December 13, 2014

FINAL REPORT

PROJECT NUMBER: **DE-FG02-08ER64687**

PROJECT TITLE: **Biohydrogenesis in the Thermotogales**

PERIOD COVERED: **SEPTEMBER 2008 – SEPTEMBER 2014**

Principal Investigator: **Robert M. Kelly**
Department of Chemical and Biomolecular Engineering
North Carolina State University

Collaborators: **Paul H. Blum**
Beadle Center for Genetics
University of Nebraska-Lincoln

Kenneth M. Noll
Department of Molecular and Cellular Biology
University of Connecticut

EXECUTIVE SUMMARY: The production and consumption of molecular hydrogen drives the physiology and bioenergetics of many microorganisms in hydrothermal environments. As such, the potential of these microorganisms as model systems to probe fundamental issues related to biohydrogen production merits consideration. It is important to understand how carbon/energy sources relate to the disposition of reducing power and, ultimately, the formation of molecular hydrogen by high temperature microorganisms. This project focused on bacteria in the thermophilic order Thermotogales, fermentative anaerobes that produce H_2 from simple and complex carbohydrates. The major thrusts of the project are summarized in the Objectives listed below:

OBJECTIVE 1: Examine the regulation of substrate catabolic proteins and pathways as this relates to carbon partitioning, disposition of reducing power, and H_2 generation in *Thermotoga maritima*.

OBJECTIVE 2: Apply classical genetics and develop molecular genetic tools for Thermotoga species to dissect catabolic and regulatory pathways related to sugar metabolism and H_2 evolution.

OBJECTIVE 3: Thermotogales biodiversity arises from adaptive specialization that expands on a conserved minimal genome; physiological characterization of selected novel traits will be done to expand understanding of biohydrogenesis.

Four species within the genus Thermotoga were examined to understand similarities and differences in the mechanisms by which simple and complex carbohydrates were utilized and converted to molecular hydrogen. Although the core genome of these four species represented 75% of open reading frames (ORFs), there were significant differences in carbohydrate utilization patterns. New ABC transporters were identified within the Thermotogales through genomic and biochemical analysis. Molecular genetics tools were developed to examine *Thermotoga maritima* physiology. Cell lines were created in which both H_2 and acetate levels were elevated on a per cell basis relative to the wild type, while lactate remained undetectable. Genome resequencing indicated that the primary genetic target for these phenotypic changes was the ATP binding component of a maltose ABC transporter. High temperature anaerobic [^{14}C]-maltose transport assays demonstrated maltose uptake was reduced in the H_2 overproducing cell lines. This suggested normal rates of maltose transport in the wild type organism lead to a metabolic imbalance that limited H_2 synthesis. The microbial ecology of *T. maritima* was examined through functional genomics experiments. Under low nutrient conditions, *T. maritima* was observed to produce a range of putative peptides, some of which were related to α-carbon cyclic peptides produced by *Bacillus subtilus*. Finally, the role of 'toga' in these novel microorganisms was shown to involve association with insoluble growth substrates. The 'toga' distends from the cytoplasmic membrane-enclosed portion of the cells as they enter the late exponential/stationary phase of growth. Some of the genes encoding toga-associated proteins were up-regulated during this phase of growth and the distension is caused by continued growth of the toga, and not shrinkage of the cytoplasmic aspect of the cells. This increase in cell surface area may have selective value to provide a larger anchor for polysaccharide hydrolytic enzymes during a time of nutritional stress. This project led to many interesting insights about the Thermotogales that have both scientific and technological implications. Ongoing work will leverage these developments to further elucidate many interesting features of these novel microorganisms.

Carbohydrate utilization patterns in *Thermotoga sp*.

Four hyperthermophilic members of the bacterial genus *Thermotoga* (*T. maritima, T. neapolitana, T. petrophila,* and *T.* sp. RQ2) share a core genome of 1470 open reading frames (ORFs), or about 75% of their genomes (see **Figure 1**) (Frock et al., 2012). Nonetheless, each species exhibited certain distinguishing features during growth on simple and complex carbohydrates that correlated with genomic inventories of specific ABC sugar transporters and glycoside hydrolases. These differences were consistent with transcriptomic analysis based on a multi-species cDNA microarray (see **Table 1**). Growth on a mixture of six pentoses and hexoses showed no significant utilization of galactose or mannose by any of the four species. *T. maritima* and *T. neapolitana* exhibited similar monosaccharide utilization profiles, with a strong preference for glucose and xylose over fructose and arabinose. *T.* sp. RQ2 also used glucose and xylose, but was the only species to utilize fructose to any extent, consistent with a phosphotransferase system (PTS) specific for this sugar encoded in its genome. *T. petrophila* used glucose to a significantly lesser extent than the other species. In fact, the XylR regulon was triggered by growth on glucose for *T. petrophila*, which was attributed to the absence of a glucose transporter (XylE2F2K2), otherwise present in the other Thermotoga species. This suggested that *T. petrophila* acquires glucose through the XylE1F1K1 transporter, which primarily serves to transport xylose in the other three Thermotoga species. The results here show that subtle differences exist among the hyperthermophilic Thermotogales with respect to carbohydrate utilization that support their designation as separate species. Of course, ultimately more quantitative measures are needed to differentiate among species, perhaps not only based on genome sequence but also on other 'omics' features, including transcriptome.

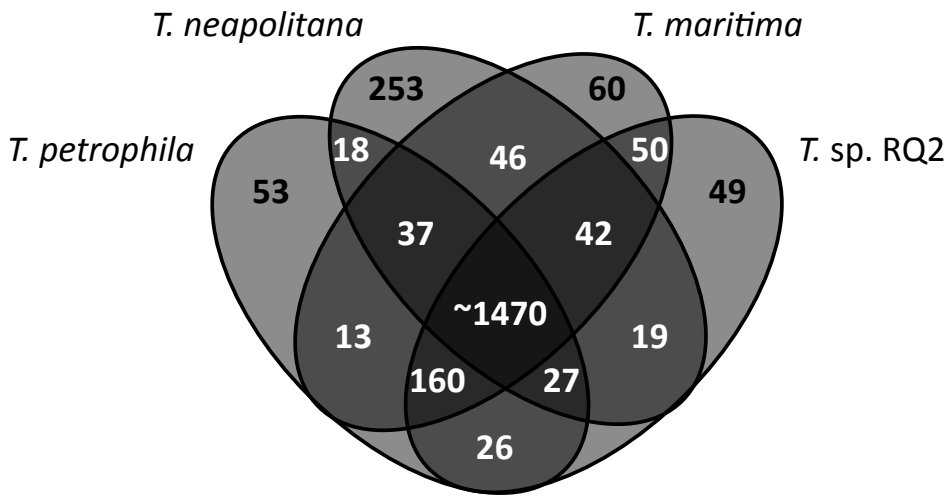


Figure 1. Venn diagram of the shared ORFs of the selected *Thermotoga* species, based on 70% identity and 80% coverage at the amino acid level over the entire ORF. Numbers indicate the ORFs shared by subsets of species.

Attempts to establish a mixed culture of the four Thermotoga species growing on a mixed α/β-polysaccharide medium were unsuccessful. *T. sp.* RQ2 was found to quickly outgrow the other three species, which likely related to the wider array of glycoside hydrolases present in its genome relative to the other species. Within a few generations, the other three species (based on qPCR-based analysis using genes unique to individual species) reached very low levels in the culture, having been outcompeted. Clearly, this result reflects the complexity of community interactions in natural environments and the critical role that ABC sugar transporters and glycoside hydrolases play in survival and proliferation of Thermotoga species.

Table 1. ORFs responding in hyperthermophilic *Thermotoga* species during growth on polysaccharide mixture compared to growth on glucose. Values displayed are fold changes. Cells are shaded as follows: significantly up-regulated on the polysaccharide mix (black), significantly up-regulated on glucose (white), no statistically significant change (light gray), absent from the indicated genome (dark gray). Column labels represent *T. maritima* (M), *T. neapolitana* (N), *T. petrophila* (P), and *T.* sp. RQ2 (R). The designation "TRQ2_1647a" indicates that there is more than one microarray probe for TRQ2_1647.

Probe	Product	M	N	P	R
Mannan utilization					
TM1746	mannan ABC transporter, periplasmic oligopeptide-binding protein	4.6			3.7
TM1747	mannan ABC transporter, permease protein	3.7			8.3
TM1748	mannan ABC transporter, permease protein				
TM1749	mannan ABC transporter, ATP-binding protein	6.4			4.3
TM1750	mannan ABC transporter, ATP-binding protein	5.5			6.7
TM1751	endoglucanase	5.9			2.5
ManR operon					
TM1224	ManR transcriptional regulator, ROK family	3.3	2.4		8.7
TM1226	mannooligosaccharide ABC transporter, sugar-binding protein	5.4			18.3
TM1227	endo-1,4-beta-mannosidase	24.1	12.7	3.9	15.3
CelR regulon					
TM0312	predicted dehydrogenase	2.1	2.1		
TM0313	predicted aldo/keto reductase	2.0		2.6	
TM1218	CelR transcriptional regulator, LacI family	3.8		3.9	
TM1219	mannobiose ABC transporter, ATP-binding protein		2.6		2.5
TM1223	mannobiose ABC transporter, sugar-binding protein	10.0	25.4	4.9	18.8
TM1524	endoglucanase	7.7		6.8	3.9
TM1525	endoglucanase	9.5		7.3	
TM1848	cellobiose phosphorylase	10.2	4.7	19.7	9.4
TRAP transporter					
TM0322	TRAP transporter, periplasmic substrate-binding protein, putative		8.1		14.7
TM0323	TRAP transporter, small transmembrane component				7.0
TM0324	TRAP transporter, large transmembrane component				3.9
TM0325	predicted sugar dehydrogenase		2.8		12.0
TM0326	transcriptional regulator, RpiR family	-3.6			2.6
TM0327	phosphoglycerate dehydrogenase, putative				6.2
Processing of α-glucans					
TM0752	alpha-glucosidase, putative	5.3			2.2
TM0767	maltodextrin glycosyltransferase		-2.7		-6.1
TM1834	alpha-glucosidase				-3.6
TM1839	maltose ABC transporter, periplasmic maltose-binding protein		-3.8	-2.3	-6.6
TM1841	hypothetical protein				-2.5
TM1842	hypothetical protein		-2.6	-3.2	-5.9
TM1843	hypothetical protein				-3.2
TM1844	hypothetical protein				-5.1
TM1845	pullulanase			-5.1	
CTN_1407	alpha-glucan phosphorylase				-2.3
TRQ2_1647a			-2.2		-2.5
BglR operon					
TM0023	methyl-accepting chemotaxis protein			-7.8	-2.4
TM0024	laminarinase	-2.6		-11.8	
TM0025	beta-glucosidase	-2.1	2.1	-6.7	
TM0026	hypothetical protein			-3.1	
TM0027	cellobiose/laminaribiose ABC transporter, ATP-binding protein	-2.0		-4.7	-2.6
TM0028	cellobiose/laminaribiose ABC transporter, ATP-binding protein	-2.0		-8.0	-2.4
TM0029	cellobiose/laminaribiose ABC transporter, permease protein			-5.0	-2.6
TM0031	cellobiose/laminaribiose ABC transporter, sugar-binding protein	-2.7		-11.6	-2.7
TM0032	BglR transcriptional regulator, ROK family			-4.2	

Microbial ecology of *Thermotoga sp.* The genome of the hyperthermophilic bacterium *Thermotoga maritima* encodes numerous putative peptides/proteins of 100 amino acids or less. While most of these 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 up-regulation of several ORFs encoded in the TM1298-1336 genome locus (Frock et al., 2013). This region contains putative HicAB toxin/antitoxin pairs, hypothetical proteins, radical SAM enzymes, and ABC transporters. Of particular note was the TM1315-1319 operon, which includes a putative 31-amino acid peptide (TM1316) that was the most highly transcribed gene in the transcriptome during stationary phase (see **Figure 2**). 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 (see **Figure 3**). The only functionally characterized locus with similar organization to the TM1315-1319 is in *Bacillus subtilis*, which contains subtilosin A, a cyclic peptide with Cys to α-carbon linkages that functions as an anti-listerial bacteriocin. While the organization of TM1316 resembled the Bacillus peptide (e.g., number of amino acids, spacing of Cys residues) (see **Figure 4**), preparations containing high levels of TM1316 affected neither the growth of 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 TM1298-1336 locus, and, while their roles are also unclear, they merit examination as potential anti-microbial agents in hyperthermophilic biotopes.

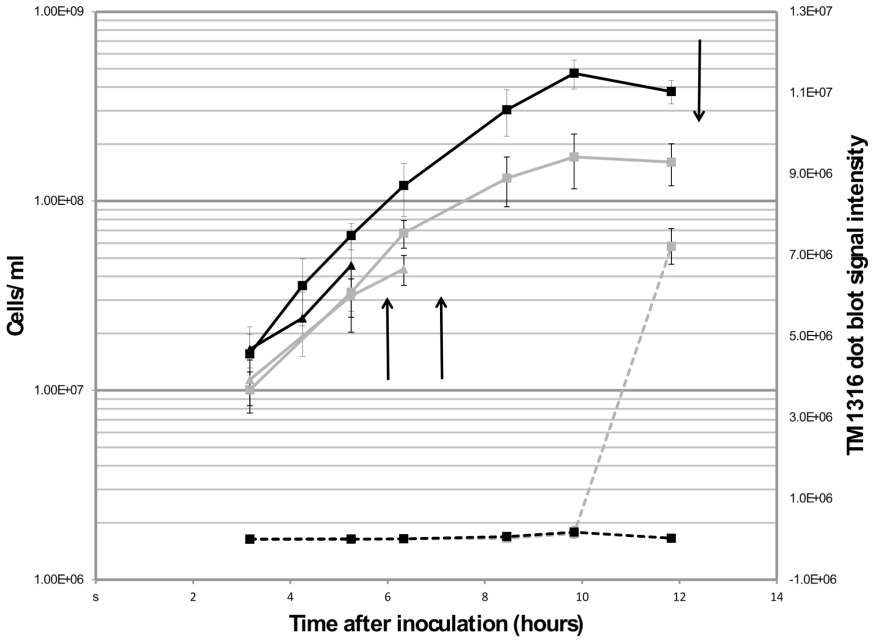


Figure 2. Growth of *T. maritima* and production of a small peptide encoded in TM1316 varies with the levels of yeast extract in culture medium. Cell density of *T. maritima* when grown on 1 g/l (grey) versus 5 g/l (black) yeast extract are denoted by solid lines. TM1316 abundance, estimated by immunoblot intensity, is denoted by dotted lines for 1 g/l (grey) and 5 g/l (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/l yeast extract medium.

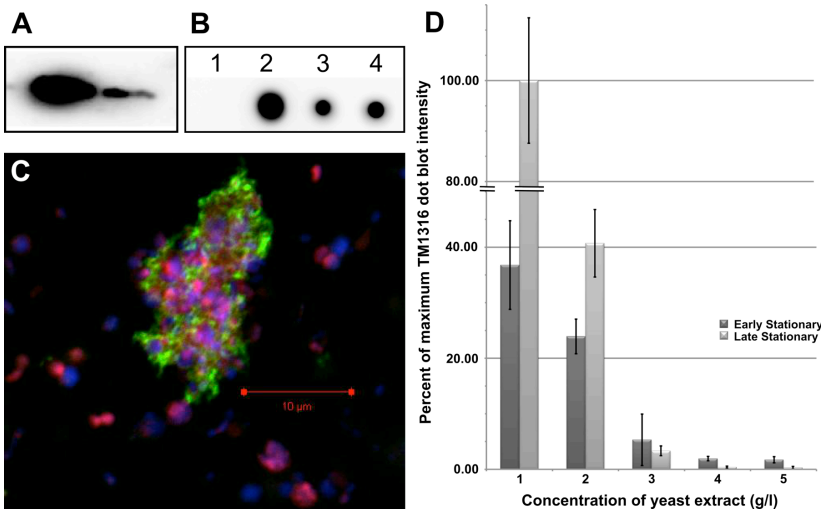
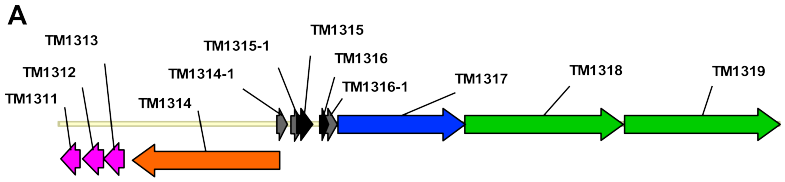


Figure 3. Detection of the small peptide TM1316 in fractionated *T. maritima* cells. (A) A single band is detected in whole cell lysate by western blot. (B) Dot blot of cell fractions including: (1) cell-free culture supernatant, (2) whole-cell lysate, (3) lysate clarified by centrifugation, (4) lysate clarified by centrifugation and sterile filtration. (C) Detection of TM1316 associated with aggregated cells by immunofluorescence. Cells were stained with 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 immunoblot of unprocessed stationary phase culture broth after 13.5 (early stationary) and 40.5 hours (late stationary) 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/l yeast extract.



B

Probe ID	Fold Change			
	1E vs. 5E	1S vs. 5S	1S vs. 1E	5S vs. 5E
TM1311	-	-	-	-
TM1312	-	-	-	-
TM1313	-	-	-	-
TM1314	-	-	-	-
TM1315	-	3.6	4.5	1.6
TM1316	6.1	27.3	34.7	7.7
TM1317	-	7.4	12.6	2.4
TM1318	-	7.1	5.4	-
TM1319	-	3.1	1.9	-

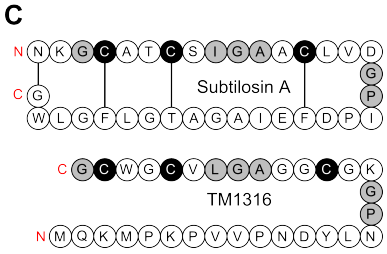


Figure 4. TM1316 locus organization and gene response during growth on 1 or 5 g/l yeast extract. (A) The TM1316 locus includes small hypothetical ORFs (black), radical SAM proteins (blue), transporters (green), toxins/antitoxins (pink), previously unannotated small ORFs predicted by RASTA (gray), and a transcriptional regulator based on identifiable TPR and XRE family HTH domains (orange). Gene organization was generated using Vector NTI (Invitrogen). (B) Transcriptional response of genes in the TM1316 locus. Genes that did not change more than 1.6-fold (no change in expression) are denoted in grey with a dash. Increasing fold-changes are marked by an increase of red hue. Colors shading the Probe IDs correspond with the TM1316 locus map. (C) Comparison of TM1316 and the mature form of subtilosin A. Black lines indicate known intra-chain linkages in subtilosin A. Cysteines are highlighted in black, 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 note the N- and C- termini of the peptides.

Draft genome sequence of an evolved *Thermotoga maritima* isolate. *Thermotoga maritima* (Tma) is a hyperthermophilic bacterium with a small genome (1.86 Mbp). A new derivative was isolated by experimental microbial evolution that underwent significant deletion formation. A draft genome sequence of this isolate, Tma200, was completed and a manuscript will be submitted on this subject in the near future.

Uncoupling fermentative synthesis of molecular hydrogen from biomass formation in *Thermotoga* (see Figures 5-10)**.** The yield of biologically produced metabolites is constrained by the energy inherent to the starting substrates or by the toxicity of the resulting products. For example, when carbohydrates are fermented by the hyperthermophilic anaerobe, *Thermotoga maritima*, the end product molecular hydrogen (H_2), is formed in strict proportion to substrate availability while inhibiting growth. Consequently, organic acid excretion is an essential trait required for H_2 production to maintain redox homeostasis. Here we describe the use of transient gene inactivation (TGI) targeting organic acid synthesis to isolate hydrogen overproducing cell lines of this organism. Lactate dehydrogenase (*ldh*) was disrupted using a truncated copy of *ldh* fused to a kanamycin resistance cassette expressed from a native groESLp promoter. Passage of the unstable recombinant resulted in cell line Tma100 that exhibited decreased utilization of maltose and loss of the disruption cassette along with a maltose-specific small colony phenotype. Continued passage in the presence of maltose yielded additional cell lines including a large colony variant termed Tma200. Both H_2 and acetate levels were elevated on a per cell basis for Tma100 and Tma200 relative to the wild type while lactate remained undetectable. Metabolite overproduction was evident with or without headspace exchange eliminating a direct effect of H_2. Genome resequencing indicated that the primary genetic target for these phenotypic changes was the ATP binding component of a maltose ABC transporter. High temperature anaerobic [^{14}C]-maltose transport assays demonstrated maltose uptake was reduced in the H_2 overproducing cell lines. This suggested normal rates of maltose transport in the wild type organism lead to a metabolic imbalance that limited H_2 synthesis. TGI could be applied to other organisms to modulate yields of metabolic products where target genes are essential.

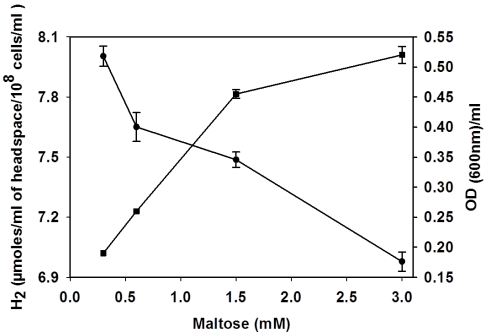


Figure 5 Relationship between maltose concentration and H_2 production. *Tma* (*wild type*) was grown in various concentrations of maltose and H_2 production (●) and growth of Wild type (■) was estimated. Accumulated H_2 from each sugar concentration was normalized to 10^8 cells/ml. The error bar represents the standard deviations from biological replicates.

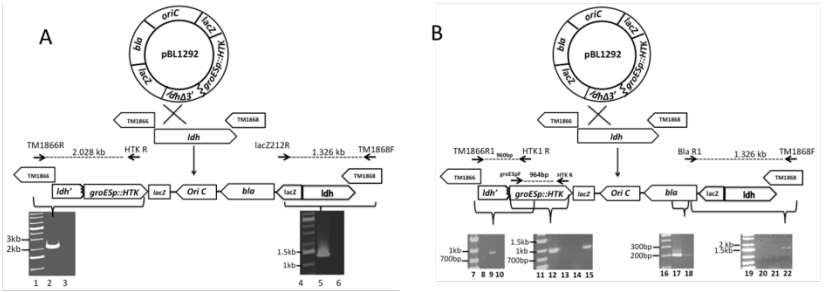


Figure 6 Targeted disruption of lactate dehydrogenase *(ldh)* via homologous recombination in *Tma*. Schematic of *ldh* disruption by single crossover and PCR amplification of predicted amplicons (2A liquid enrichment and 2B *Tma100*). Lanes1,4,7,11,16 and 19 represent molecular marker, Lanes 2,3,8,9,10 represent unique 5' fusion joint at *ldh* locus in liquid enrichment, *Tma*, *Kan*^{R} mutant, *Tma100* and *Tma*, respectively. Lanes 5,6,20,21 and 22 represent and 3' fusion joint at *ldh* locus in liquid enrichment, *Tma*, *Kan*^{R} mutant, *Tma* and *Tma100*, respectively. Lanes 12,13,14 and 15 represent selectable marker in pBL1292, *Tma*, *Kan*^{R} mutant and *Tma100*, respectively. Lanes 17 and 18 represents bla gene in pBL1292 and *Tma100*.

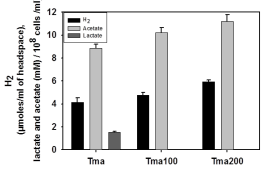


Figure 7 Comparison of Metabolites (H_2 and organic acids) of wild type, *Tma100* and *Tma200* in batch culture. Hydrogen and organic acid production was normalized to 10^8 cell/ml for all three strains. Error bar represents the standard deviations from biological replicates.

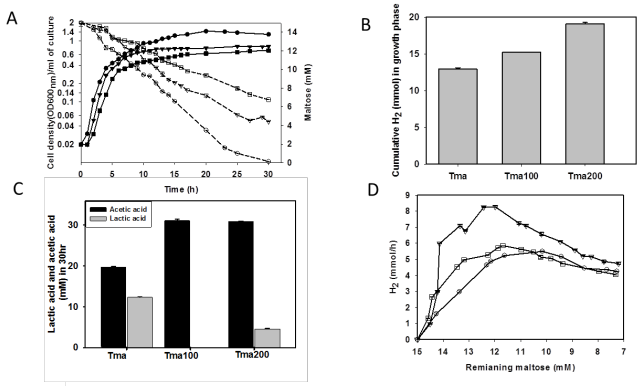


Figure 8 Fermentation profile of *Tma, Tma100* and *Tma200* grown in a 2L fermenter. 4**A.** A relationship between growth (filled symbols; *Tma*(●), *Tma100*(■) and *Tma200*(▼)) and maltose utilization (open symbols; *Tma*(○), *Tma100*(□) and *Tma200*(▽)). 4**B.** Comparison of H_2 produced (cumulative) by *Tma*, *Tma100* and *Tma200* in growth phase. 4**C.** Normalized organic acids (lactate and acetate) of *Tma100* and *Tma200* to *Tma* biomass produced in 30hr. 4**D.** A relationship between H_2 production rate and maltose utilization in *Tma* (○), *Tma100* (□) and *Tma200* (▽). Error bar represents the standard deviations from a triplicate analysis.

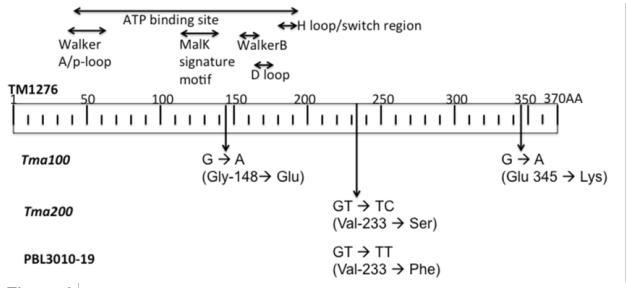


Figure 9 Domain structure of TM1276 (MalK) along with mutations present in *Tma100* and *Tma200*. Mutated amino acids sequence with coordinates are shown.

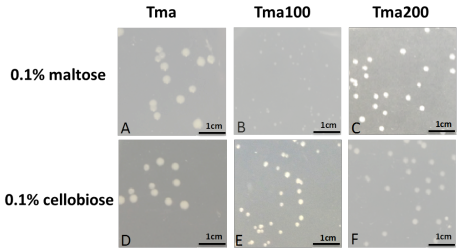


Figure 10 Colony size variation in *Tma* (wild-type), *Tma100* and *Tma200* grown on complex medium plates supplemented with 0.1% maltose (panel A, B and C) or 0.1% cellobiose (panel D, E and F), respectively. The *Tma100* formed smaller colonies on CM plates with 0.1% maltose (panel B) and bigger colonies on CM plates with 0.1% cellobiose (panel E) comparative to *Tma200* colonies on CM plates with 0.1% cellobiose (panel F). *Tma200* formed same size colonies on 0.1% maltose (panel C) and 0.1% cellobiose (panel F) as compared to the *Tma* control on CM plates (Panel A and D) containing different carbon sources. Scale bar is equal to 1cm.

Thermotogales biodiversity arises from adaptive specialization that expands on a conserved minimal genome. Characterization of ABC transporters in *T. maritima* was done by developing a new method to more rapidly screen recombinant substrate binding proteins for their ligand affinities (10). We modified a differential scanning fluorimetry (DSF) method to allow us to measure substrate binding at high temperatures, up to the limit of the PCR thermocycler, ~96°C. Using this method, numerous ligands could be screened over several days. This allows the reexamination of those that showed possible binding using fluorescence spectroscopy to measure binding affinities. We found that DFS also allowed us to measure the thermal stability of these binding proteins as well as other thermophile proteins. DSF allowed us to discover two new ABC transporter operons in *T. maritima*, one for trehalose and the other a second xylose transporter, that were missed in the original genome sequencing project due to an apparent deletion mutation in the sequenced strain (see Table 2). We have also used DSF and fluorescence spectroscopy to examine the substrate ranges of several evolutionarily related mannose-binding proteins from Thermotogales species. We found slight differences in ligand affinities and related these to the evolutionary history of these operons. This work provides information about the evolutionary processes that lead to new ligand affinities and may impact future genome annotations and efforts to evolve these transporters for new functions.

We contributed to a study of myo-inositol (MI) degradation by *T. maritima* in an effort to increase our knowledge of the catabolic capacities of this organism (Rodionova et al., 2013). A novel variant of the MI catabolic pathway was discovered in *T. maritima*. Recombinant enzymes of the pathway were examined using in vitro biochemical assays. *T. maritima* was shown to be unable to grow on myo-inositol as a single carbon source. We showed that InoEFGK (TM0418–TM0421) encodes an ABC transporter for myoinositol- phosphate suggesting that this novel pathway utilizes a phosphorylated derivative of inositol.

TABLE 2. Apparent K_d values at 60°C of TreE and XylE2 for different sugars as measured by changes in intrinsic fluorescence upon ligand binding

Protein	Sugar	K_d value at 60°C (μM)
TreE	Trehalose	0.024
	Sucrose	0.300
	Glucose	56.78
XylE2	Glucose	0.059
	Xylose	0.042
	L-Fucose	1.436

Related to the above study, we examined one enzyme of the MI degradation pathway, the myo-inositol-3-phosphate synthase (MIPS) in detail (Butzin et al., 2013). The *Thermotoga* lineage apparently inherited the MIPS gene from an archaeal Thermococcales donor (see Table 3). We obtained recombinant enzymes from several Thermotogales and some Thermococcales species to measure their catalytic activities and thermal stabilities. We compared these properties with those of 8 artificially synthesized MIPS proteins that we determined should be similar to the ancestors of extant MIPS proteins. We found that the ancestral proteins from both the apparent archaeal donor and the *Thermotoga* most recent common ancestor recipient were more thermostable than their descendants. We found a correlation between the thermostability of MIPS proteins and the optimal growth temperatures of their hosts, suggesting that the ancestral species grew at higher optimal growth temperatures than their descendants. Our method for reconstruction of ancestral genes that created derivatives more thermostable than their extant homologs could be used to engineer novel properties into organisms or enzymes.

Table 3. Measures of thermostability of extant and ancestral MIPSs.

Organism	Protein	OGT (°C)	T_{opt} (°C)	T_m (°C)			IVYWREL value
				pH 7.0	pH 4.2	pH 3.5	
Tc. sibiricus MM 739	TSIB_1788	78	ND	81.1±0.06	63.5±0.20	ND	0.4184
Tt. maritima MSB8	TM1419	80	75	81.0±0.16	64.5±0.17	ND	0.4031
Thermotoga sp. str. RQ2	TRQ2_1313	80	80	85.2±0.07	63.6±0.14	ND	0.4005
	ATM_T1		90	88.6±0.05	66.7±0.09	44.7±0.47	0.4084
	ATM_T2		83	88.8±0.02	67.5±0.09	47.9±0.07	0.4110
	ATM_T3		85	88.9±0.08	68.1±0.17	47.4±0.13	0.4110
	ATM_T4		83	88.9±0.01	68.3±0.39	49.2±0.13	0.4136
	ACM_C1		95	>99	76.6±0.11	55.2±0.11	0.4188
	ACM_C2		95	>99	76.5±0.04	55.3±0.03	0.4162
Tc. kodakarensis KOD1	TK2278	85	95	>99	80.2±0.12	57.7±0.23	0.4215
P. furiosus DSM 3638	PF1616	100	99	>99	81.6±0.08	69.1±0.15	0.4334
	AAM_A1		99	>99	>99	89.7±0.25	0.4491
	AAM_A2		99	>99	>99	90.6±0.17	0.4465

The T_m values of MIPSs were determined using DSF. The T_{opt} values were determined using a MIPS/malachite green assay. Standard deviations for MIPS T_m values were determined from three replicates. T_{opt} values were determined from at least four replicates. ND, not determined.

Another effort to characterize a novel trait in the Thermotogales was initiated by the discovery that genes encoding a vitamin B12 biosynthetic pathway were acquired by some Thermotogales by horizontal gene transfer (Swithers et al., 2012). We experimentally demonstrated that *Thermotoga lettingae* does not synthesize B12 *de novo*, but salvages cobinamide, as we had predicted based on its gene content. We also found that regulation of expression of these genes is controlled by upstream B12 riboswitches. We identified a putative B12/cobinamide ABC transporter and measured B12 and cobinamide binding by BtuF using DSF. Our results suggest

that Thermotogales species can use the BtuFCD transporters to import both B12 and cobinamide, even if they cannot salvage cobinamide.

Other Thermotogales species possess a variety of catabolic and transport proteins useful for biomass digestion. We characterized new Thermotogales isolates both phenotypically and through genome sequencing to extend our knowledge of the catabolic capacities of this lineage (Nesbo et al., 2012; Swithers et al., 2011A; Swithers et al., 2011B). Both *Kosmotoga olearia* and *Mesotoga prima* were examined, and both grow at mesophile temperatures. *K. olearia* has hydrogenase genes not found in other Thermotogales, so perhaps has capacities for hydrogen production or tolerance that can be utilized in efforts to produce hydrogen from biomass. Genes in *M. prima* indicate that it may play a role in reductive dehalogenation in its natural habitat, a role not previously found in the Thermotogales.

Toga-associated proteins are necessary to allow cells to remain in close association with insoluble growth substrates. We undertook the first proteomic examination of the outer envelope, or "toga," from *T. maritima* in an effort to understand its role in carbohydrate catabolism and cell development (Swithers et al., 2011A). We identified the gene encoding the

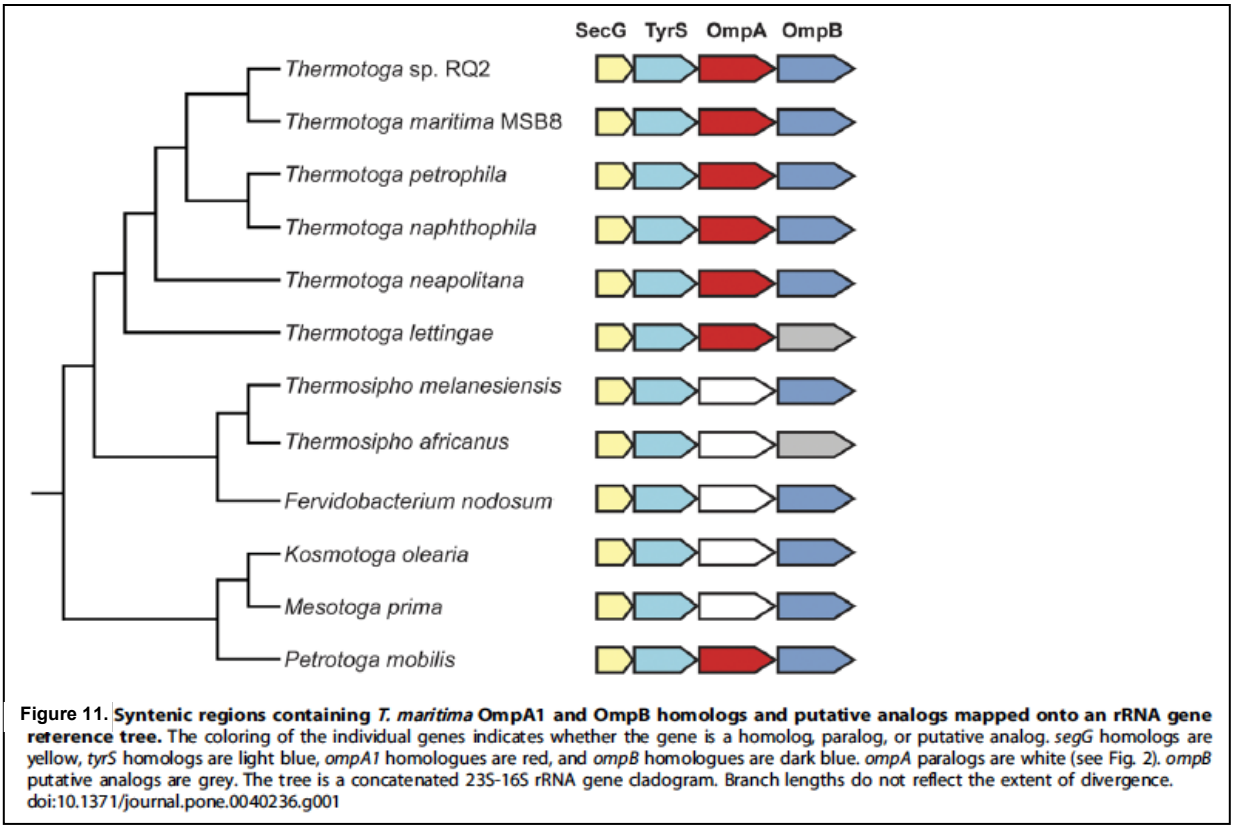


Figure 11. Syntenic regions containing *T. maritima* OmpA1 and OmpB homologs and putative analogs mapped onto an rRNA gene reference tree. The coloring of the individual genes indicates whether the gene is a homolog, paralog, or putative analog. *segG* homologs are yellow, *tyrS* homologs are light blue, *ompA1* homologues are red, and *ompB* homologues are dark blue. *ompA* paralogs are white (see Fig. 2). *ompB* putative analogs are grey. The tree is a concatenated 23S-16S rRNA gene cladogram. Branch lengths do not reflect the extent of divergence.
doi:10.1371/journal.pone.0040236.g001

major outer envelope porin, OmpB, since it had not been identified after the native protein was purified or during genome sequence annotation. The gene TM0476, has all the characteristics reported for OmpB and characteristics expected of a porin. We highly enriched a toga fraction of cells for OmpB and analyzed it by LC/MS/MS in collaboration with the Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory and found that the only porin candidate that it contained was the TM0476 product. A phylogenetic analysis of OmpB found orthologs encoded in syntenic locations in the genomes of all but two Thermotogales species and those without orthologs have putative isofunctional genes in their place. Phylogenetic analyses of the other major envelope protein, OmpA, revealed that each species of the

Thermotogales has one or two OmpA homologs. *T. maritima* has two OmpA homologs, encoded by ompA1 (TM0477) and ompA2 (TM1729), both of which were found in the toga protein-enriched cell extracts (Petrus et al., 2012). We have examined the physiology of the toga further and demonstrated that it distends from the cytoplasmic membrane-enclosed portion of the cells as they enter the late exponential/stationary phase of growth. We find that some of the genes encoding toga-associated proteins are up-regulated during this phase of growth and that the distension is caused by continued growth of the toga and not shrinkage of the cytoplasmic aspect of the cells. We speculate that this increase in cell surface area may have selective value to provide a larger anchor for polysaccharide hydrolytic enzymes during a time of nutritional stress.

Summary. This project led to many interesting insights about the Thermotogales that have both scientific and technological implications (Frock et al., 2010; Frock et al., 2012B). Ongoing work will leverage these developments to further elucidate many interesting features of these novel microorganisms.

PUBLICATIONS ACKNOWLEDGING THIS AWARD

Boucher, N. and K. M. Noll. 2011. Ligand screening by differential scanning fluorimetry of thermophilic transporters encoded in a newly sequenced genomic region of *Thermotoga maritima* MSB8. Appl. Environ. Microbiol. 77:6395-6399.

Butzin, N. C., Lapierre, P., Green, A. G., Swithers, K. S., Gogarten, J. P., and Noll, K. M. Ancestral reconstruction of myo-inositol-3-phosphate synthase, a *Thermotoga* protein of archaeal origin. PLoS One. 2013

Frock, A.D., and R.M. Kelly. 2012B. Extreme thermophiles: Moving beyond single-enzyme biocatalysis. *Curr Opin Chem Eng*. 1:363-372.

Frock, A.D., C.I. Montero, S.E. Blumer-Schuette, and R.M. Kelly. 2013. Stationary phase triggers transcription of a genomic locus containing a novel peptide (TM1316) in the hyperthermophilic bacterium *Thermotoga maritima*. Appl. Environ. Microbiol. 79:66377-6646.

Frock, A.D., J.S. Notey, and R.M. Kelly. 2010. The genus *Thermotoga*: New developments. Environ. Technol. 31:1169-1181.

Frock, A.D., S.A. Gray and R.M. Kelly. 2012A. Hyperthermophilic Thermotoga species differ with respect to specific carbohydrate transporters and glycosidase hydrolases. Appl. Environ. Microbiol. 78:1978-1986.

Nesbø, C. L., D. M. Bradnan, A. Adebusuyi, M. Dlutek, A. K. Petrus, J. Foght, W. F. Doolittle, and K. M. Noll. 2012. *Mesotoga prima* gen. nov., sp. nov., the first described mesophilic species of the Thermotogales. Extremophiles 16:387-393.

Petrus, A. K., K. S. Swithers, C. Ranjit, S. Wu, H. M. Brewer, J. P. Gogarten, L. Pasa-Tolic, and K. M. Noll. 2012. Genes for the major structural components of Thermotogales species' togas revealed by proteomic and evolutionary analyses of OmpA and OmpB homologs. PLoS One.

Rodionova, I. A., S. A. Leyn, M. D. Burkart, N. Boucher, K. M. Noll, A. L. Osterman, D. A. Rodionov. 2013. Novel inositol catabolic pathway in *Thermotoga maritima*. Environ Microbiol. 15:2254-2266.

Swithers, K. S., J. L. DiPippo, D. C. Bruce, C. Detter, R. Tapia, S. Han, E. Saunders, L. A. Goodwin, J. Han, T. Woyke, S. Pitluck, L. Pennacchio, M. Nolan, N. Mikhailova, A. Lykidis, M. L. Land, T. Brettin, K. O. Stetter, K. E. Nelson, J. P. Gogarten, and K. M. Noll. 2011A. Genome sequence of Thermotoga sp. strain RQ2, a hyperthermophilic bacterium isolated from a geothermally heated region of the seafloor near Ribeira Quente, the Azores. J. Bacteriol. 193:5869-5870.

Swithers, K. S., J. L. DiPippo, D. C. Bruce, C. Detter, R. Tapia, S. Han, L. A. Goodwin, J. Han, T. Woyke, S. Pitluck, L. Pennacchio, M. Nolan, N. Mikhailova, M. L. Land, C. Nesbø, J. P. Gogarten, and K. M. Noll. 2011B. Genome sequence of *Kosmotoga olearia* strain TBF 19.5.1, a thermophilic bacterium with a wide growth temperature range, isolated from the Troll B oil platform in the North Sea. J. Bacteriol. 193:5566-5567.

Swithers, K. S., Petrus, A. K., Secinaro, M. A., Nesbø, C. L., Gogarten, J. P., Noll, K. M., Butzin, N. C. Vitamin B12 synthesis and salvage pathways were acquired by horizontal gene transfer to the Thermotogales. 2012. Genome Biol Evol. 4:730-739. doi: 10.1093/gbe/evs057. Epub 2012 Jul 12.

STUDENTS SUPPORTED IN SOME PART BY THIS AWARD:

Frock, Andrew D. (PhD NCSU 2012) Research Engineer, Novozymes (Franklinton, NC)

Notey, Jaspreet S. (PhD NCSU 2013) Process Develop. Scientist, KBI Biopharma (RTP, NC)

Lewis, Derrick L. (PhD NCSU 2010) Research Engineer, Novozymes Biologicals (Salem, VA)

Gray, Steven R. (PhD NCSU 2009) Postdoc, Department of Food, Bioprocessing and Nutrition Sciences, NC State University)

Mukherjee, Arpan (PhD NCSU 2014) Develop. Scientist, Alexion Pharmaceuticals (Chesire, CT)

White, Derrick (PhD UN-L 2015 (expected))

Singh, Raghuveer (PhD UN-L 2015 (expected))