2	0.3
TM0636	69	N	-0.8	-1.5	-0.4	-1.2	TM1457	94	N	2.6	3.7	2.6	3.4
TM0637	41	N	0.2	0.7	0.4	0.7	TM1462	81	N	1.1	1.1	1.2	1.4
TM0693	90	N	1.3	1.2	1.2	1.4	TM1538	89	N	-0.2	0.2	0.0	0.6
TM0773	73	N	0.0	0.4	0.0	0.1	TM1567	75	N	1.5	1.5	1.2	1.6
TM0781	71	N	1.9	2.6	1.7	1.7	TM1630	100	N	2.3	2.6	2.2	2.9
TM0783	33	N	0.7	0.9	0.5	0.8	TM1649	93	N	2.6	2.1	2.3	2.4
TM0826	56	N	0.7	1.9	0.3	1.8	TM1669	34	N	-0.4	0.3	-0.5	0.2
TM0871	47	N	0.9	0.5	0.7	0.3	TM1671	31	N	-0.5	-0.1	-0.5	-0.1
TM0898	76	N	2.7	2.3	2.5	1.5	TM1680	67	N	1.7	1.3	0.6	1.1
TM0969	94	N	1.3	1.6	1.1	2.1	TM1681	71	N	0.3	-1.8	-1.6	-1.8
TM0970	41	N	0.3	0.3	0.0	0.4	TM1688	79	N	1.9	2.2	1.9	2.6
TM0971	30	N	1.3	1.1	0.8	1.2	TM1690	92	N	2.2	2.0	2.3	2.6
TM0974	96	N	-0.5	0.8	-0.6	0.9	TM1711	88	N	1.1	1.6	1.1	1.8
TM0976	33	N	0.1	0.7	-0.1	0.3	TM1712	88	N	0.5	0.7	0.5	0.4
TM0977	66	N	1.2	1.7	1.2	1.8	TM1732	98	N	0.1	0.5	0.2	0.6
TM0983	79	N	0.8	1.1	0.3	0.2	TM1786	59	N	1.1	1.2	0.7	1.3
TM0994	37	Y	0.6	0.7	0.4	0.5	TM1829	30	N	-0.1	0.1	-0.1	0.1
TM0995	38	N	-0.3	-0.2	-0.2	-0.2	TM1850	59	N	-0.2	3.9	-0.5	4.9
TM0996	80	N	0.4	0.6	0.5	0.4	TM1871	91	N	2.3	1.8	2.3	1.3

^a Least-square means (LSMs) are based on mixed-effects model estimates, with 0 as the average transcription level for all genes in this experimental loop. Coloring is proportional to transcription level, with 0 (black) as the average, >0 (red) as higher levels, and <0 (green) as lower levels. SP, signal peptide. Signal peptides were (yes [Y]) or were not (no [N]) predicted by SignalP 4.1, with “eukaryotes” selected as the organism group.

zations consisting of one or more small ORFs (31 to 79 aa), followed by one or two radical S-adenosylmethionine (SAM) proteins and then by multiple ABC-2-type transporter proteins or hypothetical proteins (see Fig. S1 in the supplemental material). Each of these four predicted operons is separated by putative transcriptional regulators and toxin-antitoxin pairs. When the ABC-2 transporter family was first described (32), known members in-

cluded drug resistance transporters and capsular polysaccharide transporters. All five of the radical SAM proteins in the TM1300s region (TM1301, TM1317, TM1324, TM1325, and TM1334) have a CX₃CX₂C motif in the N-terminal portion, which is very highly conserved among nearly all radical SAM proteins (33). This motif coordinates a [4Fe-4S] center, which donates an electron to form a 5'-deoxyadenosyl radical from cleavage of SAM (34). One of the

TABLE 5 Putative small ORFs in the TM1300s region^a

ID	Coordinate		No. of aa	pI	aa sequence
	Starting	Ending			
TM1299	1324834	1324947	37	7.8	MVLSKVDDFSILSLQNYFYAILSHCCVTLGVLHKT
TM1300	1325051	1325290	79	9.3	<u>MKIVRRIL</u> <u>TIVILL</u> <u>SLLYTLVFP</u> <u>SPSDDYQTTVRKKT</u> <u>TVVHTLLLEKSH</u> <u>SEDLT</u> <u>AGNQILRFN</u> <u>KS</u> <u>LARLL</u> ISNERGDKG
TM1305	1329161	1329292	43	6.0	MKIGLVQAGGLVAVLVVAIFLEKSARGCYFAKISALSESADID
TM1307-1	1329812	1329913	33	10.1	MPLKRSSIRFFGCSFTQHFLHMDIKRCTMRELS
TM1308	1330202	1330297	31	4.4	MAALGLLTRDINMLLGVAEMMVFILSVQVFL
TM1309	1330309	1330515	68	9.5	MFFRWLSQLPLSRVESIIYALQGFKMKAWISSAWAAVGLAYLVLAYFTLRLLIEFAARKYGNIDMY
TM1314-1	1334887	1335000	37	7.8	MVLSKVDDFSILSLQNYFYAILSHCCVTLGVLHKT
TM1315	1335104	1335277	57	6.0	<u>MKIVRRIL</u> <u>TIVILL</u> <u>SLLYTLVFP</u> <u>SPSDDYQTTVRKKT</u> <u>TVVHTFLIEKLH</u> <u>SEDLT</u> <u>VG</u> <u>N</u>
TM1315-1	1335040	1335186	48	10.0	MAFQNRKSLKDKHRSVLIRKGGENREKDHSHNSYSAFSLIHACVSQSF
TM1316	1335354	1335449	31	8.6	<u>MQKMPKPVVPNDYLN</u> <u>PGKGGAGLV</u> <u>CGWCG</u>
TM1316-1	1335410	1335547	45	12.4	MWRSWSRLWLVLWRKNSSRFFPPLAKGQGKGFIKRLYFILPRRF
TM1322-1	1342532	1342696	54	6.0	MKSGAENVLLERILEQPKLENHIFDKVMTGDQNAKGFDNLFDTRSPVYNHFR
TM1323	1342623	1342829	68	5.6	<u>MIRMRRLVLMII</u> <u>FLILVLLSTIIF</u> <u>ADKVK</u> <u>TDNETHSWKSEITEQVQVAPK</u> <u>SAATCEVTFKG</u> <u>STAGNQSF</u>
TM1323-1	1342934	1343101	55	3.7	MEFLNNPITIGKIGLE---CCQSNNGNCECFNPDCNGICDPS-CNAVCD-CNAL--CTLIVP
TM1323-2	1342992	1343132	46	8.8	MGTVLSVSIRIVMGYVTLHAMRYVIVMLYAHCVLCHKKEGENLWSS
TM1323-3	1343120	1343236	38	3.9	MEFLNDPLEVFKPVGDIMCCDKNGNCD-CINTVCND-CAC
TM1323-4	1343262	1343432	56	3.6	MEFLNFPDNLSP---CCSTNPNC-DWDLNCSCG-CFDDMCAG-CLVCSNPANCTSLAKN
TM1325-1	1346088	1346273	61	4.6	MEFLNDPVVDEKVLVSFCDLSLDRCGDRVGNCECTNSHCVGLCNTCEDHVPWWKRIFLWW
TM1332-1	1352560	1352724	54	8.1	MKSGAENVLLKRILEQPKLENHIFDKVMTGDQNAKGFDNLFDTRSPVYNHFR
TM1333	1352651	1352857	68	5.6	<u>MIRMRRLVLMII</u> <u>FLILVLLSTIIF</u> <u>ADKVK</u> <u>TDNETHSWKSEITEQVQVAPK</u> <u>SAATCEVTFKG</u> <u>STAGNQSF</u>
TM1333-1	1353021	1353185	54	10.6	MAPCRCGKVPPELLRRILMARQTQGTCKKLFERYIHSFNILFRRSEKDRWLK

^a ORFs highlighted in gray are supported by microarray data from this study or tiling array data (GEO series accession no. GSE28822). ORFs that were not originally annotated have been given a unique identifier (ID) by appending "-n" to the nearest annotated ID. Underlined amino acids (aa) are predicted signal peptides. Boldface amino acids are cysteines.

small ORFs in this region (TM1316) was the most highly transcribed gene in the genome during stationary phase on 1 g/liter YE, the gene with the highest response (27-fold) on 1 g/liter versus 5 g/liter YE during stationary phase and the gene with the highest response (35-fold) in stationary phase versus exponential phase during growth on 1 g/liter YE (Table 4; Fig. 3).

Production and localization of TM1316. To gain further insight into TM1316 and track its production in *T. maritima*, rabbit antibodies were raised against a synthetic version of the peptide. Using anti-TM1316 serum, a single band below 20 kDa was detected by Western blotting, consistent with the predicted size of TM1316 (3.2 kDa) (Fig. 4A). TM1316 was not strongly detected during growth on 1 g/liter YE until stationary phase, and the amount of TM1316 increased significantly as stationary phase progressed (Fig. 1 and 4D). During growth on 5 g/liter YE, much lower levels of TM1316 were detected in early stationary phase, and TM1316 signal disappeared as stationary phase progressed. Further investigation showed that TM1316 production progressively decreased as YE concentration was increased (Fig. 4D), and TM1316 production was also repressed on media containing 1 g/liter YE plus 5 g/liter tryptone (data not shown).

To further investigate the regulation of TM1316 and the rest of the TM1300s region, two proteins in this region with identified helix-turn-helix domains (TM1314 and TM1330) were cloned and expressed in *E. coli*. Expression of TM1330, but not TM1314, resulted in soluble protein. Untagged TM1330 could be copurified with 6×His-tagged TM1331 by affinity chromatography (data not shown), suggesting that these proteins interact. This observation is consistent with the prediction that these two proteins comprise a toxin-antitoxin pair (35), although the specific targets and regulatory function of TM1330 to TM1331 are not clear.

For initial attempts to detect TM1316 in *T. maritima*, cultures were fractionated and concentrated to prepare extracellular supernatant, soluble lysate, and insoluble lysate samples. It was subsequently discovered that TM1316 could be readily detected in *T. maritima* culture without any concentration, cellular lysis, or processing; culture samples could simply be spotted onto nitrocellulose for dot blot analysis. If the cells were removed by centrifugation or filtration, most (if not all) of the TM1316 signal disappeared (Fig. 4B). TM1316 was not released by a freeze-thaw method previously used for examination of maltose binding activity (20), suggesting that it is not located in the periplasm or the outer membrane. When cells were lysed, TM1316 was distributed approximately equally between the soluble and insoluble fractions (Fig. 4B). Each fraction clearly exhibited less TM1316 signal than the whole-cell lysate or the intact cells. All of this evidence taken together suggests that TM1316 is associated with either the cytoplasmic membrane or an extracellular matrix/biofilm, and it is partially released from this membrane/matrix when the cells are lysed. This hypothesis is supported by Fig. 4C, where TM1316 antibodies demonstrated that the peptide was closely associated with cells that had formed aggregates during stationary-phase growth on 1 g/liter YE.

DISCUSSION

The small ORFs in the TM1300s region have no apparent homology to proteins outside the *Thermotogales*. The radical SAM enzymes, which are adjacent to the small ORFs, do have conserved motifs identifiable in similar types of enzymes encoded in other genomes. In addition to the N-terminal CX₃CX₂C motif that is present in nearly all members of the radical SAM superfamily, seven cysteines with fairly conserved spacing near the C terminus

FIG 3 TM1316 locus organization and gene response during growth on 1 or 5 g/liter yeast extract. (A) The TM1316 locus includes small hypothetical ORFs (black), radical SAM proteins (blue), transporters (green), toxins-antitoxins (pink), previously unannotated small ORFs (gray), and a transcriptional regulator based on identifiable TPR and XRE family helix-turn-helix (HTH) domains (orange). Gene organization was generated using Vector NTI (Invitrogen). (B) Transcriptional responses of genes in the TM1316 locus. Results for genes that did not change more than 1.6-fold (no change in expression) are shaded in gray (indicated with a dash). Increasing fold changes are marked by an increase in red hue. Colors shading the probe identifiers correspond with the TM1316 locus map. (C) Comparison of TM1316 and the mature form of subtilisin. Black lines indicate known intrachain linkages in subtilisin A. Cysteines are highlighted in black, and other similar or identical amino acids common to the two peptides are highlighted in gray. Amino acids are noted by their one-letter code. Red letters denote the N and C termini of the peptides.

that presumably coordinate another Fe-S cluster are present in four of the proteins in TM1300s (Fig. S2). The fifth protein (TM1334) has five of these seven cysteines, with slightly different spacing. The subset of radical SAM enzymes with such a C-terminal domain is described by the hidden Markov model (HMM) for TIGRFAM accession number [TIGR04085](#) (36). These proteins usually occur in genomes near small peptides, suggesting that they are involved in peptide modification.

Unfortunately, to date, very few of these radical SAM proteins have been characterized biochemically. These include PqqE, a radical SAM enzyme that is involved in pyrroloquinoline quinone (PQQ) biosynthesis (37); Rv0693 from *Mycobacterium tuberculosis*, another such enzyme, is involved in the maturation of mycofactocin, a putative electron carrier (38). In *B. subtilis*, the radical SAM protein Alba has been shown to contribute to the maturation of the anti-listerial bacteriocin subtilisin A by catalyzing the formation of three linkages between cysteine sulfurs and amino acid α -carbons (39). Similar radical SAM enzymes are presumed to be involved in the production of the sporulation killing factor in *B. subtilis* and the bacteriocins thuricin CD and thurincin H from *B. thuringiensis* (40, 41). To further explore the function of radical SAM proteins in TM1300s, it is important to consider the neighboring ORFs. PqqE and Rv0693-like radical SAM enzymes often cooccur with cofactor-dependent enzymes, while genes involved in bacteriocin synthesis generally cooccur with transporters. From this perspective, the genomic organization of TM1300s is clearly more consistent with that of known bacteriocin production genes (see Fig. S1 in the supplemental material and Fig. 3). However, because these types of radical SAM enzymes and their neighboring genes are widespread and not well characterized, other yet-to-be-discovered roles are likely (36). It is also worth noting that the

TM1300s region does not include an identifiable protease for leader peptide cleavage that is common among known bacteriocins.

The small peptide examined here (TM1316) does have a spacing of cysteines and a length similar to those of subtilisin A (Fig. 3). However, TM1316 is apparently primarily localized to the cell membrane (Fig. 4C), and it is unclear whether this peptide is a bacteriocin. Most characterized bacteriocins have been purified from cell-free supernatants, but it is possible that a bacterium can direct a bacteriocin to its membrane, where that bacteriocin either awaits a signal or is slowly released from the cell. To test for bacteriocidal activity, log-phase cultures of *T. maritima*, *T. neapolitana*, and the hyperthermophilic archaeon *Pyrococcus furiosus* were spiked with concentrated cell-free supernatants from *T. maritima*, but no effects on growth were observed. Because TM1316 was generally not detected in cell-free supernatants, log-phase *T. maritima* cultures were spiked with concentrated cell lysate containing TM1316, but again, no effect was observed. Alternatively, it is possible that under these conditions, TM1316 was detected primarily in the membranes of cells targeted for degradation, in a manner similar to that of the sporulation killing factor from *B. subtilis* (42). If this is the case, it would be interesting to determine whether or not TM1316 is detected primarily in the coccoid cells that were observed during stationary phase on 1 g/liter YE (Fig. 2B and 4C). However, there is no evidence at present to support this theory, and it is more likely that the coccoid morphology can be attributed to nutritional limitations. TM1316 might also be an immunity protein for a bacteriocin encoded elsewhere in the TM1300s region or serve as a modifier of other membrane proteins or signal transduction systems.

Nevertheless, the TM1300s region is intriguing because of its

FIG 4 Detection of the small peptide TM1316 in fractionated *T. maritima* cells. (A) A single band is detected in whole-cell lysate by Western blotting. (B) Dot blot analysis of cell fractions, including cell-free culture supernatant (lane 1), whole-cell lysate (lane 2), lysate clarified by centrifugation (lane 3), and lysate clarified by centrifugation and sterile filtration (lane 4). (C) Detection of TM1316 associated with aggregated cells by immunofluorescence. Cells were stained with 4',6-diamidino-2-phenylindole (DAPI) (blue) and CellMask deep red membrane stain (red) and incubated with rabbit anti-TM1316 serum and goat anti-rabbit IgG-DyLight 488 conjugate (green). (D) TM1316 detected by immunoblotting of unprocessed stationary-phase culture broth after 13.5 (early stationary phase) and 40.5 (late stationary phase) h of growth. Percentages are based on the maximum intensity level of TM1316 detected in the supernatant from a *T. maritima* culture in late stationary phase grown on 1 g/liter yeast extract.

repetitive organization, including four loci made up of small hypotheticals, radical SAM proteins, and transporters. We have been unable to locate a similar genomic region consisting of more than one or two loci in other species outside the *Thermotogales*. There are a number of interesting small ORFs in the TM1300s region, and it seems highly possible that at least one of them (if not TM1316) is a bacteriocin. More-detailed study of the region would benefit from more-accurate identification of small ORFs that are transcribed into protein. Nine ORFs less than 100 amino acids were originally annotated in this region (four of which appear to have signal peptides), although we have identified at least 12 more possible ORFs (Table 5). Certainly, some of these possible ORFs may not be translated, but others look promising and are preceded by possible ribosome binding sites. One of the originally annotated ORFs (TM1299) is nearly identical to one that was not annotated (TM1314-1) (Table 5), so it seems likely that either both of these ORFs are translated or neither of them is. Limited tiling array data are publicly available for *T. maritima* (GEO series accession number GSE28822), and these data were gathered under growth conditions where transcription from the TM1300s region is low. Despite this limitation, the tiling array data do provide evidence for the transcription of two unannotated ORFs identified here (TM1323-1 and TM1333-1) (Table 5). TM1323-1 (55 aa) is especially interesting because it contains 11 cysteines, and there are three other similar ORFs nearby that each contain at least 7 cysteines. A number of antimicrobial peptides have similarly cysteine-rich sections, including defensins (43) and lantibiotics (44). However, comparative genomic analysis of another uncharacterized group of cysteine-rich peptides, the SCIFF peptides (6 cys-

teines in 45 residues) suggests that they might not be pheromones or bacteriocins (36), so further investigation is needed to determine the function of small proteins in the TM1300s region.

Conclusion. Efforts to determine the role of specific ORFs, and small ORFs in particular, in the TM1300s region are ongoing. However, as indicated in Tables 1 and 4, microbial genomes bear many small ORFs, which, as is the case for *T. maritima*, are transcribed and potentially translated. Experimental and bioinformatics tools for studying the roles of these small ORFs are currently limited, as is knowledge about growth conditions that trigger their activation. However, the roles of these small ORFs in microbial physiology and ecology merit further attention, as insights into how microorganisms interact with other organisms could be forthcoming.

ACKNOWLEDGMENTS

This work was supported by the U.S. Department of Energy GTL Program (grant DG-FG02-08ER64687) and the National Science Foundation (grant CBET-0730091). A.D.F. acknowledges support from NIH Predoctoral Biotechnology Traineeships (NIH grant T32 GM008776-06).

Helpful discussions with T. Klenhammer (North Carolina State University), K. Noll (University of Connecticut), and D. Rodionov (Sanford-Burnham Medical Research Institute, La Jolla, CA) are also acknowledged.

REFERENCES

- Basrai MA, Hieter P, Boeke JD. 1997. Small open reading frames: beautiful needles in the haystack. *Genome Res.* 7:768–771.
- Hobbs EC, Fontaine F, Yin X, Storz G. 2011. An expanding universe of small proteins. *Curr. Opin. Microbiol.* 14:167–173.

3. Alix E, Blanc-Potard AB. 2009. Hydrophobic peptides: novel regulators within bacterial membrane. *Mol. Microbiol.* 72:5–11.
4. Waters LS, Sandoval M, Storz G. 2011. The *Escherichia coli* MntR mini-regulon includes genes encoding a small protein and an efflux pump required for manganese homeostasis. *J. Bacteriol.* 193:5887–5897.
5. Eguchi Y, Ishii E, Yamane M, Utsumi R. 2012. The connector SafA interacts with the multi-sensing domain of PhoQ in *Escherichia coli*. *Mol. Microbiol.* 85:299–313.
6. Gassel M, Mollenkamp T, Puppe W, Altendorf K. 1999. The KdpF subunit is part of the K(+)–translocating Kdp complex of *Escherichia coli* and is responsible for stabilization of the complex in vitro. *J. Biol. Chem.* 274:37901–37907.
7. Magnuson R, Solomon J, Grossman AD. 1994. Biochemical and genetic characterization of a competence pheromone from *B. subtilis*. *Cell* 77: 207–216.
8. Johnson MR, Montero CI, Connors SB, Shockley KR, Bridger SL, Kelly RM. 2005. Population density-dependent regulation of exopolysaccharide formation in the hyperthermophilic bacterium *Thermotoga maritima*. *Mol. Microbiol.* 55:664–674.
9. Montero CI, Lewis DL, Johnson MR, Connors SB, Nance EA, Nichols JD, Kelly RM. 2006. Colocation of genes encoding a tRNA-mRNA hybrid and a putative signaling peptide on complementary strands in the genome of the hyperthermophilic bacterium *Thermotoga maritima*. *J. Bacteriol.* 188:6802–6807.
10. Kaletta C, Entian KD. 1989. Nisin, a peptide antibiotic: cloning and sequencing of the *nisA* gene and posttranslational processing of its peptide product. *J. Bacteriol.* 171:1597–1601.
11. Zheng G, Yan LZ, Vederas JC, Zuber P. 1999. Genes of the *sbo-alb* locus of *Bacillus subtilis* are required for production of the antilisterial bacteriocin subtilisin. *J. Bacteriol.* 181:7346–7355.
12. Chhabra SR, Shockley KR, Ward DE, Kelly RM. 2002. Regulation of endo-acting glycosyl hydrolases in the hyperthermophilic bacterium *Thermotoga maritima* grown on glucan- and mannan-based polysaccharides. *Appl. Environ. Microbiol.* 68:545–554.
13. Hobbie JE, Daley RJ, Jasper S. 1977. Use of Nuclepore filters for counting bacteria by fluorescence microscopy. *Appl. Environ. Microbiol.* 33:1225–1228.
14. Chhabra SR, Shockley KR, Connors SB, Scott KL, Wolfinger RD, Kelly RM. 2003. Carbohydrate-induced differential gene expression patterns in the hyperthermophilic bacterium *Thermotoga maritima*. *J. Biol. Chem.* 278:7540–7552.
15. Frock AD, Gray SR, Kelly RM. 2012. Hyperthermophilic *Thermotoga* species differ with respect to specific carbohydrate transporters and glycoside hydrolases. *Appl. Environ. Microbiol.* 78:1978–1986.
16. Huber R, Langworthy TA, König H, Thomm M, Woese CR, Sleytr UB, Stetter KO. 1986. *Thermotoga maritima* sp.-nov represents a new genus of unique extremely thermophilic eubacteria growing up to 90 degrees C. *Arch. Microbiol.* 144:324–333.
17. Nanavati DM, Thirangoon K, Noll KM. 2006. Several archaeal homologs of putative oligopeptide-binding proteins encoded by *Thermotoga maritima* bind sugars. *Appl. Environ. Microbiol.* 72:1336–1345.
18. Boucher N, Noll KM. 2011. Ligands of thermophilic ABC transporters encoded in a newly sequenced genomic region of *Thermotoga maritima* MSB8 screened by differential scanning fluorimetry. *Appl. Environ. Microbiol.* 77:6395–6399.
19. Nanavati DM, Nguyen TN, Noll KM. 2005. Substrate specificities and expression patterns reflect the evolutionary divergence of maltose ABC transporters in *Thermotoga maritima*. *J. Bacteriol.* 187:2002–2009.
20. Nanavati D, Noll KM, Romano AH. 2002. Periplasmic maltose- and glucose-binding protein activities in cell-free extracts of *Thermotoga maritima*. *Microbiology* 148:3531–3537.
21. Le Fourn C, Fardeau ML, Ollivier B, Lojou E, Dolla A. 2008. The hyperthermophilic anaerobe *Thermotoga maritima* is able to cope with limited amount of oxygen: insights into its defence strategies. *Environ. Microbiol.* 10:1877–1887.
22. Novichkov PS, Laikova ON, Novichkova ES, Gelfand MS, Arkin AP, Dubchak I, Rodionov DA. 2010. RegPrecise: a database of curated genomic inferences of transcriptional regulatory interactions in prokaryotes. *Nucleic Acids Res.* 38:D111–D118.
23. Ravcheev DA, Li X, Latif H, Zengler K, Leyn SA, Korostelev YD, Kazakov AE, Novichkov PS, Osterman AL, Rodionov DA. 2012. Transcriptional regulation of central carbon and energy metabolism in bacteria by redox-responsive repressor Rex. *J. Bacteriol.* 194:1145–1157.
24. Yang C, Rodionov DA, Rodionova IA, Li X, Osterman AL. 2008. Glycerate 2-kinase of *Thermotoga maritima* and genomic reconstruction of related metabolic pathways. *J. Bacteriol.* 190:1773–1782.
25. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero DA, Horvath P. 2007. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 315:1709–1712.
26. Pysz MA, Ward DE, Shockley KR, Montero CI, Connors SB, Johnson MR, Kelly RM. 2004. Transcriptional analysis of dynamic heat-shock response by the hyperthermophilic bacterium *Thermotoga maritima*. *Extremophiles* 8:209–217.
27. Stockdreher Y, Venceslau SS, Josten M, Sahl HG, Pereira IA, Dahl C. 2012. Cytoplasmic sulfurtransferases in the purple sulfur bacterium *Allochrochromatium vinosum*: evidence for sulfur transfer from DsrEFH to DsrC. *PLoS One* 7:e40785. doi:10.1371/journal.pone.0040785.
28. Ikeuchi Y, Shigi N, Kato J, Nishimura A, Suzuki T. 2006. Mechanistic insights into sulfur relay by multiple sulfur mediators involved in thiouridine biosynthesis at tRNA wobble positions. *Mol. Cell* 21:97–108.
29. Zdobnov EM, Apweiler R. 2001. InterProScan—an integration platform for the signature-recognition methods in InterPro. *Bioinformatics* 17: 847–848.
30. Okusu H, Ma D, Nikaido H. 1996. AcrAB efflux pump plays a major role in the antibiotic resistance phenotype of *Escherichia coli* multiple-antibiotic-resistance (Mar) mutants. *J. Bacteriol.* 178:306–308.
31. Montero C. 2005. Molecular microbial ecology of the hyperthermophilic bacterium *Thermotoga maritima*: transcriptional and physiological response to antibiotic challenge and inter-species interactions. Ph.D. thesis. North Carolina State University, Raleigh, NC.
32. Reizer J, Reizer A, Saier MH, Jr. 1992. A new subfamily of bacterial ABC-type transport systems catalyzing export of drugs and carbohydrates. *Protein Sci.* 1:1326–1332.
33. Sofia HJ, Chen G, Hetzler BG, Reyes-Spindola JF, Miller NE. 2001. Radical SAM, a novel protein superfamily linking unresolved steps in familiar biosynthetic pathways with radical mechanisms: functional characterization using new analysis and information visualization methods. *Nucleic Acids Res.* 29:1097–1106.
34. Mulliez E, Padovani D, Atta M, Alcouffe C, Fontecave M. 2001. Activation of class III ribonucleotide reductase by flavodoxin: a protein radical-driven electron transfer to the iron-sulfur center. *Biochemistry* 40: 3730–3736.
35. Sevin EW, Barloy-Hubler F. 2007. RASTA-Bacteria: a web-based tool for identifying toxin-antitoxin loci in prokaryotes. *Genome Biol.* 8:R155. doi:10.1186/gb-2007-8-8-r155.
36. Haft DH, Basu MK. 2011. Biological systems discovery *in silico*: radical S-adenosylmethionine protein families and their target peptides for post-translational modification. *J. Bacteriol.* 193:2745–2755.
37. Wecksler SR, Stoll S, Tran H, Magnusson OT, Wu SP, King D, Britt RD, Klinman JP. 2009. Pyrroloquinoline quinone biogenesis: demonstration that PqqE from *Klebsiella pneumoniae* is a radical S-adenosyl-L-methionine enzyme. *Biochemistry* 48:10151–10161.
38. Haft DH. 2011. Bioinformatic evidence for a widely distributed, ribosomally produced electron carrier precursor, its maturation proteins, and its nicotinoprotein redox partners. *BMC Genomics* 12:21.
39. Fluhe L, Knappe TA, Gattner MJ, Schafer A, Burghaus O, Linne U, Marahiel MA. 2012. The radical SAM enzyme AlbA catalyzes thioether bond formation in subtilisin A. *Nat. Chem. Biol.* 8:350–357.
40. Lee H, Churey JJ, Worobo RW. 2009. Biosynthesis and transcriptional analysis of thurincin H, a tandem repeated bacteriocin genetic locus, produced by *Bacillus thuringiensis* SF361. *FEMS Microbiol. Lett.* 299:205–213.
41. Rea MC, Sit CS, Clayton E, O'Connor PM, Whittall RM, Zheng J, Vederas JC, Ross RP, Hill C. 2010. Thuricin CD, a posttranslationally modified bacteriocin with a narrow spectrum of activity against *Clostridium difficile*. *Proc. Natl. Acad. Sci. U. S. A.* 107:9352–9357.
42. Gonzalez-Pastor JE, Hobbs EC, Losick R. 2003. Cannibalism by sporulating bacteria. *Science* 301:510–513.
43. Bensch KW, Raida M, Magert HJ, Schulz-Knappe P, Forssmann WG. 1995. hBD-1: a novel beta-defensin from human plasma. *FEBS Lett.* 368: 331–335.
44. Banerjee S, Hansen JN. 1988. Structure and expression of a gene encoding the precursor of subtilin, a small protein antibiotic. *J. Biol. Chem.* 263:9508–9514.