Conners, Shannon Burns. Carbohydrate utilization pathway analysis in the hyperthermophile *Thermotoga maritima*. (Under the direction of Robert Kelly)

Carbohydrate utilization and production pathways identified in *Thermotoga* species likely contribute to their ubiquity in hydrothermal environments. Many carbohydrate-active enzymes from *Thermotoga maritima* have been characterized biochemically; however, sugar uptake systems and regulatory mechanisms that control them have not been well defined. Transcriptional data from cDNA microarrays were examined using mixed effects statistical models to predict candidate sugar substrates for ABC (ATP-binding cassette) transporters in *T. maritima*. Genes encoding proteins previously annotated as oligopeptide/dipeptide ABC transporters responded transcriptionally to various carbohydrates. This finding was consistent with protein sequence comparisons that revealed closer relationships to archaeal sugar transporters than to bacterial peptide transporters. In many cases, glycosyl hydrolases, co-localized with these transporters, also responded to the same sugars. Putative transcriptional repressors of the LacI, XylR, and DeoR families were likely involved in regulating genomic units for β-1,4-glucan, β-1,3-glucan, β-1,4-mannan, ribose, and rhamnose metabolism and transport. Carbohydrate utilization pathways in *T. maritima* may be related to ecological interactions within cell communities. Exopolysaccharide-based biofilms composed primarily of β-linked glucose, with small amounts of mannose and ribose, formed under certain conditions in both pure *T. maritima* cultures and mixed cultures of *T. maritima* and *M. jannaschii*. Further examination of transcriptional differences between biofilm-bound sessile cells and planktonic cells revealed differential expression of β-glucan-specific degradation enzymes, even though maltose, an α-1,4 linked glucose disaccharide, was used as a

growth substrate. Higher transcripts of genes encoding iron and sulfur compound transport, iron-sulfur cluster chaperones, and iron-sulfur cluster proteins suggest altered redox environments in biofilm cells. Further direct comparisons between cellobiose and maltose-grown cells suggested that transcription of cellobiose utilization genes is highly sensitive to the presence of cellobiose, or a cellobiose-maltose mixture. Increased transcripts of genes related to polysulfide reductases in cellobiose-grown cells and biofilm cells suggested that *T. maritima* cells in pure culture biofilms escaped hydrogen inhibition by preferentially reducing sulfur compounds, while cells in mixed culture biofilms form close associations with hydrogen-utilizing methanogens. In addition to probing issues related to the microbial physiology and ecology of *T. maritima*, this work illustrates the strategic use of DNA microarray-based transcriptional analysis for functional genomics studies.

CARBOHYDRATE UTILIZATION PATHWAY ANALYSIS IN THE HYPERTHERMOPHILE *THERMOTOGA MARITIMA*

by

SHANNON B. CONNERS

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

BIOINFORMATICS

Raleigh

2005

BIOGRAPHY

A native of Rome, NY, Shannon Burns Conners graduated summa cum laude from Mount Holyoke College in South Hadley, MA in 1998 with a B.A. in Biochemistry and a minor in Culture, Health and Science. She entered the Bioinformatics program at North Carolina State University in 2000 after transferring from Duke University, where she spent two years as a graduate student in molecular genetics. She joined the Kelly Lab in 2001 and has had the good fortune to participate in a number of interesting projects, including the work collected in this thesis.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1:

CHAPTER 2:

CHAPTER 3:

CHAPTER 4:

LIST OF FIGURES

CHAPTER 1:

CHAPTER 2:

CHAPTER 3:

Introduction to

Carbohydrate utilization pathway analysis in the hyperthermophile *Thermotoga maritima*

Shannon B. Conners

In recent years, advances in sequencing technology have enabled the rapid sequencing of numerous genomes from species in all three domains of life (bacteria, archaea, eukaryotes). Microbial genomes have been a common target because of their relatively small size and compact nature. To date, sequencing projects for ~760 prokaryotic genomes have been undertaken, with 242 completed bacterial and 24 completed archaeal genome sequences now available (http://genomesonline.org). Among these completed genomes are a number from hyperthermophiles, including archaea and bacteria with optimum growth temperatures above 80^{o}C, isolated from high temperature environments around the world. Although the protein sequences of these organisms are related to mesophilic sequences, the high thermostabilty of the encoded proteins has made them attractive targets for characterization.

Thermotoga maritima was originally isolated from a geothermal feature off Vulcano Island, Italy (7). This species is a member of the phylum *Thermotogae*, a group of extremely thermophilic rod-shaped, non-sporulating bacteria with an outer sheath-like envelope referred to as a "toga". The heterotrophic *Thermotogae* ferment a wide variety of sugars to acetic acid, carbon dioxide and hydrogen gas (H_2) gas (7, 20, 21). Early interest in *T. maritima* and other *Thermotoga* species revolved mainly around phylogenetic issues and the wide array of carbohydrate utilization enzymes isolated from these organisms (1-3, 10, 12, 25-27). The genome of *Thermotoga maritima* MSBB was completed in 1999 (14), approximately 13 years after the reported identification of *T. maritima* (7). This genome sequence hinted at massive lateral gene transfer with archaea, confirmed by subsequent work (4, 15, 16, 23)

The genome sequence of MSB8 did not reveal mobile elements in the form of intact transposons or phages which might explain the large numbers of 'archaeal-like' gene sequences in the genome. Although the MSB8 strain had no apparent plasmids, a plasmid isolated from *Thermotoga* sp. RQ7 has been used for transient uptake and expression of genes in *T. maritima* and *T. neapolitana*, though this technique has not been widely applied (17, 24). Despite the lack of easily applied genetic tools, many *T. martima* proteins have been characterized in detail biochemically. However, over the past few years, the tools of functional genomics have provided new opportunities for scientists interested in examining gene regulation and pathways in this organism. In particular, the use of expression-based studies with cDNA microarrays has provided important insights into carbohydrate utilization strategies and ecological interactions of *T. maritima* (5, 6, 8, 18). By coupling sequence information with biochemical data and expression data, potential regulators for carbohydrate utilization pathways and likely binding sites for these regulators have been identified (5, 6). Conservation of important heat shock transcriptional regulatory strategies (e.g., HrcA-mediated CIRCE control) has also been observed, and differences between the response of *T. maritima* and well-characterized mesophiles uncovered (19). Transcriptional responses of *T. maritima* cells in monoculture biofilm communities and in co-culture association with *M. jannaschii* have been examined, revealing concerted networks of response under different conditions (8, 18). Transcriptional data led to the identification of a small orf which encodes a signaling peptide involved in quorum sending in *T. maritima*, indicating for the first time the operation of peptide-based signaling systems in hyperthermophilic bacteria (8).

The expansion of *T. maritima* expression studies to examine multiple treatments has required the use of experimental design and analysis methods suited to large datasets. Often, cost must be balanced with replication in an experiment, and the use of loop designs has proven an efficient way to compare multiple treatments while minimizing the number of arrays required. Microarray analysis techniques such as Analysis of Variance (ANOVA) are well-suited to analyzing such data, and allow the construction of multiple pair-wise comparisons between treatments within an experiment (9, 22). Visualization of results by genomic order and hierarchical clustering methods offer different benefits when analyzing microarray data from prokaryotes, suggesting potential operon-level and regulon-level groupings of genes, respectively. Genes within these groupings can then be examined for commonalities in promoter structure, which may indicate candidate binding sites and potential transcriptional regulatory mechanisms.

Tools of functional genomics have also been applied to questions relating to genomic evolution, as microarrays have been utilized for comparative genomics purposes to highlight the remarkable diversity present within the *Thermotoga* genus (13). The completion of the *T. neapolitana* genome (Nelson, unpublished observation) provides an important opportunity for whole-genome comparison within the *Thermotoga* genus, and perhaps eventually comparison of expression responses between the two species. A structural genomics initiative targeting *T. maritima* proteins provides an additional functional genomics resource for this and related species (11). Clearly, the integration of expression information gained from microarray experiments with structural and comparative genomics data will provide guidance for future experiments addressing the potential functional roles of hypothetical proteins in *T. maritima*.

INTRODUCTION REFERENCES

1. **Bandlish, R. K., J. Michael Hess, K. L. Epting, C. Vieille, and R. M. Kelly.** 2002. Glucose-to-fructose conversion at high temperatures with xylose (glucose) isomerases from *Streptomyces murinus* and two hyperthermophilic *Thermotoga* species. Biotechnol Bioeng **80:**185-94.
2. **Bibel, M., C. Brettl, U. Gosslar, G. Kriegshauser, and W. Liebl.** 1998. Isolation and analysis of genes for amylolytic enzymes of the hyperthermophilic bacterium *Thermotoga maritima*. FEMS Microbiol Lett **158:**9-15.
3. **Bronnenmeier, K., A. Kern, W. Liebl, and W. L. Staudenbauer.** 1995. Purification of *Thermotoga maritima* enzymes for the degradation of cellulosic materials. Appl Environ Microbiol **61:**1399-407.
4. **Calteau, A., M. Gouy, and G. Perriere.** 2005. Horizontal transfer of two operons coding for hydrogenases between bacteria and archaea. J Mol Evol **60:**557-65.
5. **Chhabra, S. R., K. R. Shockley, S. B. Conners, K. L. Scott, R. D. Wolfinger, and R. M. Kelly.** 2003. Carbohydrate-induced differential gene expression patterns in the hyperthermophilic bacterium *Thermotoga maritima*. J Biol Chem **278:**7540-52.
6. **Conners, S. B., C. I. Montero, D. A. Comfort, K. R. Shockley, M. R. Johnson, S. R. Chhabra, and R. M. Kelly.** 2005. An expression-driven approach to the prediction of carbohydrate transport and utilization regulons in the hyperthermophilic bacterium *Thermotoga maritima*. J Bacteriol **187:**7267-82.

7. **Huber, R., T. A. Langworthy, H. Konig, M. Thomm, C. R. Woese, U. B. Sleytr, and K. O. Stetter.** 1986. *Thermotoga maritima* sp. nov. represents a new genus of unique extremely thermophilic eubacteria growing up to 90 DegreesC. Archives of Microbiology **144:**324-333.

8. **Johnson, M. R., C. I. Montero, S. B. Conners, K. R. Shockley, S. L. Bridger, and R. M. Kelly.** 2005. Population density-dependent regulation of exopolysaccharide formation in the hyperthermophilic bacterium *Thermotoga maritima*. Mol Microbiol **55:**664-74.

9. **Kerr, M. K., M. Martin, and G. A. Churchill.** 2000. Analysis of variance for gene expression microarray data. J Comput Biol **7:**819-37.

10. **Kluskens, L. D., G. J. van Alebeek, A. G. Voragen, W. M. de Vos, and J. van der Oost.** 2003. Molecular and biochemical characterization of the thermoactive family 1 pectate lyase from the hyperthermophilic bacterium *Thermotoga maritima*. Biochem J **370:**651-9.

11. **Lesley, S. A., P. Kuhn, A. Godzik, A. M. Deacon, I. Mathews, A. Kreusch, G. Spraggon, H. E. Klock, D. McMullan, T. Shin, J. Vincent, A. Robb, L. S. Brinen, M. D. Miller, T. M. McPhillips, M. A. Miller, D. Scheibe, J. M. Canaves, C. Guda, L. Jaroszewski, T. L. Selby, M. A. Elsliger, J. Wooley, S. S. Taylor, K. O. Hodgson, I. A. Wilson, P. G. Schultz, and R. C. Stevens.** 2002. Structural genomics of the *Thermotoga maritima* proteome implemented in a high-throughput structure determination pipeline. Proc Natl Acad Sci U S A **99:**11664-9.

12. **Liebl, W., P. Ruile, K. Bronnenmeier, K. Riedel, F. Lottspeich, and I. Greif.** 1996. Analysis of a *Thermotoga maritima* DNA fragment encoding two similar thermostable cellulases, CelA and CelB, and characterization of the recombinant enzymes. Microbiology **142 (Pt 9):**2533-42.

13. **Mongodin, E. F., I. R. Hance, R. T. Deboy, S. R. Gill, S. Daugherty, R. Huber, C. M. Fraser, K. Stetter, and K. E. Nelson.** 2005. Gene transfer and genome plasticity in *Thermotoga maritima*, a model hyperthermophilic species. J Bacteriol **187:**4935-44.

14. **Nelson, K. E., R. A. Clayton, S. R. Gill, M. L. Gwinn, R. J. Dodson, D. H. Haft, E. K. Hickey, J. D. Peterson, W. C. Nelson, K. A. Ketchum, L. McDonald, T. R. Utterback, J. A. Malek, K. D. Linher, M. M. Garrett, A. M. Stewart, M. D. Cotton, M. S. Pratt, C. A. Phillips, D. Richardson, J. Heidelberg, G. G. Sutton, R. D. Fleischmann, J. A. Eisen, C. M. Fraser, and et al.** 1999. Evidence for lateral gene transfer between Archaea and bacteria from genome sequence of *Thermotoga maritima*. Nature **399:**323-9.

15. **Nesbo, C. L., S. L'Haridon, K. O. Stetter, and W. F. Doolittle.** 2001. Phylogenetic analyses of two "archaeal" genes in *Thermotoga maritima* reveal multiple transfers between archaea and bacteria. Mol Biol Evol **18:**362-75.

16. **Nesbo, C. L., K. E. Nelson, and W. F. Doolittle.** 2002. Suppressive subtractive hybridization detects extensive genomic diversity in *Thermotoga maritima*. J Bacteriol **184:**4475-88.

17. **Noll, K. M., and M. Vargas.** 1997. Recent advances in genetic analyses of hyperthermophilic archaea and bacteria. Arch Microbiol **168:**73-80.

18. **Pysz, M. A., S. B. Conners, C. I. Montero, K. R. Shockley, M. R. Johnson, D. E. Ward, and R. M. Kelly.** 2004. Transcriptional analysis of biofilm formation processes in the anaerobic, hyperthermophilic bacterium *Thermotoga maritima*. Appl Environ Microbiol **70:**6098-112.

19. **Pysz, M. A., D. E. Ward, K. R. Shockley, C. I. Montero, S. B. Conners, M. R. Johnson, and R. M. Kelly.** 2004. Transcriptional analysis of dynamic heat-shock response by the hyperthermophilic bacterium *Thermotoga maritima*. Extremophiles **8:**209-17.

20. **Van Ooteghem, S. A., S. K. Beer, and P. C. Yue.** 2002. Hydrogen production by the thermophilic bacterium *Thermotoga neapolitana*. Appl Biochem Biotechnol **98-100:**177-89.

21. **Van Ooteghem, S. A., A. Jones, D. Van Der Lelie, B. Dong, and D. Mahajan.** 2004. H(2) production and carbon utilization by *Thermotoga neapolitana* under anaerobic and microaerobic growth conditions. Biotechnol Lett **26:**1223-32.

22. **Wolfinger, R. D., G. Gibson, E. D. Wolfinger, L. Bennett, H. Hamadeh, P. Bushel, C. Afshari, and R. S. Paules.** 2001. Assessing gene significance from cDNA microarray expression data via mixed models. Journal of Computational Biology **8:**625-637.

23. **Worning, P., L. J. Jensen, K. E. Nelson, S. Brunak, and D. W. Ussery.** 2000. Structural analysis of DNA sequence: evidence for lateral gene transfer in *Thermotoga maritima*. Nucleic Acids Res **28:**706-9.

24. **Yu, J. S., M. Vargas, C. Mityas, and K. M. Noll.** 2001. Liposome-mediated DNA uptake and transient expression in *Thermotoga*. Extremophiles **5:**53-60.

25. **Zverlov, V., K. Piotukh, O. Dakhova, G. Velikodvorskaya, and R. Borriss.** 1996. The multidomain xylanase A of the hyperthermophilic bacterium *Thermotoga neapolitana* is extremely thermoresistant. Appl Microbiol Biotechnol **45:**245-7.

26. **Zverlov, V. V., I. Y. Volkov, T. V. Velikodvorskaya, and W. H. Schwarz.** 1997. Highly thermostable endo-1,3-beta-glucanase (laminarinase) LamA from *Thermotoga neapolitana*: nucleotide sequence of the gene and characterization of the recombinant gene product. Microbiology **143 (Pt 5):**1701-8.

27. **Zverlov, V. V., I. Y. Volkov, T. V. Velikodvorskaya, and W. H. Schwarz.** 1997. *Thermotoga neapolitana bglB* gene, upstream of *lamA*, encodes a highly thermostable beta-glucosidase that is a laminaribiase. Microbiology **143 (Pt 11):**3537-42.

Chapter 1:

Microbial Biochemistry, Physiology, Ecology and Biotechnology of Hyperthermophilic *Thermotoga* Species

Shannon B. Conners[1], Emmanuel F. Mongodin[2], Matthew R. Johnson[1], Clemente I. Montero[1], Karen E. Nelson[2] and Robert M. Kelly[1]

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

[2] The Institute for Genomic Research
Rockville, MD 20850

OUTLINE

1. Introduction

The members of the phylum *Thermotogae* are a group of extremely thermophilic rod-shaped, non-sporeforming bacteria with an outer sheath-like envelope also known as a "toga". They demonstrate heterotrophic growth, with acetic acid, carbon dioxide and hydrogen gas (H_2) gas as the main products from fermentation (56, 224, 225). Along with the *Aquificales*, these organisms represent the bacteria with the highest growth temperatures (from 65 to 90°C). While the hyperthermophilic bacterium *Thermotoga maritima* was originally isolated from a geothermal feature off Vulcano Island, Italy (56), *Thermotoga* species have been found in diverse locations around the globe (4, 39, 56, 63, 67, 164, 218). Among the *Thermotogales* that are currently described are 9 members of the genus *Thermotoga*, four species of *Fervidobacterium*, eight members of the genus *Thermosipho*, four species of *Geotoga*, four species of *Petrotoga*, two species of *Marinitoga*, and one species of *Thermopallium* (www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?mode=Undef&id=2419&lvl=3&lin=f&keep=1&srchmode=1&unlock). The ubiquity of these species is likely the result of genomic and metabolic versatility, which includes an extraordinary array of enzymes involved in diverse carbohydrate utilization pathways in some genus members (5, 10, 15, 21, 27, 74, 78, 136, 215, 216, 253, 254). There has been substantial debate over the phylogenetic placement of the *Thermotogales* (for a recent review of phylogenetics of the *Thermotogales* as well as details of physical characteristics of this genus, see (57)), which is complicated by substantial apparent lateral gene transfer (LGT) events between *Thermotoga* and the archaeal species living in the same environment. The publication of the MSB8 genome sequence significantly raised scientific interest in *Thermotoga* species,

and provided the reference data needed for high-throughput genomics. The approach taken to the characterization of the *T. maritima* MSB8 genome (functional genomics, protein characterization and structure determination) illustrates a new paradigm for microbial genome characterization. Such approaches will be increasingly important as the sequencing of diverse microbial genomes overtakes the development of species-specific genetic systems.

One recent estimate places the number of partially sequenced microbial genomes at over 760, while more than 260 have been finished (source: www.GenomesOnline.org). As metagenomics of uncultured microbes provides additional sequence information, some of the tools and approaches developed to explore the *T. maritima* genome will undoubtedly be applied to other species. In particular, recent applications of microarrays and high-throughput structural genomics to *T. maritima* will be highlighted in this review, including microarray experiments aimed at characterizing the transcriptional response of *T. maritima* to heat shock, carbon source variation, growth in cell communities, and co-culture with *Methanococcus jannaschii* (Table 1.1). The results of comparative genomics studies which have revealed high levels of variability between *Thermotoga* species will also be discussed. Novel findings from structural genomics efforts based on high-throughput crystallization and characterization of *T. maritima* proteins will also be described.

2. Genomics of *Thermotoga* species

2.1. Insights from the genome sequence of the *Thermotoga* model strain : *T. maritima* MSB8

The genome of the hyperthermophilic bacterium *Thermotoga maritima* strain MSB8 was sequenced in 1999 (132). At the time, the 1,860,725-base-pair circular chromosome (G+C content: 46%) was predicted to contain 1,877 coding regions, 1,014 (54%) of which were assigned functional assignments and 863 (46%) of which were of unknown function (132). All but 5% of the 1.86 Mb genome was predicted to be covered by open reading frames (ORFs). The motivations to sequence the genome of *T. maritima* MSB8 were multiple. First, this organism has an optimum growth temperature of 80°C and is able to metabolize many simple and complex carbohydrates including glucose, sucrose, starch, cellulose and xylan (56). Both cellulose and xylan, through conversion to fuels (such as hydrogen), have great potential as renewable carbon and energy sources (224). From the genome sequence, approximately 7% of the *T. maritima* genes were predicted to be involved in carbohydrate utilization, breakdown and metabolism, consistent with observed growth of *T. maritima* on various sugars (132).

The *T. maritima* genome sequencing project was therefore carried out to elucidate further the unique metabolic properties of this organisms, as well as the evolutionary relationship with other microbial species, and address potential industrial applications that could derived from its predicted protein sequences. Comparative genome analysis of the chromosome with the genomes of other available completely sequenced microbial species presented evidence for LGT, with approximately 24% of the *T. maritima* ORFs having their best matches to genes from archaeal species. Many of these genes were also

of atypical composition and were often found in clusters (termed “archaeal islands”): 81 archaeal-like genes were found to be clustered in 15 chromosomal regions, with a size ranging from 4 to 20 kbps. Conservation of gene order between *T. maritima* and Archaea in many of the clustered regions suggested that LGT may have occurred between *T. maritima* and archaeal species. Possibly due to the extensive LGT with archaeal species, the MSB8 genome sequence did not unequivocally clarify the phylogenetic position of *Thermotoga* species. Examination of a subset of 33 genes conserved among all known microbial species could not verify the phylogenetic position of *T. maritima* as proposed by 16S rDNA phylogeny. Additional support for massive LGT between *T. maritima* and archaea was found when the periodicity of the genome was examined relative to other bacteria and archaea (239).

The *T. maritima* genome sequence did not reveal any obvious mechanisms for LGT, such as mobile elements in the form of intact transposons or phages, that may have contributed to the observed levels of 'archaeal-like' gene sequences in the genome. Competence has not been demonstrated in *T. maritima*, but various type II secretion pathway proteins and type IV pilin-related proteins that function in natural competence in other bacterial species could be identified from an analysis of the whole genome (132). In addition, homologs of various competence genes could also be identified including *dprA*, *comM*, and *comE* of *Haemophilus influenzae*, suggesting that there may be an inherent (but yet to be identified) system for the uptake of exogenous DNA and thereby facilitating the exchange of DNA with other organisms. A direct consequence of this is the lack of effective tools, such as vectors and knock-out systems, needed to perform molecular biology experiments with *Thermotoga*. A plasmid vector has been developed

for *Thermotoga* based on a cryptic miniplasmid fused with the *E. coli* plasmid pBluescript and the heat-stable *Staphylococcus aureus* chloramphenicol acetyltransferase (52, 247). Vargas and Noll (138) have developed anti-metabolite resistant and auxotrophic mutants of *T. neopolitana*, as well as a defined minimal medium that will support growth of this organism.

The original *T. martima* genome annotation was re-evaluated by Kyripides et al. (84) and revealed agreement in 895 cases, 29 discordant annotations, and 863 remaining hypothetical proteins. Examination of the conflicting annotation predictions using tools such as the PSI-BLAST algorithm (www.ncbi.nlm.nih.gov/BLAST/tutorial/Altschul-2.html), presence of PRO-SITE patterns (http://au.expasy.org/prosite/), matches to the protein family databases Pfam (www.sanger.ac.uk/Software/Pfam/) and COGs (http://www.ncbi.nlm.nih.gov/COG/), revealed 163 new functional assignments and 29 amendments to functions predicted by Nelson and colleagues (132). Of particular interest are the prediction of a rhamnose-rhamnulose utilization cluster (TM1071-TM1073) and confirmation of additional homologs for large protein families including ABC transporters, as well as identification of proteins with CBS, transmembrane, and zinc-finger domains. Kyripides and colleagues also identified several cases of fusion proteins, e.g. the annotated chorismate mutase from the tyrosine and phenylalanine biosynthesis pathway is a fusion protein consisting of both chorismate mutase and phospho-2-dehydro-3-deoxyheptonate-aldolase domains. In addition to bioinformatics re-analysis, biochemical studies of a number of enzymes from *T. maritima* have been undertaken since the publication of the MSB8 genome sequence, clarifying or confirming predicted functions (Table 1.2).

2.2. Genetic and phylogenetic studies in the *Thermotoga* lineage

LGT in the *Thermotoga* lineage was also investigated in a series of studies by Nesbo and colleagues. In the first of these studies, the patterns of acquisition of two predicted "archaeal-like" genes, glutamate synthase large subunit (*gltB*, TM0397) and myo-inositol-1-P synthase (*inol1*, TM1419) were investigated (134). Amplification of *inol1* and *gltB* was attempted in 15 species of the *Thermotogales* genus, but was only successful for species within the *Thermotoga* lineage. Phylogenetic analysis of the *Thermotoga* homologs with those of other species suggested the possibility of several independent transfers of the archaeal *gltB* genes into bacterial species. The presence in *Thermotoga* of three ORFs with regions of similarity to bacterial *gltB*, a feature characteristic of Archaea, supports this hypothesis. The gene *inol1*, which is found only in hyperthermophiles, is known to be used in the production of the osmolyte and supposed thermoprotectant DIP (di-myo-inositol-1,1-P). The pattern of amplification of *inol1* along the species of the *Thermotogales* examined was consistent with the distribution of DIP among *Thermotoga* species as previously determined (113). Here, intraspecific intragenic recombination was suggested to explain the mosaic nature of the *inol1* homolog of *Thermotoga sp.* strain RQ2. Phylogenetic analyses also suggested that all *inol1* homologs present in bacteria likely resulted from LGT with archaea.

In a broader study, Nesbo and workers used suppressive subtractive hybridization (SSH) to compare the gene content of *T. maritima* MSB8 to that of *Thermotoga* sp. strain RQ2, which was isolated from the geothermally heated sea floor in Ribiera Quente, the Azores, and whose 16S rDNA sequence shares high identity (~99.7%) with that of strain

MSB8 (135). Over 300 unique RQ2-specific sequences were obtained from clones generated by SSH, and it was estimated that 20% of the RQ2 genome was not present in the genome of strain MSB8. Genes with highest similarity to archaeal genes as well as genes most closely related to homologs in distant bacterial genomes were found, suggesting LGT with species from both domains of life. Consistent with this hypothesis, RQ2-specific genes appeared to be found in gene clusters missing from strain MSB8. Ratios of synonymous to non-synonymous mutations in the divergent genes did not suggest positive selection but rather LGT followed by random mutations. Homology searches among RQ2-specific clones revealed numerous genes related to carbohydrate processing and ABC transport. These included carbohydrate active enzymes (e.g. arabinosidases, xylosidases) as well as ABC transporter subunits. Numerous clones encoding sugar ABC transporter subunits identified in strain RQ2 by SSH were found by subsequent sequencing to be related to, but divergent from, homologs in strain MSB8 (<85% identity). It was suggested that the expansion of sugar transporters in strain RQ2 may represent lineage-specific gene expansion possibly used as a strategy for environmental adaptation. Variation in genes related to surface structure formation was observed between the two species, including the acquisition of a set of rhamnose biosynthesis proteins by RQ2. Using Southern blots, the authors proposed that other *Thermotoga* strains shared scattered RQ2-specific sequences in patterns again suggestive of LGT.

Using RQ2-specific probe sequences derived from the SSH studies described above, Nesbo and co-workers (133) screened lambda libraries that were created from strain RQ2 DNA for five regions that are absent from the MSB8 genome. Among the

gene clusters found to be unique to strain RQ2 were an archaeal-type ATPase, the rhamnose biosynthesis operon mentioned above, and an arabinosidase island. Again, both the phylogenetic patterns and G+C content suggested the acquisition of these RQ2 specific genes through LGT.

2.3. Genetic diversity among the *Thermotogales*

Six years after the completion of the strain MSB8 genome, additional sequences from *Thermotoga* species have been made available in Genbank, including a whole-genome shotgun draft sequence from *T. neapolitana* strain NS-E (Genbank accession NC_006811; Nelson et al. - Manuscript in preparation), and selected sequences from *Thermotoga napthophila*, *Thermotoga petrophila*, and *Thermotoga* species KOL6, RQ2, RQ7, SL7, and *Thermotoga maritima* FjSS3-B.1. Many of the deposited sequences from other *Thermotoga* species are homologous to MSB8 genes, including enzymes active against various carbohydrates (176, 181, 205). Plasmid sequences from *T. petrophila* and sequences generated by SSH studies with *Thermotoga* strains MSB8 and RQ2 (135) are among those sequences not shared with the MSB8 genome.

Using comparative genomic hybridization (CGH), Mongodin and colleagues (127) examined the genetic differences of nine *Thermotoga* species, including strain RQ2, in comparison to the sequenced *T. maritima* strain MSB8 (Figure 1.1). Patterns of shared and species-specific sequences between the different species were not restricted by 16S rDNA sequence relationships. Subsequent DNA sequencing of the RQ2-specific islands was performed, in order to differentiate between divergent sequences and absence of genes in one strain compared to the reference, both of which events could not be

distinguished by CGH using the MSB8 array. Apparent large acquisitions by strain MSB8, which are not found in strain RQ2, included a rhamnose utilization locus (TM1063-TM1071) (see also (27)), putative sugar utilization genes (TM0411-TM0423), and a region containing phosphate transport genes (TM1261-TM1271). Sequencing of other regions which displayed poor hybridization with *T. maritima* probes revealed recombination events within genes. In some cases, these resulted in homologous genes retaining conserved N- and C- termini separated by strain-specific intervening protein sequences. Additionally, some genes which were not detected by CGH were indeed present but displayed high divergence compared to probe sequences.

Many of the genes found to differ in their presence and absence patterns in the different *Thermotoga* species from the CGH study of Mongodin and colleagues (127) seem to be related to substrate utilization, and reflect the importance of LGT events in enabling a quick evolution and adaptation to specific habitats. For example, a large number of MSB8 carbohydrate utilization loci were found to be lacking in *Thermotoga* species PB1platt, isolated from an oil field where plant polymers are scarce. Multiple paralogous sugar transport systems and carbohydrate active enzymes have been identified in RQ2 by SSH with MSB8 (133). As a result of these findings, it is clear that the details of carbohydrate utilization patterns in MSB8 that are gleaned from functional studies cannot be automatically extended to all *Thermotoga* species and strains. However, these initial studies have been beneficial in elucidating the pathways by which *T. maritima* MSB8 hydrolyzes, transports, and utilizes a variety of substrates.

2.4. CRISPR elements in the *Thermotogales*

Analysis of the *T. maritima* genome showed the presence of Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) in eight distinct loci on the chromosome. CRISPR elements have a remarkable structure that consists of a 30-bp repeat element interspersed with a variable and non-repetitive 39- to 40-bp sequence called the "spacer." They are thought to increase in size by duplicating the repeat sequences and adding at least one new spacer by a mechanism that is still not known; the origin of the variable spacer sequences also remains elusive (125). These CRISPR elements have been identified in a broad range of microbial species, including *Salmonella enterica* serovar *Typhimurium*, *Streptococcus pyogenes*, *Mycobacterium tuberculosis* and *Campylobacter jejuni* (64, 126, 153, 185). The unique structure of CRISPRs and their association with a group of conserved genes (called *cas* genes, for CRISPR-associated sequences), which are potentially involved in DNA recombination and repair (111), provide additional clues for an active role of CRISPR elements in the mobilization of DNA, and could play an active role in LGT in *T. maritima*. In a recent study by Mongodin and colleagues, CRISPRs regions were found to vary in length and in number of repeats, between different *Thermotoga* species, in many cases encompassing large sets of unique gene sequences. CRISPR repeats have also been found flanking inversion sites detected by comparison of the whole genome sequences of *T. maritima* MSB8 and *T. neapolitana* NS-E (Nelson et al, manuscript in preparation; DeBoy et al., manuscript in preparation).

3. Insights into *Thermotoga* physiology revealed by biochemical and functional genomic efforts

3.1. Carbohydrate utilization pathways of *Thermotoga* species

Important functional information about sugar utilization pathways used by *T. maritima* has been gained from both biochemical characterization of carbohydrate active enzymes and expression-based work. By combining data from these studies with comparative genomics analyses, many details of *T. maritima* carbohydrate processing pathways can be predicted. However, there is still much to learn about the local and global mechanisms of regulation of these pathways. The functions and specificities of members of the LacI and XylR families of carbohydrate-responsive regulators in *T. maritima* have yet to be biochemically determined. No mechanism of carbon catabolite repression (CCR) has yet been defined for *T. maritima*, although catabolite repression of lactose utilization genes has been demonstrated in the presence of glucose in *T. neapolitana* (226). However, glucose has not been shown to be a preferred substrate for any *Thermotoga* species, perhaps due to its thermolability at high temperatures. In fact, growth of *T. maritima* on glucose has been observed to be slower than growth on other polysaccharides (21). Neither the general (HPr, EI) nor sugar-specific (EII) components of a PTS system are apparent in *T. maritima* or *neapolitana* species (43, 132). Rather, uptake of most carbohydrates, and many other substances, is likely accomplished via binding protein dependent ABC transporters (129, 132) (Table 1.3). Multiple homologs from known families of bacterial sugar transporters (e.g. CUT1 and CUT2) (184) have been identified in *Thermotoga* species (132, 135). A number of oligopeptide/dipeptide

family ABC transporters found in *T. maritima* are closely related to archaeal homologs shown to act as sugar transporters (34, 80). If this is also the case for *T. maritima*, it could explain the proximity of many of these genes to glycoside hydrolases and members of carbohydrate-responsive regulator families.

In the absence of a PTS system, *T. maritima* cells must perform specific hydrolysis, phosphorylation, and/or isomerization steps to use each imported sugar. Free phosphate (e.g., cellobiose phosphorylase) or polyphosphate (e.g., PPi-dependent PFK) can be used to generate phosphorylated sugars in some cases. In other cases ATP is required; e.g., glucose is phosphorylated via ATP hydrolysis by an ATP-dependent glucokinase (51). Since ATP is also used by different ABC transporters to import substrates of varying chain lengths, the number of sugar molecules acquired per molecule of ATP hydrolyzed may also vary by substrate. It remains to be seen what effects these differing ATP requirement have on the metabolism of *T. maritima* cells.

3.1.1. Alpha-linked glucans oligo- and poly- saccharides

The most exhaustively studied pathways for polysaccharide utilization in *T. maritima* involve the hydrolysis of α-linked polysaccharides and disaccharides. Genomic and biochemical data suggest that *T. maritima* can use the α-1,4 linked glucose disaccharide maltose as well as pullulan and starch, which contain mixed α-1,4 and α-1,6 linkages (10, 21). Reconstruction of the pathways, and biochemical characterization of the enzymes used by *T. maritima* MSB8 to break down α-linked polysaccharides in some cases preceded the sequencing of its genome (10, 102, 187). Using data from these

functional studies, pathways for the uptake and utilization of various α-linked glucans sugars can be inferred (Figure 1.2).

In the cases of both pullulan and starch, extracellular enzymes perform the reactions necessary to break down the complex polysaccharides for transport into the cell. Debranching of α-1,6 linkages in pullulan and starch is likely accomplished by the type I pullulanase Pul13A (TM1845) (82), which yields mainly maltotriose as a product (10). Recent studies of this enzyme characterized a novel N-terminal domain (TmCBM41) which binds α-1,4 linked glucans and α-1,4 glucans with occasional α-1,6 linkages (85). Amy13A (TM1840), a Ca^{2+}-requiring membrane-bound α-amylase, hydrolyzes extracellular starch, amylase and amylopectin but has lowered activity towards the more highly branched polysaccharides glycogen and pullulan (102, 187).

Binding of hydrolyzed products of α-glucans and maltose for transport into *T. maritima* MSB8 cells apparently involves at least two different maltose ABC binding proteins. The maltose binding protein encoded by TM1839 (TMMBP, *malE2*) has been characterized both functionally and structurally, and binds maltose, maltotriose, and trehalose (129, 130, 232). Transcripts of TM1839 were observed to be more than 3 fold higher during growth on maltose than glucose (136). TM1839 co-localizes in the genome with two ABC permease subunits (TM1836, TM1838), the family 4 *a*-glucosidase AglA (TM1834) (162), and a cyclomaltodextrinase (TMG, TM1835) (94). A gene string divergently transcribed from the maltose and trehalose transport locus encodes the lipoprotein α-amylase Amy13A, pullulanase Pul13A, and hypothetical proteins which have sequence similarity to putative α-glucan processing enzymes (132). The *T.*

neapolitana locus is organized similarly, but several additional inserted genes are present, further underscoring the genome plasticity observed in *Thermotoga* species (8).

The high identity between TMMBP and the protein encoded by TM1204 (79% id/391 aa) may suggest possible recent duplication in the *Thermotoga* lineage (130). While the second MBP homolog (TM1204, *malE1*) was not differentially expressed during growth on maltose as compared to glucose, large differences in transcript levels during growth on lactose as compared to glucose were observed (>5 fold) (136). In addition to binding maltose and maltotriose, TM1204 has been shown to bind β-1,4-mannotetraose, suggesting divergence of the two maltose transporters to fulfill different transport capabilities (130). Consistent with previous findings, we have observed no differential expression of the TM1202-1204 transporter genes during growth on starch (Conners and Kelly, unpublished data). The proximity of this transporter to lactose/galactose utilization genes (TM1190-1199) (75, 103), a LacI family regulator (TM1200) and putative arabinogalactan endo-1,4-β-galactosidase (TM1201) suggests a role in the uptake of mixed sugar oligosaccharides.

Both sets of characterized maltose transport proteins from *T. maritima* lack ATP-binding subunits, as do several other sets of CUT1 transporters found in the genome. We have observed increased transcription of a MalK ATP-binding subunit (TM1276) during growth on maltose, and presume it may interact with both maltose transporters and perhaps other related sets of permeases and substrate binding proteins. Interaction of the same ATP-binding subunit with separate ABC permeases and substrate binding proteins for maltose and cellobiose has been observed in *Streptomyces olivaceoviridis* (183).

Following transport into the cell, maltose hydrolysis in *T. maritima* cells likely involves the α-glucosidase AglA (TM1834), which is active on maltose and maltotriose but not starch, amylopectin or amylose (10). AglA requires reducing conditions, NAD+, and Mn2+ for activity (109, 161, 162). Two other family 4 glycoside hydrolases initially also annotated as α-glucosidases have been shown to be α-glucuronidases (215, 216). Longer maltooligosaccharides may be hydrolyzed by a CDase (TM1835) active on maltooligosaccharides of 3-7 glucose units and cyclodextrins, which yields mainly glucose and maltose as products (91).

An additional pathway for the utilization of maltodextrins may exist in some *Thermotoga* species, although its operation in MSB8 has not been confirmed. This alternative pathway might use 4-α-glucanotransferase and maltodextrin phosphorylase activities to hydrolyze maltose or maltooligosaccharides. *E. coli* 4-α-glucanotransferase (MalQ) can not use maltose as a sole substrate, but uses maltose and glucose as acceptors for endogenously produced maltotriose and higher maltodextrins (155). Maltodextrin phosphorylase (MalP) then hydrolyzes the resulting maltodextrins to glucose-1-phosphate using free phosphate (140). Two intracellular *T. maritima* proteins with 4-α-glucanotransferase activity have been characterized. The 4-α-glucanotransferase activity identified by Liebl and colleagues (99) can be attributed to TM0364 (MgtA). This enzyme transfers α-1,4 linked glucanosyl segments from starch, amylose, amylopectin, and maltooligosaccharides M3 and longer, but can not use maltose and maltotriose as sole substrates. TM0364 has been characterized structurally (173, 174). The family 13 maltosyltransferase (MTase; mmtA, TM0767) characterized by Meissner and Liebl (118) transfers only maltose units from α-1,4 glucans to maltotriose or higher order

maltooligosaccharides. A double-displacement mechanism of maltose unit transfer and a structural basis for the maltose specificity of the enzyme have since been inferred from studies of the crystal structure of this protein in complex with maltose (19, 172). Although a functional maltodextrin phosphorylase (AgpA) was isolated from *T. maritima* MSB8 cells (10), the sequence of the gene encoding this protein (TM1168) was found to be frameshifted in the MSB8 genome (132). A homologous sequence has been found in *Thermotoga* sp. RQ2 (127) (GI:69954028). Further work is needed to determine the biological significance of these apparent inter-strain differences. Although synthesis of intracellular storage polysaccharides has not been demonstrated experimentally for *T. maritima*, other hyperthermophiles have been shown to synthesize such polymers (e.g., *Thermococcus hydrothermalis* (48). The hydrolysis of intracellular α-1,4 glucans could be accomplished by Amy13B, an intracellular alpha-amylase active on starch (104, 187).

Questions remain about the mechanism of regulation of starch and maltose utilization genes in *T. maritima,* as no maltose regulator has yet been characterized in this species. Transcriptional regulators specific for maltose have been identified in many organisms, including LacI/GalR family repressor proteins in various bacteria (2, 46, 86, 137, 165), the activator MalT in *E. coli* (24) a Crp homolog in *Bacteroides thetaiotaomicron* and TrmB in the archaeon *Thermococcus litoralis* (95). While the gene encoding the LacI family protein TM1200 lies downstream of the maltose/β-mannotetraose transporter (130), it is differentially expressed on lactose as compared to glucose but not during growth on maltose (136). Clearly, further work beyond expression studies will be needed to determine the mechanism of regulation of maltose utilization genes.

3.1.2. Beta-linked oligo- and poly- saccharides

3.1.2.1.β-1,4-Glucan

Processing of β-1,4 linked glucans (e.g. carboxymethylcellulose, β-1,4 linked degradation products from barley glucan, cellobiose) by *T. maritima* cells appears to involve hydrolysis of oligosaccharides by extracellular β-glucosidases, transport of small chains into the cell via an Opp family transporter, and a phosphorylation and cleavage step mediated by a cellobiose phosphorylase. Cellulase I and cellulase II were first isolated from *T. maritima* by Bronnenmeir and colleagues (15). This work and later work by Liebl and colleagues confirmed that cellulase I (Cel12A, TM1524) is an endo-beta-1,4-glucanase, while cellulase II (Cel12B, TM1525), located downstream of Cel12A, has a signal peptide and is an exo-beta-1,4 glucanase (101).

T. maritima cells likely import β-glucan degradation products via a putative cello-oligosaccharide transporter related to the characterized *Pyrococcus furiosus* cellobiose transporter (80). The binding protein of the putative transporter (TM1223, crystal structure PDB:1VR5) shares ~60% identity with *P. furiosus* CbtA, which binds cellobiose through cellopentaose as well as β-1,3 linked glucan di- and tri-saccharides and sophorose (2-O-β-D-glucopyranosyl-D-glucose) (80). TM1223 is transcribed at high levels in the presence of carboxymethylcellulose and barley glucan (21), as well as cellobiose (Montero and Kelly, manuscript in preparation).

After import into *T. maritima* cells, cello-oligosaccharides must be broken down further for utilization. A homolog to the intracellular 1,4-β-D-glucan glucohydrolase A of *T. neapolitana*, also known as BglA (115), has not been identified in the genome

sequence of *T. maritima* MSB8, but was found in MSB8 by Liebl et al (100). A transcript for BglA was also detected by Nguyen and colleagues using real time PCR (136). It remains unclear whether this protein is present in all isolates of *T. maritima* MSB8, but the characterized protein from *T. neapolitana* is active on cellotetraose, cellotriose, cellobiose and lactose (115). The cellobiose phosphorylases of *T. neapolitana* (CbpA) and *T. maritima* (CepA) have both been characterized (163, 245). In the absence of a cellobiohydrolase in these organisms, cellobiose phosphorylase performs phosphorolysis of cellobiose, releasing one glucose-1-phosphate and one glucose molecule per disaccharide. The *T. maritima* enzyme (corresponding to TM1848) has also been shown to perform synthetic reactions using various monosaccharides (D-glucose, D-mannose, D-xylose, D-glucosamine, 2- and 6-deoxy-D-glucose and methyl-β-D-glucoside) as glucosyl acceptors (163). It has been proposed that cellobiose phosphorylase is part of an ATP-conserving pathway of β-glucan hydrolysis in *Thermotoga* species which includes BglA in *T. neapolitana* (245).

A likely regulator for the β-1,4-glucan-cellobiose regulon is TM1218, a LacI family regulator found with the genes encoding the likely *T. maritima* cellooligosaccharide transporter. An inverted repeat resembling a LacI consensus binding site is found upstream of CbtA (TM1223), suggesting that TM1218 may regulate expression of the transporter genes (21). Highly similar sequences upstream of the other cellobiose-responsive genes discussed above (e.g., Cel12A and Cel12B, CepA) are consistent with microarray expression patterns which show high transcript levels of these genes during growth on β-1,4-glucans (21, 27).

3.1.2.2.β-1,3 glucan and β-1,6 glucan

Laminarin, a β-1,3-linked glucose polymer, is a plant-derived polysaccharide used as a growth substrate by *T. maritima* and *T. neapolitana* (15). Extracellular hydrolysis of laminarin by these species is likely accomplished by the laminarinase Lam16A (TM0024, EC 3.2.1.6), versions of which have been characterized in both species (13, 15, 252, 253). Transcriptional data suggest that import of laminarin oligosaccharides may be accomplished via a nearby Opp family transporter, whose substrate binding protein is highly transcribed in the presence of laminarin (TM0031) (27). The high identity observed between TM0031 and both *P. furiosus* CbtA and *T. maritima* CbtA (TM1223) suggests possible LGT followed by divergence of substrate and regulation. A GC-rich inverted repeat which may be capable of forming a hairpin structure is found downstream of TM1223 and TM0031 perhaps explaining the difference in transcript levels and changes observed for these substrate binding proteins as compared to other transporter subunits (27). Intracellular hydrolysis of laminarin disaccharides is likely performed by the *T. maritima* homolog (TM0025) of the characterized *T. neapolitana* laminarinase *bglB* (254). Incubation of laminarin with Lam16A and BglB has been shown to accomplish complete hydrolysis of laminarin to glucose. Although not yet verified experimentally, transcription of the laminarin responsive operon may be under the control of the XylR family transcriptional regulator encoded by TM0032, which is found with the predicted laminarin transporter and laminarin hydrolases (27).

Pustulan, a β-1,6 linked glucose polysaccharide, also triggers increased transcription of the *T. maritima* laminarin utilization pathway. The transcriptional response of *T. maritima* to barley glucan, a mixed linkage β-1,3(4) glucose

polysaccharide, incorporates elements of both the responses to β-1,4 glucans and β-1,3 glucans, including increased transcription of the main operons responsive to cellobiose and laminarin. These findings are consistent with the presence in *T. maritima* Lam16A of two family 4 carbohydrate binding domains, one of which binds laminarin and mixed β-1,3/β-1,4 linkage polysaccharides, and the other of which binds pustulan, curdlan (β-1,3 glucan) and glucans derived from fungal cell walls (252). These carbohydrate binding domains enhance the hydrolysis of their respective substrates (252). A likely pathway for the utilization of laminarin and pustulan by *T. maritima* is shown in Figure 1.2.

3.1.2.3.Mannans and Mannose

Thermotoga species produce a variety of extracellular enzymes to degrade mannan polysaccharides. Hydrolysis of the backbone of β-1,4 linked mannans in the extracellular environment is likely accomplished by an endo-1,4-β-mannanase Man5/ManB (TM1227) (32, 147), which is active on glucomannan and galactomannan and responds transcriptionally to the presence of both sugars (21, 22, 27). Endomannanase degradation products are likely imported via ABC transporters, followed by further processing within the cells. Mannan oligosaccharides are likely hydrolyzed by the intracellular mannanase ManA/Man2 (TM1624) whose properties have been shown to be similar in *T. maritima* and *T. neapolitana* (32, 147). Hydrolysis of intracellular glucomannan degradation products apparently involves two glucomannanases, Cel5A (TM1751) and Cel5B (TM1752) (22). The galactose side chains attached to galactomannan are likely cleaved by the intracellular α-1,6-galactosidase GalA/Gal36A (TM1192), first isolated from *T. neapolitana* (32) and later characterized in *T. maritima*

(103). Incidentally, work in *T. neapolitana* revealed the presence of an active transport system induced by lactose and galactose (42), and later microarray experiments suggested that an oligopeptide family transporter system (TM1194, TM1196-9) might be responsible for this activity (136).

A number of candidates for mannan degradation product uptake have been suggested by functional genomics or biochemical characterization of *T. maritima* transporters. Some components of the Opp family ABC transporter (TM1746-TM1750) upstream of the Cel5A and Cel5B intracellular glucomannanases are transcribed more highly on mannose and β-mannans than other carbohydrates (21, 27), although the specificity of the substrate binding protein (TM1746) has not yet been determined. Increased transcription of the substrate binding protein TM1226 has been observed during growth on β-mannans and mannose (27). The close proximity of this gene to the binding protein of the putative *T. maritima* cellobiose/β-glucan transporter (TM1223, CbtA) and the high identity between the two (~60%) might indicate duplication of the ancestral binding protein and specialization of the paralogs to accommodate binding of different oligosaccharides. Confirmation of this hypothesis will require biochemical verification of the substrates for the two binding proteins. Biochemical characterization of the Opp family substrate binding protein encoded by TM1204 showed that despite a transcriptional response by the corresponding gene to lactose, the protein did not bind this sugar; rather, it binds maltose and β-mannotetraose, perhaps suggesting a role in the uptake of β-mannan degradation products (130).

Alpha linked mannans may also be imported for utilization by *T. maritima* cells. α–Mannobiose is a plausible substrate for the CUT1 ABC transporter encoded by

TM1853-TM1855, found with an α-mannosidase (TMM, Man38A, TM1851) capable of hydrolyzing α-1,2, α -1,3, α -1,4, and α-1,6 mannobiose (128). Similar LacI family sequence motifs are present upstream of the LacI regulator TM1856 and substrate binding protein TM1855. The function of another protein encoded within this locus (TM1852) remains unknown, but a homologous gene (TM1225) is highly expressed in the presence of β-mannans and co-localizes with Man5B. The TM1225 gene product has been crystallized and classified as a hypothetical protein with no apparent signal peptide, although it does display some similarity to glycoside hydrolases of family 32. It is also related (29% id/331 aa) to a protein of unknown function (Unk1) from a *Cellvibrio mixtus* galactomannan utilization locus (AAS19693). The conserved positioning of homologs to this gene in three separate mannan utilization loci suggests some undetermined function related to mannan utilization. A possible role as an α-mannosidase has been predicted for TM1231 (Man38B), and transcripts of this gene have been observed to be 2-4 fold higher on mannose and glucomannan than for other sugars (Conners and Kelly, unpublished data).

3.1.2.4.Mannose incorporation

After transport and complete hydrolysis of mannans has been accomplished by *T. maritima*, galactose derived from galactomannan side chains is likely phosphorylated by the putative galactokinase encoded by TM1190 (103). Cellobiose resulting from glucomannan breakdown is likely acted upon by cellobiose phosphorylase, yielding glucose-1-phosphate and glucose which can be phosphorylated by glucokinase. However, the pathway for incorporation of mannose into central metabolism it is not clear. Without

a PTS system, direct phosphorylation of mannose during transport does not occur. While the ATP-dependent glucokinases of some archaeal species (e.g. *Aeropyrum pernix* (70), *Thermoproteus tenax* (31)) also accommodate mannose, the *T. maritima* ATP-dependent glucokinase is not active on mannose (51). Other sugar kinases are found within the genome, including xylulokinase homologs and a predicted rhamnulokinase, leaving open the possibility that one of these may accomplish phosphorylation of mannose. It is also possible that an alternative pathway for the incorporation of mannose into central metabolism may be used by *T. maritima*.

Extremely low production of H_2 relative to carbon dioxide (H_2 1:CO_2 19) has been observed during growth on mannose, indicating that mannose metabolism differs substantially from glucose (2:1), maltose (4:3) and cellobiose (1:2) metabolism, among others (237). In addition, several proteins with similarity to alcohol dehydrogenases have been observed to be upregulated during growth on mannose, suggesting that perhaps mannose is preferentially converted into one or more alcohols (e.g., mannitol, mannitol-1-phosphate) which are isomerized into fructose or fructose-6-phosphate before incorporation into central metabolism. The demand of such reactions on the cellular NAD+/NADH pool might reduce the availability of NADH for use by *T. maritima* hydrogenases, explaining the reduced H_2 production during growth on mannose. While the interconversion of mannose to mannitol with NAD+/NADH has been noted in celery plants as a result of enzymatic activity of mannitol:mannose 1-oxidoreductase (213), no characterized bacterial enzyme has been shown to catalyze this reversible reaction.

If the conversion of mannose into mannitol does take place in *T. maritima* cells as part of a mannose utilization pathway, candidates for possible mannitol dehydrogenases

can be suggested by transcriptional data. For example, transcripts of TM0068 are detected at highest levels during growth on mannose and glucomannan (Conners and Kelly, unpublished observation). Although TM0068 is believed to play the role of D-mannonate oxidoreductase in the pathway leading from glucuronates to KDG (see next section), it belongs to a family of proteins (COG0246) which also includes NADH and NADPH dependent D-mannitol-2-dehydrogenases (EC 1.1.1.67, mannitol +NAD+ → fructose + NADH) from *Lactobacillus* and *Pseudomonas* species (18, 107, 178), and D-mannitol-1-phosphate 5-dehydrogenases (E.C. 1.1.1.17, Mannitol-1-phosphate+NAD+ → D-fructose-6-phosphate + NADH (87, 234).

A candidate locus for the metabolism of mannitol consists of a transaldolase-related protein (TM0295) with similarity to fructose-6-phosphate aldolases (PFAM00923), and a fructokinase (TM0296) divergently transcribed from two alcohol dehydrogenase related proteins (TM0297 and TM0298). Both TM0295 and TM0297 are highly expressed during growth on arabinose, ribose, mannose, and xylan (Conners and Kelly, unpublished data). A potential pathway for the conversion of mannose to mannitol might involve TM0297, a member of the FabG family of oxidoreductases and dehydrogenases. Surprisingly, its highest identity (50% identical/255 aa) is to a domain within a hypothetical protein from *Gallus gallus* (COG1028). TM0298 is related to a *Lactobacillus reuteri* ATCC 53608 NADH-dependent mannitol-2-dehydrogenase (E.C. 1.1.1.138, COG1063, *mdh,* 30% id/330 aa) (180), and might be used to convert mannitol into fructose. A fructokinase (TM0296) displaying similar levels of homology with a *Clostridium* homolog (40%/318 aa) and an *Arabidopsis* protein (39% id/318 aa) is likely involved in the phosphorylation of fructose to fructose-6-P. Finally, TM0295 is a

transaldolase-like fructose-6-phosphate aldolase (pfam00923) which has been crystallized (1VPX). A possible pathway for the incorporation of mannose into central metabolism is proposed in Figure 1.3.

3.1.2.5. Pectin

Growth of *T. maritima* on the plant polymer polygalacturonic acid (pectin) has been demonstrated (78). Two main gene clusters have been implicated in the breakdown and catabolism of pectin, and are likely involved in hydrolysis of the polysaccharide. They include a characterized family 1 extracellular pectinase (PelA, TM0433) (78, 116, 142) and a family 28 extracellular exo-poly-α-D-galacturonosidase (3.2.1.82, TM0437) (143). *T. maritima* appears likely to utilize ABC transporters to take up pectin degradation products rather than a single protein transporter like ExuT (169). TM0430 and TM0431 encode ABC permease subunits showing 56% and 57% similarity respectively to the TogMN permeases utilized by *Erwinia chrysanthemi* 3937 for the uptake of oligogalacturonides (58, 59). These subunits and the co-localized substrate binding protein (TM0432) are highly transcribed in the presence of pectin (Kluskens and Kelly, unpublished observation), and appear likely candidates for the uptake of uronic acid sugars.

Hydrolysis of pectin to individual sugar units is most likely followed by metabolism to glyceraldehyde-3-phosphate via enzymes of the pentose and glucuronate interconversion pathway, encoded in a second *T. maritima* locus. The composition of the uronic acids catabolic pathway of *T. maritima* revealed by its genome sequence (132) is most similar to that of *Bacillus subtilis*. The *T. maritima* uronate isomerase (TM0064)

(188) is found divergently transcribed from other uronic acid utilization genes encoding a putative fructuronate reductase (*uxuB*, TM0068), D-mannonate hydrolase (*uxuA*, TM0069), 2-keto-3-deoxygluconate kinase (*kdgK*, TM0067), and 2-dehydro-3-deoxyphosphogluconate aldolase/4-hydroxy-2-oxoglutarate aldolase (*kdgA*, TM0066). Like *B. subtilis*, *T. maritima* appears to lack an *uxaA* ortholog and contains only one *uxaB/uxuB* homolog. This suggests that the same enzymes catabolize galacturonic and glucuronic acid. However, in most other microbial genomes, *kdgA* and *kdgK* are not usually found nearby the *uxu/uxa* genes, and these sets of genes are often regulated by different proteins whose identities vary by species. While *B. subtilis* uses two separate LacI family proteins (ExuR and KdgR) to regulate the two gene sets (120, 156), *E. coli* utilizes the FadR repressor ExuR (152). An IclR family protein acts as the pectin repressor KdgR in *E. chrysanthemi* (131, 166), and is released from its operator upon binding the uronic acid metabolite 2-keto-3-deoxygluconate (KDG) (131, 166). A gene encoding a putative transcriptional regulator of the IclR (isocitrate lyase regulator) family is found within the *T. maritima* uronic acids catabolic locus (TM0065), and the crystal structure of the corresponding protein has been determined (249). Based on this structure, it has been predicted that TM0065 likely binds DNA as a tetramer, perhaps interacting with a target sequence containing two closely spaced palindromic operator sites (249). Although the sugar binding specificity of the *T. maritima* repressor has not yet been determined, two similar palindromic sequences sites separated by a 7bp spacer are found upstream of TM0065 (Conners and Kelly, unpublished observation).

3.1.2.6.Chitin

Despite efforts to grow *T. maritima* on the β-1,4 linked polysaccharide chitin, which is an effective growth substrate for the hyperthermophilic archaeons *T. kodakarensis* (219) and *P. furiosus* (44), growth of *T. maritima* on this polysaccharide was similar to control cultures lacking the substrate. Although genes with putative functions relating to chitin or N-acetylglucosamine utilization and transport are present in the *T. maritima* genome (TM0810-0814), no differential expression of these genes was observed in the presence of chitin. Since no identifiable chitinase is found in the *T. maritima* genome, *T. maritima* may not be capable of hydrolyzing chitin. Instead, other N-acetylglucosamine-containing sugars found in the natural environment of *T. maritima* may be substrates for the proteins encoded by this locus.

3.1.2.7.Xylan

Two related xylanases have been characterized in both *T. maritima* (60, 68, 119, 233, 236) and *T. neapolitana* (228, 251). The large difference in the overall sizes of these related xylanases (1059 aa for Xyn10A vs 328 aa for Xyn10B) is explained by their domain architecture; while Xyn10A contains multiple carbohydrate binding domains, Xyn10B is a single domain xylanase. The two C-terminal carbohydrate binding domains of XynA have been shown to be specific for cellulose, while the two N-terminal domains facilitate interaction with xylan and mixed linkage glucans (119, 233). These differences may relate to specialization of the xylanases and nearby genes (e.g., transporters, hydrolases) for specific cellular (or extracellular) roles. The fusion of a carbohydrate binding domain from *Streptomyces thermoviolaceus* STX-II was shown to increase the activity of Xyn10B towards soluble xylan (77).

Both *T. maritima* xylanases are divergently transcribed from sets of Opp family ABC transporter subunits. However, this similar orientation is not the result of simple duplication. The different organization of the transporter subunits, domain organization of the xylanases, and a phylogeny of this family of transporters in *T. maritima* suggests that these operons arose from more distantly related sets of proteins (27). Expression studies have revealed that transcripts of components of both xylanase-containing divergons are higher in the presence of xylan and xylose than other sugars (21, 27). *T. maritima* XynB has recently been shown to be activated in the presence of alcohols and capable of hydrolyzing xylan to xylosides while also producing alkyl β-xylosides via transglycosylation reactions with alcohols (68).

3.1.3. Monosaccharides

3.1.3.1.Glucose

Glucose transport in *T. neapolitana* has been suggested to be accomplished via an ion-coupled system (43) whose activity is apparently constitutive (42). Activity of a periplasmic glucose-binding protein has also been detected in *T. maritima* (129), although its protein identity remains unknown. Transcriptional studies have been of limited use in identifying transporters which are transcriptionally regulated in the presence of glucose (21, 27, 136), which would be consistent with constitutive expression of these systems. However, a candidate glucose ABC transporter may be encoded by TM1149-53. These proteins are located downstream of the glucose-6-phosphate

dehydrogenase (TM1155) and 6-phosphogluconolactonase (TM1154), and the putative substrate binding protein is most similar to a *Pyrobaculum aerophilum* gene (PAE2391).

Phosphorylation of imported glucose occurs via the ATP-dependent glucokinase TM1469, a member of the ROK family of sugar kinases, which is highly specific for glucose and and 2-deoxyglucose (51). The glucokinase apparently lacks the helix-turn-helix domain found in some related glucokinases (e.g., *Streptomyces coelicolor* A3(2) (83), *Staphylococcus xylosus* (229) which function both as glucokinases and transcriptional regulators. C-13 labeling experiments have determined that the bulk of glucose utilization in *T. maritima (~87%)* proceeds mainly via the Embden-Meyerhof-Parnas (EMP) pathway, although a fraction (~13%) of glucose is processed via the Entner-Doudoroff (ED) pathway (196). Certain EMP genes have been observed to be regulated transcriptionally, including TM0208-9, TM0273, TM0688-9, TM0877, and TM1469; however, a mechanism for the regulation of these genes has not yet been determined., *T. maritima* is the only known prokaryote with both an ATP-dependent (TM0209) and a poly-phosphate dependent (TM0289) 6-phosphofructokinase (6-PFK) (30). The poly-phosphate-dependent PFK uses triphosphate and polyphosphate preferentially to diphosphate, and pyrophosphate inhibits the ATP-PFK, perhaps representing an ATP-conserving mechanism.

3.1.3.2.Rhamnose

A *T. maritima* rhamnose utilization locus has been predicted by sequence comparison (84), and increased transcripts corresponding to genes within this locus are detected during growth in the presence of rhamnose (27). The locus is present in MSB8

but not in RQ2 (127), and contains a set of oligopeptide family transporter subunits divergently transcribed from rhamnose catabolic enzymes. A number of these genes, particularly an Opp-family substrate binding protein (TM1068), are highly transcribed in the presence of rhamnose. A putative regulatory protein of the DeoR family (TM1069) is also present, suggesting that it may act as a regulator of rhamnose catabolism for *T. maritima*. Recent work in *Rhizobium leguminosarum* bv. trifolii has revealed a rhamnose regulator of the DeoR family (167) although the sequence identity to the *T. maritima* protein is only 24% over 254 residues. The *R. leguminosarum rhaR* gene is also oriented upstream of a gene for an ABC binding protein (*rhaS*), but *rhaS* encodes a sugar binding protein related to ribose and xylose binding proteins rather than oligopeptide binding proteins. However, the two gene sets may share a second mechanism for transcriptional modulation in the form of a hairpin structure found downstream of the ABC binding protein. In *R. leguminosarum*, a hairpin structure downstream of rhaS prevents transcription of other transporter subunits under non-inducing conditions (167). Similar hairpin structures are also found downstream of a number of Opp family ABC substrate binding proteins in other *T. maritima* ABC transport operons (27).

3.1.3.3.Xylose, Arabinose and Ribose

Growth of *T. maritima* on the simple sugars ribose, xylose, and arabinose has been demonstrated (27). Unlike the polysaccharides described previously, hydrolysis steps prior to transport of these pentose sugars are unnecessary. Transport of one or more of these sugars may occur via a CUT2 family ABC transporter with subunits homologous to RbsBAC ribose transporters from other species (TM0955-TM0956, TM0958). It is not

yet clear which sugars are bound by the substrate binding protein of the transport system, but transcripts for the transporter genes are detected at high levels during growth on ribose, arabinose, and xylose (27). Also found within this locus are homologs to genes implicated in ribose utilization in other species, including a homolog to the cytoplasmic ribose binding protein RbsD (76), and a putative LacI family regulator which has been shown to cluster phylogenetically with RbsRs from other bacterial species (41). Further examination of hypothetical proteins within the *T. maritima* ribose utilization gene cluster may reveal insights into strategies used by other species to process this monosaccharide, as homologs to the hypothetical proteins TM0950 and TM0957 are found with a *Lactobacillus johnsonii* locus with a similar gene content (154).

While a final determination of the specificity of the ribose-responsive substrate binding protein TM0958 will await biochemical characterization efforts, the genomic content of *T. maritima* also lends support its involvement in arabinose transport. A frameshift in a second sugar binding protein gene (TM0277) found in an arabinose utilization locus with a family 51 α-L-arabinofuranosidase (TM0281) (123) and L-arabinose isomerase (TM0276) (88, 89) suggests the functionality of the binding protein may have been lost.

3.2. Heat shock response of *T. maritima*

Little was known about the transcriptional response of *T. maritima* cells to temperature stress prior to recent functional genomics studies (159). Genomic data suggested that *T. maritima* lacked an ortholog to *rpoH/sigB* encoding the sigma factor σ^{32}/σ^{B}, a major positive regulator of the heat shock response in *E. coli*, *B. subtilis,* and

other mesophiles (6, 47). Genes encoding other known regulators of stress response (e.g. *rpoS*, *ctsR*) were also lacking. However, the *T. maritima* genome sequence contained a homolog to the *E. coli* extracytoplasmic stress sigma factor σ^E and several important heat shock proteins, including HrcA, GroESL, DnaK, DnaJ and GrpE (132). To gain insight into transcriptional mechanisms of heat shock response employed by this bacterium, a targeted microarray was used to examine transcriptional profiles of selected *T. maritima* genes before and at multiple time points after a temperature shift from 80^o to 90^oC (159).

In the absence of an *rpoH/sigB* ortholog, transcripts levels of genes encoding σ^A and σ^E orthologs increased after heat shock. It may be that a heat-inducible vegetative sigma factor is beneficial to *T. maritima* in environments which experience temperature fluctuations. Homologs to the heat shock regulator HrcA have been detected in more than 50 microbial genomes (207), including *T. maritima* (132). In most species, HrcA acts as a thermosensor and repressor of heat shock chaperones which binds the Controlling Inverted Repeat of Chaperone *E*xpression (CIRCE) element until deactivated by temperature stress (250). The *T. maritima* HrcA protein has been shown to repress transcription of a *B. subtilis* reporter construct containing the dnaK promoter-operator region (235), and transcriptional data coupled with sequence analysis strongly suggest that an HrcA-mediated heat shock response is present in *T. maritima*. Transcripts of *T. maritima* homologs to CIRCE regulon genes (e.g. *hrcA*, *groESL*, *dnaK*) were detected at much higher levels after heat shock, consistent with the presence of palindromic CIRCE sequences upstream of both *hrcA* and *groEL*. We have also observed that smaller temperature increases (e.g., 80 to 85) can increase transcript levels of these heat shock

genes (203). The recent determination of the structure of *T. maritima* HrcA represents the first structural information for this family of proteins (105).

Recent work in *B. subtilis* and *Caulobacter crescentus* has demonstrated that HrcA is a target of the GroESL chaperone machinery (124, 217). A complex strategy for regulating transcript levels of the various components of the heat shock chaperone complex has been described in *B. subtilis* (55). The *T. maritima* strategy for differential regulation of chaperone subunits has not yet been determined; however, some differential regulation of transcript levels may be accomplished by a genomic rearrangement which pairs DnaK (TM0373) with a small heat shock protein (TM0374) in a locus distant from the *hrcA-grpE-dnaJ* gene string, a unique arrangement is also seen in *Chlorobium tepidum* (33).

A major difference between *T. maritima* and model mesophiles is the apparent lack of regulation of most of its proteases in response to temperature stress. This may be because homologs to σ^{32} and the repressor CtsR, a major regulator of the mesophilic proteolytic response, are missing in *T. maritima*. Alternatively, *T. maritima* cells may gain a survival advantage from constitutive expression of most proteases. Proteases whose transcripts were detected at higher levels during heat stress included a homolog to the HrtA heat shock serine protease (TM0571), and two Clp ATPase subunit homologs, ClpC-1 (TM0198, 820 aa) and ClpC-2 (TM0873, 314 aa). A frameshift is present in orf TM0873, leaving the functionality of the resulting protein in question.

3.3. Biofilm formation in *T. maritima*

Mimicking conditions which might arise in natural habitats of *T. maritima* has yielded new insights into ecological interactions in which the cells participate. Microarrays have been used to probe transcriptional changes in response to growth of *T. maritima* in homogeneous and heterogeneous cell communities. A full genome microarray was used to compare expression profiles of planktonic and sessile cells from a continuous culture reactor (158). Despite the heterogeneity of the cell populations, distinct patterns of differential expression were observed for a number of predicted operons. Consistent with observations from transcriptional studies of mesophilic biofilms, transcripts from heat shock (e.g., *dnaK, smHSP*) genes were observed to be higher in biofilm cells, while transcripts of a cold shock gene (150) were lower in biofilm cells. Cellobiose phosphorylase was highly up-regulated in biofilm cells, along with endoglucanases involved in the processing of β-1,4 linked polysaccharides (TM1524-TM1525, CelA and CelB), possibly as a result of the assembly or recycling of biofilm material. Additionally, transcripts relating to iron-sulfur cluster synthesis and repair were detected at higher levels in sessile cells, including chaperones, transporters, and iron-sulfur cluster proteins. More recent work has clarified the relationship between heat shock chaperones and iron-sulfur cluster chaperones. Wu and co-workers used the *T. maritima* versions of IscU, DnaK and DnaJ used to demonstrate that DnaK stabilizes Fe-S clusters bound by IscU and inhibits the transfer of these clusters to other proteins (240).

3.4. Quorum sensing in *T. maritima*

Single-organism cultures can reveal much about the growth physiology or organisms; however, mixed cultures containing more than one organism can offer additional insights into mechanisms relevant to growth in cell communities. Extremely high cell densities have been achieved during growth of *T. maritima* cells in the presence of the methanogenic archaeon *M. jannaschii*, which uses the growth inhibitory hydrogen produced by *T. maritima* cells to produce methane (71). Full genome transcriptional comparisons between *T. maritima* cells in pure culture and high density co-culture with *M. jannaschii* indicated changes in sugar utilization and transport genes, including glycosyl transferases and genes encoding known glucomannan (TM1752) and a-glucan (TM1834) hydrolases, which correlated with the appearance of extensive exopolysaccharide (EPS) formation in the culture. Calcofluor staining of pure and mixed cultures indicated the presence of β-1,4 linked glycans. Analysis of biofilm material from the high-density co-culture revealed a polysaccharide composed mainly of glucose (~92%), ribose, and mannose. Transcriptional patterns indicated possible involvement of cyclic di-GMP in regulation of biofilm formation. Several GGDEF-domain containing proteins displayed different expression patterns between the pure culture and co-culture, including a putative diguanylate cyclase (TM1163) and cyclic-di-GMP phosphodiesterase (TM1184). Subsequent characterization of TM1163 has since confirmed the diguanylate cyclase activity of the protein (177). Additionally, transcripts of a small unknown orf (TM0504) were observed to increase in the co-culture. Further examination of the orf and its genomic neighborhood revealed that the ABC peptide transporter co-localized with the peptide lacks a substrate-binding protein, suggesting a possible role in peptide export.

Differential expression of a orphan ABC transporter subunit (TM0043) with a likely proteolytic motif (COG2274) suggested a possible mechanism for peptide processing prior to secretion. The addition of a synthesized form of the peptide to pure low density *T. maritima* cultures triggered formation of EPS within 30 minutes, although EPS did not form in undosed control cultures. Although peptide-mediated quorum sensing has been shown to operate in numerous gram-positive bacteria, this report is the first indication of its importance in hyperthermophilic habitats.

3.5. Growth phase effects on *T. maritima*

A comparison of growth-phase dependent transcriptional changes between pure *T. maritima* cells and cells grown in a *T. maritima/M. jannaschii* co-culture revealed that many more genes changed during the transition into stationary phase in the co-culture than the pure culture (Johnson 2005, in press). Cell aggregates formed in the co-culture but not the pure culture, likely to facilitate heterotroph-methanogen hydrogen transfer. During stationary phase in the co-culture, genes encoding numerous carbohydrate utilization enzymes and transporters were up-regulated, including members of known β-glucan degradation pathways, although the carbon source in the media was maltose (α-1,4 linked glucan disaccharide). Presumably, the up-regulation of these genes in early stationary phase is connected to aggregate disintegration, as growth-phase dependent biofilm degradation has been observed in other species (73, 212).

4. Biotechnology applications of *Thermotoga* enzymes

The extreme thermostability of *Thermotoga* enzymes is attractive for an array of biotechnology applications (for a recent review, see (26)). In particular, a number of potential uses have been described for carbohydrate active enzymes of *Thermotoga* species. Galactomannan-utilizing enzymes of *T. maritima* may be useful for breaking down guar gum used as a fracturing fluid in oil/gas wells (26). The *T. maritima* maltosyltransferase has been used to synthesize maltooligosaccharide-daidzein glycosides, which greatly increase the water solubility of daidzein isoflavone (98). High-temperature glucose-to-fructose isomerization has been demonstrated using the *T. neapolitana* and *T. maritima* (TM1667) xylose (glucose) isomerases (5), suggesting that these enzymes may improve fructose concentrations by allowing the production of high-fructose corn syrup at high temperatures. The single-domain xylanase XynB isolated from *Thermotoga* sp. strain FjSS3-B.1 has high thermostability and has been shown to be active on kraft pulp (181). Use of this enzyme or similar enzymes during paper manufacturing processes could potentially reduce the use of chemical methods for bleaching. A recombinant version of *T. maritima* XynB (TM0070) has shown high alkaline stability during pulp pretreatment processes and reduces the need for chlorine during biobleaching of wheat straw pulp (69).

5. Structural studies of *T. maritima* enzymes

A substantial fraction of the >150 solved crystal structures available from PDB for *T. maritima* proteins are the result of a structural genomics effort aimed at the *T. maritima* proteome (29, 96). This high-throughput pipeline developed by the Joint Center

for Structural Genomics at the Scripps Institute has resulted in the production of 469 crystal hits for *T. maritima* proteins (29). Priority for structural determination within this group was given to 269 targets with low similarity to known structures. To date, a number of novel folds have been identified using this approach (97, 114, 168, 241). A website is available to check the progress of structure determination efforts for all *T. maritima* proteins: http://www.jcsg.org/scripts/prod/public_targets/pub_target_list.cgi. Publications resulting from the JCSG crystallization pipeline and other investigations since 1999 are listed in Table 1.4. A global examination of protein properties from *T. maritima* structures has revealed that 73% have higher contact order than their structurally characterized mesophilic counterparts (171).

Independent structural studies have also illustrated the value of structural approaches for revealing insights into the biology of *T. maritima*. Aspartate dehydrogenase was identified as a nonhomologous substitution for L-aspartate oxidase (NadC) in the *T. maritima* NAD biosynthesis pathway through structural and functional studies of TM1643, a previously uncharacterized reading frame situated between the *T. maritima* homologs of *nadA* and *nadB* (244). Functional studies of aspartate dehydrogenase revealed that like aspartate oxidase, it catalyzes the oxidation of L-aspartate to iminoaspartate. However, NAD(P) is reduced in contrast to FAD, used as a cofactor by most NadC proteins. Previously uncharacterized homologs to TM1643 are apparent in the genomes of numerous archaea and bacteria. Insight into the function of another large family of proteins was obtained through structural studies of TM0841, a putative fatty-acid binding protein (186). Members of this family, related to *B. subtilis*

DegV, are classified into COG1307 and may play a role in the transport or metabolism of fatty acids.

6. Summary

Since its isolation in 1986, *T. maritima* has become an important model organism for bacterial hyperthermophiles. Its genome sequence revealed massive lateral gene transfer with archaea (132) and comparative studies have uncovered stunning diversity within the *Thermotoga* genus (127). Despite the lack of a well-developed genetic system for *T. maritima*, alternative strategies have been used to predict pathways and gene functions of this organism, resulting in a wealth of new information. In particular, the application of high-throughput genomics methods (e.g., microarrays, high-throughput structural genomics) to *T. maritima* has provided complementary information to single-enzyme studies. Expression studies have provided clues about the regulation of carbohydrate transport and utilization pathways, and important insights into the ecological behavior of *T. maritima*. Global studies of protein structure made possible by high-throughput structural genomics efforts have revealed new insights into the extreme thermostability of *T. maritima* enzymes. The success of these efforts will undoubtedly guide further work on *T. maritima* and other interesting organisms for which genetic systems have not yet been developed.

CHAPTER 1 REFERENCES

1. **Alberto, F., C. Bignon, G. Sulzenbacher, B. Henrissat, and M. Czjzek.** 2004. The three-dimensional structure of invertase (beta-fructosidase) from *Thermotoga maritima* reveals a bimodular arrangement and an evolutionary relationship between retaining and inverting glycosidases. J Biol Chem **279:**18903-10.

2. **Andersson, U., and P. Radstrom.** 2002. Physiological function of the maltose operon regulator, MalR, in *Lactococcus lactis*. BMC Microbiol **2:**28.

3. **Bakolitsa, C., R. Schwarzenbacher, D. McMullan, L. S. Brinen, J. M. Canaves, X. Dai, A. M. Deacon, M. A. Elsliger, S. Eshagi, R. Floyd, A. Godzik, C. Grittini, S. K. Grzechnik, L. Jaroszewski, C. Karlak, H. E. Klock, E. Koesema, J. S. Kovarik, A. Kreusch, P. Kuhn, S. A. Lesley, T. M. McPhillips, M. D. Miller, A. Morse, K. Moy, J. Ouyang, R. Page, K. Quijano, A. Robb, G. Spraggon, R. C. Stevens, H. van den Bedem, J. Velasquez, J. Vincent, F. von Delft, X. Wang, B. West, G. Wolf, K. O. Hodgson, J. Wooley, and I. A. Wilson.** 2004. Crystal structure of an orphan protein (TM0875) from *Thermotoga maritima* at 2.00-A resolution reveals a new fold. Proteins **56:**607-10.

4. **Balk, M., J. Weijma, and A. J. Stams.** 2002. *Thermotoga lettingae* sp. nov., a novel thermophilic, methanol-degrading bacterium isolated from a thermophilic anaerobic reactor. Int J Syst Evol Microbiol **52:**1361-8.

5. **Bandlish, R. K., J. Michael Hess, K. L. Epting, C. Vieille, and R. M. Kelly.** 2002. Glucose-to-fructose conversion at high temperatures with xylose (glucose)

isomerases from *Streptomyces murinus* and two hyperthermophilic *Thermotoga* species. Biotechnol Bioeng **80:**185-94.

6. **Benson, A. K., and W. G. Haldenwang.** 1993. The sigma B-dependent promoter of the *Bacillus subtilis sigB* operon is induced by heat shock. J Bacteriol **175:**1929-35.

7. **Bent, C. J., N. W. Isaacs, T. J. Mitchell, and A. Riboldi-Tunnicliffe.** 2004. Crystal structure of the response regulator 02 receiver domain, the essential YycF two-component system of Streptococcus pneumoniae in both complexed and native states. J Bacteriol **186:**2872-9.

8. **Berezina, O. V., N. A. Lunina, V. V. Zverlov, D. G. Naumov, W. Liebl, and G. A. Velikodvorskaia.** 2003. [*Thermotoga neapolitana* gene clusters participating in degradation of starch and maltodextins: molecular structure of the locus]. Mol Biol (Mosk) **37:**801-9.

9. **Bertini, I., J. A. Cowan, C. Del Bianco, C. Luchinat, and S. S. Mansy.** 2003. *Thermotoga maritima* IscU. Structural characterization and dynamics of a new class of metallochaperone. J Mol Biol **331:**907-24.

10. **Bibel, M., C. Brettl, U. Gosslar, G. Kriegshauser, and W. Liebl.** 1998. Isolation and analysis of genes for amylolytic enzymes of the hyperthermophilic bacterium *Thermotoga maritima*. FEMS Microbiol Lett **158:**9-15.

11. **Bilwes, A. M., C. M. Quezada, L. R. Croal, B. R. Crane, and M. I. Simon.** 2001. Nucleotide binding by the histidine kinase CheA. Nat Struct Biol **8:**353-60.

12. **Bonin, I., R. Robelek, H. Benecke, H. Urlaub, A. Bacher, G. Richter, and M. C. Wahl.** 2004. Crystal structures of the antitermination factor NusB from *Thermotoga maritima* and implications for RNA binding. Biochem J **383:**419-28.

13. **Boraston, A. B., R. A. Warren, and D. G. Kilburn.** 2001. beta-1,3-Glucan binding by a thermostable carbohydrate-binding module from *Thermotoga maritima*. Biochemistry **40:**14679-85.

14. **Brinen, L. S., J. M. Canaves, X. Dai, A. M. Deacon, M. A. Elsliger, S. Eshaghi, R. Floyd, A. Godzik, C. Grittini, S. K. Grzechnik, C. Guda, L. Jaroszewski, C. Karlak, H. E. Klock, E. Koesema, J. S. Kovarik, A. Kreusch, P. Kuhn, S. A. Lesley, D. McMullan, T. M. McPhillips, M. A. Miller, M. D. Miller, A. Morse, K. Moy, J. Ouyang, A. Robb, K. Rodrigues, T. L. Selby, G. Spraggon, R. C. Stevens, H. van den Bedem, J. Velasquez, J. Vincent, X. Wang, B. West, G. Wolf, S. S. Taylor, K. O. Hodgson, J. Wooley, and I. A. Wilson.** 2003. Crystal structure of a zinc-containing glycerol dehydrogenase (TM0423) from *Thermotoga maritima* at 1.5 A resolution. Proteins **50:**371-4.

15. **Bronnenmeier, K., A. Kern, W. Liebl, and W. L. Staudenbauer.** 1995. Purification of *Thermotoga maritima* enzymes for the degradation of cellulosic materials. Appl Environ Microbiol **61:**1399-407.

16. **Brown, P. N., C. P. Hill, and D. F. Blair.** 2002. Crystal structure of the middle and C-terminal domains of the flagellar rotor protein FliG. Embo J **21:**3225-34.

17. **Brown, P. N., M. A. Mathews, L. A. Joss, C. P. Hill, and D. F. Blair.** 2005. Crystal structure of the flagellar rotor protein FliN from *Thermotoga maritima*. J Bacteriol **187:**2890-902.

18. **Brunker, P., J. Altenbuchner, K. D. Kulbe, and R. Mattes.** 1997. Cloning, nucleotide sequence and expression of a mannitol dehydrogenase gene from *Pseudomonas fluorescens* DSM 50106 in *Escherichia coli*. Biochim Biophys Acta **1351:**157-67.

19. **Burke, J., A. Roujeinikova, P. J. Baker, S. Sedelnikova, C. Raasch, W. Liebl, and D. W. Rice.** 2000. Crystallization and preliminary X-ray crystallographic studies on maltosyltransferase from *Thermotoga maritima*. Acta Crystallogr D Biol Crystallogr **56 (Pt 8):**1049-50.

20. **Canaves, J. M.** 2004. Predicted role for the archease protein family based on structural and sequence analysis of TM1083 and MTH1598, two proteins structurally characterized through structural genomics efforts. Proteins **56:**19-27.

21. **Chhabra, S. R., K. R. Shockley, S. B. Conners, K. L. Scott, R. D. Wolfinger, and R. M. Kelly.** 2003. Carbohydrate-induced differential gene expression patterns in the hyperthermophilic bacterium *Thermotoga maritima*. J Biol Chem **278:**7540-52.

22. **Chhabra, S. R., K. R. Shockley, D. E. Ward, and R. M. Kelly.** 2002. Regulation of endo-acting glycosyl hydrolases in the hyperthermophilic bacterium *Thermotoga maritima* grown on glucan- and mannan-based polysaccharides. Appl Environ Microbiol **68:**545-54.

23. **Christodoulou, E., W. R. Rypniewski, and C. R. Vorgias.** 2003. High-resolution X-ray structure of the DNA-binding protein HU from the hyper-thermophilic *Thermotoga maritima* and the determinants of its thermostability. Extremophiles **7:**111-22.

24. **Cole, S. T., and O. Raibaud.** 1986. The nucleotide sequence of the malT gene encoding the positive regulator of the *Escherichia coli* maltose regulon. Gene **42:**201-8.

25. **Columbus, L., W. Peti, T. Etezady-Esfarjani, T. Herrmann, and K. Wuthrich.** 2005. NMR structure determination of the conserved hypothetical protein TM1816 from *Thermotoga maritima*. Proteins **60:**552-7.

26. **Comfort, D. A., S. R. Chhabra, S. B. Conners, C. J. Chou, K. L. Epting, M. R. Johnson, K. L. Jones, A. C. Sehgal, and R. M. Kelly.** 2004. Strategic biocatalysis with hyperthermophilic enzymes. Green Chemistry **6:**459-465.

27. **Conners, S. B., C. I. Montero, D. A. Comfort, K. R. Shockley, M. R. Johnson, S. R. Chhabra, and R. M. Kelly.** 2005. An expression-driven approach to the prediction of carbohydrate transport and utilization regulons in the hyperthermophilic bacterium *Thermotoga maritima*. J Bacteriol **187:**7267-82.

28. **Dams, T., G. Auerbach, G. Bader, U. Jacob, T. Ploom, R. Huber, and R. Jaenicke.** 2000. The crystal structure of dihydrofolate reductase from *Thermotoga maritima*: molecular features of thermostability. J Mol Biol **297:**659-72.

29. **DiDonato, M., A. M. Deacon, H. E. Klock, D. McMullan, and S. A. Lesley.** 2004. A scaleable and integrated crystallization pipeline applied to mining the *Thermotoga maritima* proteome. J Struct Funct Genomics **5:**133-46.

30. **Ding, Y. R., R. S. Ronimus, and H. W. Morgan.** 2001. *Thermotoga maritima* phosphofructokinases: expression and characterization of two unique enzymes. J Bacteriol **183:**791-4.

31. **Dorr, C., M. Zaparty, B. Tjaden, H. Brinkmann, and B. Siebers.** 2003. The hexokinase of the hyperthermophile *Thermoproteus tenax*. ATP-dependent hexokinases and ADP-dependent glucokinases, teo alternatives for glucose phosphorylation in Archaea. J Biol Chem **278:**18744-53.

32. **Duffaud, G. D., C. M. McCutchen, P. Leduc, K. N. Parker, and R. M. Kelly.** 1997. Purification and characterization of extremely thermostable beta-mannanase, beta-mannosidase, and alpha-galactosidase from the hyperthermophilic eubacterium *Thermotoga neapolitana* 5068. Appl Environ Microbiol **63:**169-77.

33. **Eisen, J. A., K. E. Nelson, I. T. Paulsen, J. F. Heidelberg, M. Wu, R. J. Dodson, R. Deboy, M. L. Gwinn, W. C. Nelson, D. H. Haft, E. K. Hickey, J. D. Peterson, A. S. Durkin, J. L. Kolonay, F. Yang, I. Holt, L. A. Umayam, T. Mason, M. Brenner, T. P. Shea, D. Parksey, W. C. Nierman, T. V. Feldblyum, C. L. Hansen, M. B. Craven, D. Radune, J. Vamathevan, H. Khouri, O. White, T. M. Gruber, K. A. Ketchum, J. C. Venter, H. Tettelin, D. A. Bryant, and C. M. Fraser.** 2002. The complete genome sequence of *Chlorobium tepidum* TLS, a photosynthetic, anaerobic, green-sulfur bacterium. Proc Natl Acad Sci U S A **99:**9509-14.

34. **Elferink, M. G., S. V. Albers, W. N. Konings, and A. J. Driessen.** 2001. Sugar transport in *Sulfolobus solfataricus* is mediated by two families of binding protein-dependent ABC transporters. Mol Microbiol **39:**1494-503.

35. **Erlandsen, H., J. M. Canaves, M. A. Elsliger, F. von Delft, L. S. Brinen, X. Dai, A. M. Deacon, R. Floyd, A. Godzik, C. Grittini, S. K. Grzechnik, L.**

Jaroszewski, H. E. Klock, E. Koesema, J. S. Kovarik, A. Kreusch, P. Kuhn, S. A. Lesley, D. McMullan, T. M. McPhillips, M. D. Miller, A. Morse, K. Moy, J. Ouyang, R. Page, A. Robb, K. Quijano, R. Schwarzenbacher, G. Spraggon, R. C. Stevens, H. van den Bedem, J. Velasquez, J. Vincent, X. Wang, B. West, G. Wolf, K. O. Hodgson, J. Wooley, and I. A. Wilson. 2004. Crystal structure of an HEPN domain protein (TM0613) from *Thermotoga maritima* at 1.75 A resolution. Proteins **54:**806-9.

36. **Etezady-Esfarjani, T., T. Herrmann, W. Peti, H. E. Klock, S. A. Lesley, and K. Wuthrich.** 2004. NMR structure determination of the hypothetical protein TM1290 from *Thermotoga maritima* using automated NOESY analysis. J Biomol NMR **29:**403-6.

37. **Etezady-Esfarjani, T., W. Peti, and K. Wuthrich.** 2003. NMR assignment of the conserved hypothetical protein TM1290 of *Thermotoga maritima*. J Biomol NMR **25:**167-8.

38. **Etezady-Esfarjani, T., and K. Wuthrich.** 2004. NMR assignment of TM1442, a putative anti-sigma factor antagonist from *Thermotoga maritima*. J Biomol NMR **29:**99-100.

39. **Fardeau, M. L., B. Ollivier, B. K. Patel, M. Magot, P. Thomas, A. Rimbault, F. Rocchiccioli, and J. L. Garcia.** 1997. *Thermotoga hypogea* sp. nov., a xylanolytic, thermophilic bacterium from an oil-producing well. Int J Syst Bacteriol **47:**1013-9.

40. **Fernandez, F. J., M. C. Vega, F. Lehmann, E. Sandmeier, H. Gehring, P. Christen, and M. Wilmanns.** 2004. Structural studies of the catalytic reaction

pathway of a hyperthermophilic histidinol-phosphate aminotransferase. J Biol Chem **279:**21478-88.

41. **Fukami-Kobayashi, K., Y. Tateno, and K. Nishikawa.** 2003. Parallel evolution of ligand specificity between LacI/GalR family repressors and periplasmic sugar-binding proteins. Mol Biol Evol **20:**267-77.

42. **Galperin, M. Y., K. M. Noll, and A. H. Romano.** 1997. Coregulation of beta-galactoside uptake and hydrolysis by the hyperthermophilic bacterium *Thermotoga neapolitana*. Appl Environ Microbiol **63:**969-72.

43. **Galperin, M. Y., K. M. Noll, and A. H. Romano.** 1996. The glucose transport system of the hyperthermophilic anaerobic bacterium *Thermotoga neapolitana*. Appl Environ Microbiol **62:**2915-8.

44. **Gao, J., M. W. Bauer, K. R. Shockley, M. A. Pysz, and R. M. Kelly.** 2003. Growth of hyperthermophilic archaeon *Pyrococcus furiosus* on chitin involves two family 18 chitinases. Appl Environ Microbiol **69:**3119-28.

45. **Gaspar, J. A., C. Liu, K. A. Vassall, G. Meglei, R. Stephen, P. B. Stathopulos, A. Pineda-Lucena, B. Wu, A. Yee, C. H. Arrowsmith, and E. M. Meiering.** 2005. A novel member of the YchN-like fold: Solution structure of the hypothetical protein Tm0979 from *Thermotoga maritima*. Protein Sci **14:**216-23.

46. **Goda, S. K., O. Eisa, M. Akhter, and N. P. Minton.** 1998. Molecular analysis of the malR gene of *Clostridium butyricum* NCIMB 7423, a member of the LacI-GalR family of repressor proteins. FEMS Microbiol Lett **165:**193-200.

47. **Grossman, A. D., J. W. Erickson, and C. A. Gross.** 1984. The htpR gene product of *E. coli* is a sigma factor for heat-shock promoters. Cell **38:**383-90.

48. **Gruyer, S., E. Legin, C. Bliard, S. Ball, and F. Duchiron.** 2002. The endopolysaccharide metabolism of the hyperthermophilic archeon *Thermococcus hydrothermalis*: polymer structure and biosynthesis. Curr Microbiol **44:**206-11.

49. **Guo, R. T., C. J. Kuo, C. C. Chou, T. P. Ko, H. L. Shr, P. H. Liang, and A. H. Wang.** 2004. Crystal structure of octaprenyl pyrophosphate synthase from hyperthermophilic *Thermotoga maritima* and mechanism of product chain length determination. J Biol Chem **279:**4903-12.

50. **Ha, K. S., J. E. Kwak, B. W. Han, J. Y. Lee, J. Moon, B. I. Lee, and S. W. Suh.** 2001. Crystallization and preliminary X-ray crystallographic analysis of the TM1442 gene product from *Thermotoga maritima*, a homologue of *Bacillus subtilis* anti-anti-sigma factors. Acta Crystallogr D Biol Crystallogr **57:**276-8.

51. **Hansen, T., and P. Schonheit.** 2003. ATP-dependent glucokinase from the hyperthermophilic bacterium *Thermotoga maritima* represents an extremely thermophilic ROK glucokinase with high substrate specificity. FEMS Microbiol Lett **226:**405-11.

52. **Harriott, O. T., R. Huber, K. O. Stetter, P. W. Betts, and K. M. Noll.** 1994. A cryptic miniplasmid from the hyperthermophilic bacterium *Thermotoga* sp. strain RQ7. J Bacteriol **176:**2759-62.

53. **Heine, A., J. M. Canaves, F. von Delft, L. S. Brinen, X. Dai, A. M. Deacon, M. A. Elsliger, S. Eshaghi, R. Floyd, A. Godzik, C. Grittini, S. K. Grzechnik, C. Guda, L. Jaroszewski, C. Karlak, H. E. Klock, E. Koesema, J. S. Kovarik, A. Kreusch, P. Kuhn, S. A. Lesley, D. McMullan, T. M. McPhillips, M. A. Miller, M. D. Miller, A. Morse, K. Moy, J. Ouyang, R. Page, A. Robb, K.**

Rodrigues, R. Schwarzenbacher, T. L. Selby, G. Spraggon, R. C. Stevens, H. van den Bedem, J. Velasquez, J. Vincent, X. Wang, B. West, G. Wolf, K. O. Hodgson, J. Wooley, and I. A. Wilson. 2004. Crystal structure of O-acetylserine sulfhydrylase (TM0665) from *Thermotoga maritima* at 1.8 A resolution. Proteins **56:**387-91.

54. **Hettwer, S., and R. Sterner.** 2002. A novel tryptophan synthase beta-subunit from the hyperthermophile *Thermotoga maritima*. Quaternary structure, steady-state kinetics, and putative physiological role. J Biol Chem **277:**8194-201.

55. **Homuth, G., A. Mogk, and W. Schumann.** 1999. Post-transcriptional regulation of the *Bacillus subtilis dnaK* operon. Mol Microbiol **32:**1183-97.

56. **Huber, R., T. A. Langworthy, H. Konig, M. Thomm, C. R. Woese, U. B. Sleytr, and K. O. Stetter.** 1986. *Thermotoga maritima* sp. nov. represents a new genus of unique extremely thermophilic eubacteria growing up to 90 DegreesC. Archives of Microbiology **144:**324-333.

57. **Huber, R. W., and M. Hannig.** 2004. *Thermotogales*. *In* M. W. e. a. Dworkin, eds. (ed.), The Prokaryotes: An Evolving Electronic Resource for the Microbiological Community, 3rd edition, 3.19, 3/18/2005, vol. http://link.springer-ny.com/link/service/books/10125/. Springer-Verlag, New York.

58. **Hugouvieux-Cotte-Pattat, N., N. Blot, and S. Reverchon.** 2001. Identification of TogMNAB, an ABC transporter which mediates the uptake of pectic oligomers in *Erwinia chrysanthemi* 3937. Mol Microbiol **41:**1113-23.

59. **Hugouvieux-Cotte-Pattat, N., and S. Reverchon.** 2001. Two transporters, TogT and TogMNAB, are responsible for oligogalacturonide uptake in *Erwinia chrysanthemi* 3937. Mol Microbiol **41:**1125-32.

60. **Ihsanawati, T. Kumasaka, T. Kaneko, C. Morokuma, S. Nakamura, and N. Tanaka.** 2003. Crystallization and preliminary X-ray studies of xylanase 10B from *Thermotoga maritima*. Acta Crystallogr D Biol Crystallogr **59:**1659-61.

61. **Ilin, S., A. Hoskins, O. Ohlenschlager, H. R. Jonker, H. Schwalbe, and J. Wohnert.** 2005. Domain reorientation and induced fit upon RNA binding: solution structure and dynamics of ribosomal protein L11 from *Thermotoga maritima*. Chembiochem **6:**1611-8.

62. **Ishii, R., A. Minagawa, H. Takaku, M. Takagi, M. Nashimoto, and S. Yokoyama.** 2005. Crystal structure of the tRNA 3' processing endoribonuclease tRNase Z from *Thermotoga maritima*. J Biol Chem **280:**14138-44.

63. **Jannasch, H. W., R. Huber, S. Belkin, and K. O. Stetter.** 1988. *Thermotoga neapolitana* sp. nov. of the extremely thermophilic, eubacterial genus *Thermotoga*. Arch Microbiol **150:**103-4.

64. **Jansen, R., J. D. Embden, W. Gaastra, and L. M. Schouls.** 2002. Identification of genes that are associated with DNA repeats in prokaryotes. Mol Microbiol **43:**1565-75.

65. **Jaroszewski, L., R. Schwarzenbacher, D. McMullan, P. Abdubek, S. Agarwalla, E. Ambing, H. Axelrod, T. Biorac, J. M. Canaves, H. J. Chiu, A. M. Deacon, M. Didonato, M. A. Elsliger, A. Godzik, C. Grittini, S. K. Grzechnik, J. Hale, E. Hampton, G. W. Han, J. Haugen, M. Hornsby, H. E.**

Klock, E. Koesema, A. Kreusch, P. Kuhn, S. A. Lesley, M. D. Miller, K. Moy, E. Nigoghossian, J. Paulsen, K. Quijano, R. Reyes, C. Rife, G. Spraggon, R. C. Stevens, H. van den Bedem, J. Velasquez, J. Vincent, A. White, G. Wolf, Q. Xu, K. O. Hodgson, J. Wooley, and I. A. Wilson. 2005. Crystal structure of Hsp33 chaperone (TM1394) from *Thermotoga maritima* at 2.20 A resolution. Proteins.

66. **Jaroszewski, L., R. Schwarzenbacher, F. von Delft, D. McMullan, L. S. Brinen, J. M. Canaves, X. Dai, A. M. Deacon, M. DiDonato, M. A. Elsliger, S. Eshagi, R. Floyd, A. Godzik, C. Grittini, S. K. Grzechnik, E. Hampton, I. Levin, C. Karlak, H. E. Klock, E. Koesema, J. S. Kovarik, A. Kreusch, P. Kuhn, S. A. Lesley, T. M. McPhillips, M. D. Miller, A. Morse, K. Moy, J. Ouyang, R. Page, K. Quijano, R. Reyes, F. Rezezadeh, A. Robb, E. Sims, G. Spraggon, R. C. Stevens, H. van den Bedem, J. Velasquez, J. Vincent, X. Wang, B. West, G. Wolf, Q. Xu, K. O. Hodgson, J. Wooley, and I. A. Wilson.** 2004. Crystal structure of a novel manganese-containing cupin (TM1459) from *Thermotoga maritima* at 1.65 A resolution. Proteins **56:**611-4.

67. **Jeanthon, C., A. L. Reysenbach, S. L'Haridon, A. Gambacorta, N. R. Pace, P. Glenat, and D. Prieur.** 1995. *Thermotoga subterranea* sp. nov., a new thermophilic bacterium isolated from a continental oil reservoir. Arch Microbiol **164:**91-7.

68. **Jiang, Z., Y. Zhu, L. Li, X. Yu, I. Kusakabe, M. Kitaoka, and K. Hayashi.** 2004. Transglycosylation reaction of xylanase B from the hyperthermophilic

Thermotoga maritima with the ability of synthesis of tertiary alkyl beta-D-xylobiosides and xylosides. J Biotechnol **114:**125-34.

69. **Jiang, Z. Q., X. T. Li, S. Q. Yang, L. T. Li, Y. Li, and W. Y. Feng.** 2005. Biobleach boosting effect of recombinant xylanase B from the hyperthermophilic *Thermotoga maritima* on wheat straw pulp. Appl Microbiol Biotechnol**:**1-7.

70. **Johnsen, U., T. Hansen, and P. Schonheit.** 2003. Comparative analysis of pyruvate kinases from the hyperthermophilic archaea *Archaeoglobus fulgidus*, *Aeropyrum pernix*, and *Pyrobaculum aerophilum* and the hyperthermophilic bacterium *Thermotoga maritima*: unusual regulatory properties in hyperthermophilic archaea. J Biol Chem **278:**25417-27.

71. **Johnson, M. R., C. I. Montero, S. B. Conners, K. R. Shockley, S. L. Bridger, and R. M. Kelly.** 2005. Population density-dependent regulation of exopolysaccharide formation in the hyperthermophilic bacterium *Thermotoga maritima*. Mol Microbiol **55:**664-74.

72. **Kaiser, J. T., T. Clausen, G. P. Bourenkow, H. D. Bartunik, S. Steinbacher, and R. Huber.** 2000. Crystal structure of a NifS-like protein from *Thermotoga maritima*: implications for iron sulphur cluster assembly. J Mol Biol **297:**451-64.

73. **Kaplan, J. B., C. Ragunath, N. Ramasubbu, and D. H. Fine.** 2003. Detachment of *Actinobacillus actinomycetemcomitans* biofilm cells by an endogenous beta-hexosaminidase activity. J Bacteriol **185:**4693-8.

74. **Kim, B. C., Y. H. Lee, H. S. Lee, D. W. Lee, E. A. Choe, and Y. R. Pyun.** 2002. Cloning, expression and characterization of L-arabinose isomerase from

Thermotoga neapolitana: bioconversion of D-galactose to D-tagatose using the enzyme. FEMS Microbiol Lett **212:**121-6.

75. **Kim, C. S., E. S. Ji, and D. K. Oh.** 2004. Characterization of a thermostable recombinant beta-galactosidase from *Thermotoga maritima*. J Appl Microbiol **97:**1006-14.

76. **Kim, M. S., J. Shin, W. Lee, H. S. Lee, and B. H. Oh.** 2003. Crystal structures of RbsD leading to the identification of cytoplasmic sugar-binding proteins with a novel folding architecture. J Biol Chem **278:**28173-80.

77. **Kittur, F. S., S. L. Mangala, A. A. Rus'd, M. Kitaoka, H. Tsujibo, and K. Hayashi.** 2003. Fusion of family 2b carbohydrate-binding module increases the catalytic activity of a xylanase from *Thermotoga maritima* to soluble xylan. FEBS Lett **549:**147-51.

78. **Kluskens, L. D., G. J. van Alebeek, A. G. Voragen, W. M. de Vos, and J. van der Oost.** 2003. Molecular and biochemical characterization of the thermoactive family 1 pectate lyase from the hyperthermophilic bacterium *Thermotoga maritima*. Biochem J **370:**651-9.

79. **Knochel, T., A. Pappenberger, J. N. Jansonius, and K. Kirschner.** 2002. The crystal structure of indoleglycerol-phosphate synthase from *Thermotoga maritima*. Kinetic stabilization by salt bridges. J Biol Chem **277:**8626-34.

80. **Koning, S. M., M. G. Elferink, W. N. Konings, and A. J. Driessen.** 2001. Cellobiose uptake in the hyperthermophilic archaeon *Pyrococcus furiosus* is mediated by an inducible, high-affinity ABC transporter. J Bacteriol **183:**4979-84.

81. **Korolev, S., Y. Ikeguchi, T. Skarina, S. Beasley, C. Arrowsmith, A. Edwards, A. Joachimiak, A. E. Pegg, and A. Savchenko.** 2002. The crystal structure of spermidine synthase with a multisubstrate adduct inhibitor. Nat Struct Biol **9:**27-31.

82. **Kriegshauser, G., and W. Liebl.** 2000. Pullulanase from the hyperthermophilic bacterium *Thermotoga maritima*: purification by beta-cyclodextrin affinity chromatography. J Chromatogr B Biomed Sci Appl **737:**245-51.

83. **Kwakman, J. H., and P. W. Postma.** 1994. Glucose kinase has a regulatory role in carbon catabolite repression in *Streptomyces coelicolor*. J Bacteriol **176:**2694-8.

84. **Kyrpides, N. C., C. A. Ouzounis, I. Iliopoulos, V. Vonstein, and R. Overbeek.** 2000. Analysis of the *Thermotoga maritima* genome combining a variety of sequence similarity and genome context tools. Nucleic Acids Res **28:**4573-6.

85. **Lammerts van Bueren, A., R. Finn, J. Ausio, and A. B. Boraston.** 2004. Alpha-glucan recognition by a new family of carbohydrate-binding modules found primarily in bacterial pathogens. Biochemistry **43:**15633-42.

86. **Le Breton, Y., V. Pichereau, N. Sauvageot, Y. Auffray, and A. Rince.** 2005. Maltose utilization in *Enterococcus faecalis*. J Appl Microbiol **98:**806-13.

87. **Lee, C. A., and M. H. Saier, Jr.** 1983. Use of cloned *mtl* genes of *Escherichia coli* to introduce *mtl* deletion mutations into the chromosome. J Bacteriol **153:**685-92.

88. **Lee, D. W., Y. H. Hong, E. A. Choe, S. J. Lee, S. B. Kim, H. S. Lee, J. W. Oh, H. H. Shin, and Y. R. Pyun.** 2005. A thermodynamic study of mesophilic,

thermophilic, and hyperthermophilic L-arabinose isomerases: the effects of divalent metal ions on protein stability at elevated temperatures. FEBS Lett **579:**1261-6.

89. **Lee, D. W., H. J. Jang, E. A. Choe, B. C. Kim, S. J. Lee, S. B. Kim, Y. H. Hong, and Y. R. Pyun.** 2004. Characterization of a thermostable L-arabinose (D-galactose) isomerase from the hyperthermophilic eubacterium *Thermotoga maritima*. Appl Environ Microbiol **70:**1397-404.

90. **Lee, H. H., J. Kim do, H. J. Ahn, J. Y. Ha, and S. W. Suh.** 2004. Crystal structure of T-protein of the glycine cleavage system. Cofactor binding, insights into H-protein recognition, and molecular basis for understanding nonketotic hyperglycinemia. J Biol Chem **279:**50514-23.

91. **Lee, H. S., J. H. Auh, H. G. Yoon, M. J. Kim, J. H. Park, S. S. Hong, M. H. Kang, T. J. Kim, T. W. Moon, J. W. Kim, and K. H. Park.** 2002. Cooperative action of alpha-glucanotransferase and maltogenic amylase for an improved process of isomaltooligosaccharide (IMO) production. J Agric Food Chem **50:**2812-7.

92. **Lee, J. Y., H. J. Ahn, K. S. Ha, and S. W. Suh.** 2004. Crystal structure of the TM1442 protein from *Thermotoga maritima*, a homolog of the *Bacillus subtilis* general stress response anti-anti-sigma factor RsbV. Proteins **56:**176-9.

93. **Lee, J. Y., J. E. Kwak, J. Moon, S. H. Eom, E. C. Liong, J. D. Pedelacq, J. Berendzen, and S. W. Suh.** 2001. Crystal structure and functional analysis of the SurE protein identify a novel phosphatase family. Nat Struct Biol **8:**789-94.

94. **Lee, M. H., Y. W. Kim, T. J. Kim, C. S. Park, J. W. Kim, T. W. Moon, and K. H. Park.** 2002. A novel amylolytic enzyme from *Thermotoga maritima*, resembling cyclodextrinase and alpha-glucosidase, that liberates glucose from the reducing end of the substrates. Biochem Biophys Res Commun **295:**818-25.

95. **Lee, S. J., A. Engelmann, R. Horlacher, Q. Qu, G. Vierke, C. Hebbeln, M. Thomm, and W. Boos.** 2003. TrmB, a sugar-specific transcriptional regulator of the trehalose/maltose ABC transporter from the hyperthermophilic archaeon *Thermococcus litoralis*. J Biol Chem **278:**983-90.

96. **Lesley, S. A., P. Kuhn, A. Godzik, A. M. Deacon, I. Mathews, A. Kreusch, G. Spraggon, H. E. Klock, D. McMullan, T. Shin, J. Vincent, A. Robb, L. S. Brinen, M. D. Miller, T. M. McPhillips, M. A. Miller, D. Scheibe, J. M. Canaves, C. Guda, L. Jaroszewski, T. L. Selby, M. A. Elsliger, J. Wooley, S. S. Taylor, K. O. Hodgson, I. A. Wilson, P. G. Schultz, and R. C. Stevens.** 2002. Structural genomics of the *Thermotoga maritima* proteome implemented in a high-throughput structure determination pipeline. Proc Natl Acad Sci U S A **99:**11664-9.

97. **Levin, I., M. D. Miller, R. Schwarzenbacher, D. McMullan, P. Abdubek, E. Ambing, T. Biorac, J. Cambell, J. M. Canaves, H. J. Chiu, A. M. Deacon, M. DiDonato, M. A. Elsliger, A. Godzik, C. Grittini, S. K. Grzechnik, J. Hale, E. Hampton, G. W. Han, J. Haugen, M. Hornsby, L. Jaroszewski, C. Karlak, H. E. Klock, E. Koesema, A. Kreusch, P. Kuhn, S. A. Lesley, A. Morse, K. Moy, E. Nigoghossian, J. Ouyang, R. Page, K. Quijano, R. Reyes, A. Robb, E. Sims, G. Spraggon, R. C. Stevens, H. van den Bedem, J. Velasquez, J.**

Vincent, X. Wang, B. West, G. Wolf, Q. Xu, O. Zagnitko, K. O. Hodgson, J. Wooley, and I. A. Wilson. 2005. Crystal structure of an indigoidine synthase A (IndA)-like protein (TM1464) from *Thermotoga maritima* at 1.90 A resolution reveals a new fold. Proteins **59:**864-8.

98. **Li, D., S. H. Park, J. H. Shim, H. S. Lee, S. Y. Tang, C. S. Park, and K. H. Park.** 2004. In vitro enzymatic modification of puerarin to puerarin glycosides by maltogenic amylase. Carbohydr Res **339:**2789-97.

99. **Liebl, W., R. Feil, J. Gabelsberger, J. Kellermann, and K. H. Schleifer.** 1992. Purification and characterization of a novel thermostable 4-alpha-glucanotransferase of *Thermotoga maritima* cloned in *Escherichia coli*. Eur J Biochem **207:**81-8.

100. **Liebl, W., J. Gabelsberger, and K. H. Schleifer.** 1994. Comparative amino acid sequence analysis of *Thermotoga maritima* beta-glucosidase (BglA) deduced from the nucleotide sequence of the gene indicates distant relationship between beta-glucosidases of the BGA family and other families of beta-1,4-glycosyl hydrolases. Mol Gen Genet **242:**111-5.

101. **Liebl, W., P. Ruile, K. Bronnenmeier, K. Riedel, F. Lottspeich, and I. Greif.** 1996. Analysis of a *Thermotoga maritima* DNA fragment encoding two similar thermostable cellulases, CelA and CelB, and characterization of the recombinant enzymes. Microbiology **142 (Pt 9):**2533-42.

102. **Liebl, W., I. Stemplinger, and P. Ruile.** 1997. Properties and gene structure of the *Thermotoga maritima* alpha-amylase AmyA, a putative lipoprotein of a hyperthermophilic bacterium. J Bacteriol **179:**941-8.

103. **Liebl, W., B. Wagner, and J. Schellhase.** 1998. Properties of an alpha-galactosidase, and structure of its gene *galA*, within an alpha-and beta-galactoside utilization gene cluster of the hyperthermophilic bacterium *Thermotoga maritima*. Syst Appl Microbiol **21:**1-11.

104. **Lim, W. J., S. R. Park, C. L. An, J. Y. Lee, S. Y. Hong, E. C. Shin, E. J. Kim, J. O. Kim, H. Kim, and H. D. Yun.** 2003. Cloning and characterization of a thermostable intracellular alpha-amylase gene from the hyperthermophilic bacterium *Thermotoga maritima* MSB8. Res Microbiol **154:**681-7.

105. **Liu, J., C. Huang, D. H. Shin, H. Yokota, J. Jancarik, J. S. Kim, P. D. Adams, R. Kim, and S. H. Kim.** 2005. Crystal structure of a heat-inducible transcriptional repressor HrcA from *Thermotoga maritima*: structural insight into DNA binding and dimerization. J Mol Biol **350:**987-96.

106. **Liu, J., Y. Lou, H. Yokota, P. D. Adams, R. Kim, and S. H. Kim.** 2005. Crystal structure of a PhoU protein homologue: a new class of metalloprotein containing multinuclear iron clusters. J Biol Chem **280:**15960-6.

107. **Liu, S., B. Saha, and M. Cotta.** 2005. Cloning, expression, purification, and analysis of mannitol dehydrogenase gene *mtlK* from *Lactobacillus brevis*. Appl Biochem Biotechnol **121-124:**391-401.

108. **Lloyd, S. A., F. G. Whitby, D. F. Blair, and C. P. Hill.** 1999. Structure of the C-terminal domain of FliG, a component of the rotor in the bacterial flagellar motor. Nature **400:**472-5.

109. **Lodge, J. A., T. Maier, W. Liebl, V. Hoffmann, and N. Strater.** 2003. Crystal structure of *Thermotoga maritima* alpha-glucosidase AglA defines a new clan of NAD+-dependent glycosidases. J Biol Chem **278:**19151-8.

110. **Maes, D., J. P. Zeelen, N. Thanki, N. Beaucamp, M. Alvarez, M. H. Thi, J. Backmann, J. A. Martial, L. Wyns, R. Jaenicke, and R. K. Wierenga.** 1999. The crystal structure of triosephosphate isomerase (TIM) from *Thermotoga maritima*: a comparative thermostability structural analysis of ten different TIM structures. Proteins **37:**441-53.

111. **Makarova, K. S., L. Aravind, N. V. Grishin, I. B. Rogozin, and E. V. Koonin.** 2002. A DNA repair system specific for thermophilic Archaea and bacteria predicted by genomic context analysis. Nucleic Acids Res **30:**482-96.

112. **Mansy, S. S., S. P. Wu, and J. A. Cowan.** 2004. Iron-sulfur cluster biosynthesis: biochemical characterization of the conformational dynamics of *Thermotoga maritima* IscU and the relevance for cellular cluster assembly. J Biol Chem **279:**10469-75.

113. **Martins, L. O., L. S. Carreto, M. S. Da Costa, and H. Santos.** 1996. New compatible solutes related to Di-myo-inositol-phosphate in members of the order *Thermotogales*. J Bacteriol **178:**5644-51.

114. **Mathews, I., R. Schwarzenbacher, D. McMullan, P. Abdubek, E. Ambing, H. Axelrod, T. Biorac, J. M. Canaves, H. J. Chiu, A. M. Deacon, M. DiDonato, M. A. Elsliger, A. Godzik, C. Grittini, S. K. Grzechnik, J. Hale, E. Hampton, G. W. Han, J. Haugen, M. Hornsby, L. Jaroszewski, H. E. Klock, E. Koesema, A. Kreusch, P. Kuhn, S. A. Lesley, I. Levin, M. D. Miller, K. Moy,**

E. Nigoghossian, J. Ouyang, J. Paulsen, K. Quijano, R. Reyes, G. Spraggon, R. C. Stevens, H. van den Bedem, J. Velasquez, J. Vincent, A. White, G. Wolf, Q. Xu, K. O. Hodgson, J. Wooley, and I. A. Wilson. 2005. Crystal structure of S-adenosylmethionine:tRNA ribosyltransferase-isomerase (QueA) from *Thermotoga maritima* at 2.0 A resolution reveals a new fold. Proteins **59:**869-74.

115. **McCarthy, J. K., A. Uzelac, D. F. Davis, and D. E. Eveleigh.** 2004. Improved catalytic efficiency and active site modification of 1,4-beta-D-glucan glucohydrolase A from *Thermotoga neapolitana* by directed evolution. J Biol Chem **279:**11495-502.

116. **McDonough, M. A., C. Ryttersgaard, M. E. Bjornvad, L. Lo Leggio, M. Schulein, S. O. Schroder Glad, and S. Larsen.** 2002. Crystallization and preliminary X-ray characterization of a thermostable pectate lyase from *Thermotoga maritima*. Acta Crystallogr D Biol Crystallogr **58:**709-11.

117. **McMullan, D., R. Schwarzenbacher, L. Jaroszewski, F. von Delft, H. E. Klock, J. Vincent, K. Quijano, P. Abdubek, E. Ambing, T. Biorac, L. S. Brinen, J. M. Canaves, X. Dai, A. M. Deacon, M. DiDonato, M. A. Elsliger, S. Eshaghi, R. Floyd, A. Godzik, C. Grittini, S. K. Grzechnik, E. Hampton, C. Karlak, E. Koesema, A. Kreusch, P. Kuhn, I. Levin, T. M. McPhillips, M. D. Miller, A. Morse, K. Moy, J. Ouyang, R. Page, R. Reyes, F. Rezezadeh, A. Robb, E. Sims, G. Spraggon, R. C. Stevens, H. van den Bedem, J. Velasquez, X. Wang, B. West, G. Wolf, Q. Xu, K. O. Hodgson, J. Wooley, S. A. Lesley,**

and I. A. Wilson. 2004. Crystal structure of a novel *Thermotoga maritima* enzyme (TM1112) from the cupin family at 1.83 A resolution. Proteins **56:**615-8.

118. **Meissner, H., and W. Liebl.** 1998. *Thermotoga maritima* maltosyltransferase, a novel type of maltodextrin glycosyltransferase acting on starch and malto-oligosaccharides. Eur J Biochem **258:**1050-8.

119. **Meissner, K., D. Wassenberg, and W. Liebl.** 2000. The thermostabilizing domain of the modular xylanase XynA of *Thermotoga maritima* represents a novel type of binding domain with affinity for soluble xylan and mixed-linkage beta-1,3/beta-1, 4-glucan. Mol Microbiol **36:**898-912.

120. **Mekjian, K. R., E. M. Bryan, B. W. Beall, and C. P. Moran, Jr.** 1999. Regulation of hexuronate utilization in *Bacillus subtilis*. J Bacteriol **181:**426-33.

121. **Miller, D. J., N. Ouellette, E. Evdokimova, A. Savchenko, A. Edwards, and W. F. Anderson.** 2003. Crystal complexes of a predicted S-adenosylmethionine-dependent methyltransferase reveal a typical AdoMet binding domain and a substrate recognition domain. Protein Sci **12:**1432-42.

122. **Miller, M. D., R. Schwarzenbacher, F. von Delft, P. Abdubek, E. Ambing, T. Biorac, L. S. Brinen, J. M. Canaves, J. Cambell, H. J. Chiu, X. Dai, A. M. Deacon, M. DiDonato, M. A. Elsliger, S. Eshagi, R. Floyd, A. Godzik, C. Grittini, S. K. Grzechnik, E. Hampton, L. Jaroszewski, C. Karlak, H. E. Klock, E. Koesema, J. S. Kovarik, A. Kreusch, P. Kuhn, S. A. Lesley, I. Levin, D. McMullan, T. M. McPhillips, A. Morse, K. Moy, J. Ouyang, R. Page, K. Quijano, A. Robb, G. Spraggon, R. C. Stevens, H. van den Bedem, J. Velasquez, J. Vincent, X. Wang, B. West, G. Wolf, Q. Xu, K. O. Hodgson,**

J. Wooley, and I. A. Wilson. 2004. Crystal structure of a tandem cystathionine-beta-synthase (CBS) domain protein (TM0935) from *Thermotoga maritima* at 1.87 A resolution. Proteins **57:**213-7.

123. **Miyazaki, K.** 2005. Hyperthermophilic alpha-L-arabinofuranosidase from *Thermotoga maritima* MSB8: molecular cloning, gene expression, and characterization of the recombinant protein. Extremophiles.

124. **Mogk, A., G. Homuth, C. Scholz, L. Kim, F. X. Schmid, and W. Schumann.** 1997. The GroE chaperonin machine is a major modulator of the CIRCE heat shock regulon of *Bacillus subtilis*. Embo J **16:**4579-90.

125. **Mojica, F. J., C. Diez-Villasenor, J. Garcia-Martinez, and E. Soria.** 2005. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. J Mol Evol **60:**174-82.

126. **Mojica, F. J., C. Diez-Villasenor, E. Soria, and G. Juez.** 2000. Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. Mol Microbiol **36:**244-6.

127. **Mongodin, E. F., I. R. Hance, R. T. Deboy, S. R. Gill, S. Daugherty, R. Huber, C. M. Fraser, K. Stetter, and K. E. Nelson.** 2005. Gene transfer and genome plasticity in *Thermotoga maritima*, a model hyperthermophilic species. J Bacteriol **187:**4935-44.

128. **Nakajima, M., H. Imamura, H. Shoun, and T. Wakagi.** 2003. Unique metal dependency of cytosolic alpha-mannosidase from *Thermotoga maritima*, a hyperthermophilic bacterium. Arch Biochem Biophys **415:**87-93.

129. **Nanavati, D., K. M. Noll, and A. H. Romano.** 2002. Periplasmic maltose- and glucose-binding protein activities in cell-free extracts of *Thermotoga maritima*. Microbiology **148:**3531-7.

130. **Nanavati, D. M., T. N. Nguyen, and K. M. Noll.** 2005. Substrate specificities and expression patterns reflect the evolutionary divergence of maltose ABC transporters in *Thermotoga maritima*. J Bacteriol **187:**2002-9.

131. **Nasser, W., S. Reverchon, and J. Robert-Baudouy.** 1992. Purification and functional characterization of the KdgR protein, a major repressor of pectinolysis genes of *Erwinia chrysanthemi*. Mol Microbiol **6:**257-65.

132. **Nelson, K. E., R. A. Clayton, S. R. Gill, M. L. Gwinn, R. J. Dodson, D. H. Haft, E. K. Hickey, J. D. Peterson, W. C. Nelson, K. A. Ketchum, L. McDonald, T. R. Utterback, J. A. Malek, K. D. Linher, M. M. Garrett, A. M. Stewart, M. D. Cotton, M. S. Pratt, C. A. Phillips, D. Richardson, J. Heidelberg, G. G. Sutton, R. D. Fleischmann, J. A. Eisen, C. M. Fraser, and et al.** 1999. Evidence for lateral gene transfer between Archaea and bacteria from genome sequence of *Thermotoga maritima*. Nature **399:**323-9.

133. **Nesbo, C. L., and W. F. Doolittle.** 2003. Targeting clusters of transferred genes in *Thermotoga maritima*. Environ Microbiol **5:**1144-54.

134. **Nesbo, C. L., S. L'Haridon, K. O. Stetter, and W. F. Doolittle.** 2001. Phylogenetic analyses of two "archaeal" genes in *Thermotoga maritima* reveal multiple transfers between archaea and bacteria. Mol Biol Evol **18:**362-75.

135. **Nesbo, C. L., K. E. Nelson, and W. F. Doolittle.** 2002. Suppressive subtractive hybridization detects extensive genomic diversity in *Thermotoga maritima*. J Bacteriol **184:**4475-88.

136. **Nguyen, T. N., A. D. Ejaz, M. A. Brancieri, A. M. Mikula, K. E. Nelson, S. R. Gill, and K. M. Noll.** 2004. Whole-genome expression profiling of *Thermotoga maritima* in response to growth on sugars in a chemostat. J Bacteriol **186:**4824-8.

137. **Nieto, C., M. Espinosa, and A. Puyet.** 1997. The maltose/maltodextrin regulon of *Streptococcus pneumoniae*. Differential promoter regulation by the transcriptional repressor MalR. J Biol Chem **272:**30860-5.

138. **Noll, K. M., and M. Vargas.** 1997. Recent advances in genetic analyses of hyperthermophilic archaea and bacteria. Arch Microbiol **168:**73-80.

139. **Page, R., M. S. Nelson, F. von Delft, M. A. Elsliger, J. M. Canaves, L. S. Brinen, X. Dai, A. M. Deacon, R. Floyd, A. Godzik, C. Grittini, S. K. Grzechnik, L. Jaroszewski, H. E. Klock, E. Koesema, J. S. Kovarik, A. Kreusch, P. Kuhn, S. A. Lesley, D. McMullan, T. M. McPhillips, M. D. Miller, A. Morse, K. Moy, J. Ouyang, A. Robb, K. Rodrigues, R. Schwarzenbacher, G. Spraggon, R. C. Stevens, H. van den Bedem, J. Velasquez, J. Vincent, X. Wang, B. West, G. Wolf, K. O. Hodgson, J. Wooley, and I. A. Wilson.** 2004. Crystal structure of gamma-glutamyl phosphate reductase (TM0293) from *Thermotoga maritima* at 2.0 A resolution. Proteins **54:**157-61.

140. **Palm, D., R. Goerl, G. Weidinger, R. Zeier, B. Fischer, and R. Schinzel.** 1987. *E. coli* maltodextrin phosphorylase: primary structure and deletion mapping of the C-terminal site. Z Naturforsch [C] **42:**394-400.

141. **Pan, G., A. L. Menon, and M. W. Adams.** 2003. Characterization of a [2Fe-2S] protein encoded in the iron-hydrogenase operon of *Thermotoga maritima*. J Biol Inorg Chem **8:**469-74.

142. **Parisot, J., A. Ghochikyan, V. Langlois, V. Sakanyan, and C. Rabiller.** 2002. Exopolygalacturonate lyase from *Thermotoga maritima*: cloning, characterization and organic synthesis application. Carbohydr Res **337:**1427-33.

143. **Parisot, J., V. Langlois, V. Sakanyan, and C. Rabiller.** 2003. Cloning expression and characterization of a thermostable exopolygalacturonase from *Thermotoga maritima*. Carbohydr Res **338:**1333-7.

144. **Park, F., K. Gajiwala, B. Noland, L. Wu, D. He, J. Molinari, K. Loomis, B. Pagarigan, P. Kearins, J. Christopher, T. Peat, J. Badger, J. Hendle, J. Lin, and S. Buchanan.** 2004. The 1.59 A resolution crystal structure of TM0096, a flavin mononucleotide binding protein from *Thermotoga maritima*. Proteins **55:**772-4.

145. **Park, S. Y., B. D. Beel, M. I. Simon, A. M. Bilwes, and B. R. Crane.** 2004. In different organisms, the mode of interaction between two signaling proteins is not necessarily conserved. Proc Natl Acad Sci U S A **101:**11646-51.

146. **Park, S. Y., C. M. Quezada, A. M. Bilwes, and B. R. Crane.** 2004. Subunit exchange by CheA histidine kinases from the mesophile *Escherichia coli* and the thermophile *Thermotoga maritima*. Biochemistry **43:**2228-40.

147. **Parker, K. N., S. R. Chhabra, D. Lam, W. Callen, G. D. Duffaud, M. A. Snead, J. M. Short, E. J. Mathur, and R. M. Kelly.** 2001. Galactomannanases Man2 and Man5 from *Thermotoga* species: growth physiology on galactomannans, gene sequence analysis, and biochemical properties of recombinant enzymes. Biotechnol Bioeng **75:**322-33.

148. **Penhoat, C. H., Z. Li, H. S. Atreya, S. Kim, A. Yee, R. Xiao, D. Murray, C. H. Arrowsmith, and T. Szyperski.** 2005. NMR solution structure of *Thermotoga maritima* protein TM1509 reveals a Zn-metalloprotease-like tertiary structure. J Struct Funct Genomics **6:**51-62.

149. **Peti, W., J. Norcross, G. Eldridge, and M. O'Neil-Johnson.** 2004. Biomolecular NMR using a microcoil NMR probe--new technique for the chemical shift assignment of aromatic side chains in proteins. J Am Chem Soc **126:**5873-8.

150. **Phadtare, S., J. Hwang, K. Severinov, and M. Inouye.** 2003. CspB and CspL, thermostable cold-shock proteins from *Thermotoga maritima*. Genes Cells **8:**801-10.

151. **Pierrel, F., H. L. Hernandez, M. K. Johnson, M. Fontecave, and M. Atta.** 2003. MiaB protein from *Thermotoga maritima*. Characterization of an extremely thermophilic tRNA-methylthiotransferase. J Biol Chem **278:**29515-24.

152. **Portalier, R., J. Robert-Baudouy, and F. Stoeber.** 1980. Regulation of *Escherichia coli* K-12 hexuronate system genes: *exu* regulon. J Bacteriol **143:**1095-107.

153. **Pourcel, C., G. Salvignol, and G. Vergnaud.** 2005. CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. Microbiology **151:**653-63.

154. **Pridmore, R. D., B. Berger, F. Desiere, D. Vilanova, C. Barretto, A. C. Pittet, M. C. Zwahlen, M. Rouvet, E. Altermann, R. Barrangou, B. Mollet, A. Mercenier, T. Klaenhammer, F. Arigoni, and M. A. Schell.** 2004. The genome sequence of the probiotic intestinal bacterium *Lactobacillus johnsonii* NCC 533. Proc Natl Acad Sci U S A **101:**2512-7.

155. **Pugsley, A. P., and C. Dubreuil.** 1988. Molecular characterization of *malQ*, the structural gene for the *Escherichia coli* enzyme amylomaltase. Mol Microbiol **2:**473-9.

156. **Pujic, P., R. Dervyn, A. Sorokin, and S. D. Ehrlich.** 1998. The *kdgRKAT* operon of *Bacillus subtilis*: detection of the transcript and regulation by the *kdgR* and *ccpA* genes. Microbiology **144 (Pt 11):**3111-8.

157. **Putnam, C. D., S. B. Clancy, H. Tsuruta, S. Gonzalez, J. G. Wetmur, and J. A. Tainer.** 2001. Structure and mechanism of the RuvB Holliday junction branch migration motor. J Mol Biol **311:**297-310.

158. **Pysz, M. A., S. B. Conners, C. I. Montero, K. R. Shockley, M. R. Johnson, D. E. Ward, and R. M. Kelly.** 2004. Transcriptional analysis of biofilm formation processes in the anaerobic, hyperthermophilic bacterium *Thermotoga maritima*. Appl Environ Microbiol **70:**6098-112.

159. **Pysz, M. A., D. E. Ward, K. R. Shockley, C. I. Montero, S. B. Conners, M. R. Johnson, and R. M. Kelly.** 2004. Transcriptional analysis of dynamic heat-shock

response by the hyperthermophilic bacterium *Thermotoga maritima*. Extremophiles **8:**209-17.

160. **Quezada, C. M., C. Gradinaru, M. I. Simon, A. M. Bilwes, and B. R. Crane.** 2004. Helical shifts generate two distinct conformers in the atomic resolution structure of the CheA phosphotransferase domain from *Thermotoga maritima*. J Mol Biol **341:**1283-94.

161. **Raasch, C., M. Armbrecht, W. Streit, B. Hocker, N. Strater, and W. Liebl.** 2002. Identification of residues important for NAD+ binding by the *Thermotoga maritima* alpha-glucosidase AglA, a member of glycoside hydrolase family 4. FEBS Lett **517:**267-71.

162. **Raasch, C., W. Streit, J. Schanzer, M. Bibel, U. Gosslar, and W. Liebl.** 2000. *Thermotoga maritima* AglA, an extremely thermostable NAD+-, Mn2+-, and thiol-dependent alpha-glucosidase. Extremophiles **4:**189-200.

163. **Rajashekhara, E., M. Kitaoka, Y. K. Kim, and K. Hayashi.** 2002. Characterization of a cellobiose phosphorylase from a hyperthermophilic eubacterium, *Thermotoga maritima* MSB8. Biosci Biotechnol Biochem **66:**2578-86.

164. **Ravot, G., M. Magot, M. L. Fardeau, B. K. Patel, G. Prensier, A. Egan, J. L. Garcia, and B. Ollivier.** 1995. *Thermotoga elfii* sp. nov., a novel thermophilic bacterium from an African oil-producing well. Int J Syst Bacteriol **45:**308-14.

165. **Reidl, J., K. Romisch, M. Ehrmann, and W. Boos.** 1989. MalI, a novel protein involved in regulation of the maltose system of *Escherichia coli*, is highly

homologous to the repressor proteins GalR, CytR, and LacI. J Bacteriol **171:**4888-99.

166. **Reverchon, S., W. Nasser, and J. Robert-Baudouy.** 1991. Characterization of *kdgR*, a gene of *Erwinia chrysanthemi* that regulates pectin degradation. Mol Microbiol **5:**2203-16.

167. **Richardson, J. S., M. F. Hynes, and I. J. Oresnik.** 2004. A genetic locus necessary for rhamnose uptake and catabolism in *Rhizobium leguminosarum* bv. trifolii. J Bacteriol **186:**8433-42.

168. **Rife, C., R. Schwarzenbacher, D. McMullan, P. Abdubek, E. Ambing, H. Axelrod, T. Biorac, J. M. Canaves, H. J. Chiu, A. M. Deacon, M. DiDonato, M. A. Elsliger, A. Godzik, C. Grittini, S. K. Grzechnik, J. Hale, E. Hampton, G. W. Han, J. Haugen, M. Hornsby, L. Jaroszewski, H. E. Klock, E. Koesema, A. Kreusch, P. Kuhn, S. A. Lesley, M. D. Miller, K. Moy, E. Nigoghossian, J. Paulsen, K. Quijano, R. Reyes, E. Sims, G. Spraggon, R. C. Stevens, H. van den Bedem, J. Velasquez, J. Vincent, A. White, G. Wolf, Q. Xu, K. O. Hodgson, J. Wooley, and I. A. Wilson.** 2005. Crystal structure of a putative modulator of DNA gyrase (pmbA) from *Thermotoga maritima* at 1.95 A resolution reveals a new fold. Proteins **61:**444-8.

169. **Rivolta, C., B. Soldo, V. Lazarevic, B. Joris, C. Mauel, and D. Karamata.** 1998. A 35.7 kb DNA fragment from the *Bacillus subtilis* chromosome containing a putative 12.3 kb operon involved in hexuronate catabolism and a perfectly symmetrical hypothetical catabolite-responsive element. Microbiology **144 (Pt 4):**877-84.

170. **Robinson, V. L., T. Wu, and A. M. Stock.** 2003. Structural analysis of the domain interface in DrrB, a response regulator of the OmpR/PhoB subfamily. J Bacteriol **185:**4186-94.

171. **Robinson-Rechavi, M., and A. Godzik.** 2005. Structural genomics of *Thermotoga maritima* proteins shows that contact order is a major determinant of protein thermostability. Structure (Camb) **13:**857-60.

172. **Roujeinikova, A., C. Raasch, J. Burke, P. J. Baker, W. Liebl, and D. W. Rice.** 2001. The crystal structure of *Thermotoga maritima* maltosyltransferase and its implications for the molecular basis of the novel transfer specificity. J Mol Biol **312:**119-31.

173. **Roujeinikova, A., C. Raasch, S. Sedelnikova, W. Liebl, and D. W. Rice.** 2002. Crystal structure of *Thermotoga maritima* 4-alpha-glucanotransferase and its acarbose complex: implications for substrate specificity and catalysis. J Mol Biol **321:**149-62.

174. **Roujeinikova, A., C. Raasch, S. Sedelnikova, W. Liebl, and D. W. Rice.** 2001. Crystallization and preliminary X-ray crystallographic studies on 4-alpha-glucanotransferase from *Thermotoga maritima*. Acta Crystallogr D Biol Crystallogr **57:**1046-7.

175. **Rubach, J. K., X. Brazzolotto, J. Gaillard, and M. Fontecave.** 2005. Biochemical characterization of the HydE and HydG iron-only hydrogenase maturation enzymes from *Thermatoga maritima*. FEBS Lett **579:**5055-60.

176. **Ruttersmith, L. D., and R. M. Daniel.** 1993. Thermostable beta-glucosidase and beta-xylosidase from *Thermotoga* sp. strain FjSS3-B.1. Biochim Biophys Acta **1156:**167-72.

177. **Ryjenkov, D. A., M. Tarutina, O. V. Moskvin, and M. Gomelsky.** 2005. Cyclic diguanylate is a ubiquitous signaling molecule in bacteria: insights into biochemistry of the GGDEF protein domain. J Bacteriol **187:**1792-8.

178. **Saha, B. C.** 2004. Purification and characterization of a novel mannitol dehydrogenase from *Lactobacillus intermedius*. Biotechnol Prog **20:**537-42.

179. **Santelli, E., R. Schwarzenbacher, D. McMullan, T. Biorac, L. S. Brinen, J. M. Canaves, J. Cambell, X. Dai, A. M. Deacon, M. A. Elsliger, S. Eshagi, R. Floyd, A. Godzik, C. Grittini, S. K. Grzechnik, L. Jaroszewski, C. Karlak, H. E. Klock, E. Koesema, J. S. Kovarik, A. Kreusch, P. Kuhn, S. A. Lesley, T. M. McPhillips, M. D. Miller, A. Morse, K. Moy, J. Ouyang, R. Page, K. Quijano, F. Rezezadeh, A. Robb, E. Sims, G. Spraggon, R. C. Stevens, H. van den Bedem, J. Velasquez, J. Vincent, F. von Delft, X. Wang, B. West, G. Wolf, Q. Xu, K. O. Hodgson, J. Wooley, and I. A. Wilson.** 2004. Crystal structure of a glycerophosphodiester phosphodiesterase (GDPD) from *Thermotoga maritima* (TM1621) at 1.60 A resolution. Proteins **56:**167-70.

180. **Sasaki, Y., M. Laivenieks, and J. G. Zeikus.** 2005. *Lactobacillus reuteri* ATCC 53608 *mdh* gene cloning and recombinant mannitol dehydrogenase characterization. Appl Microbiol Biotechnol **68:**36-41.

181. **Saul, D. J., L. C. Williams, R. A. Reeves, M. D. Gibbs, and P. L. Bergquist.** 1995. Sequence and expression of a xylanase gene from the hyperthermophile

Thermotoga sp. strain FjSS3-B.1 and characterization of the recombinant enzyme and its activity on kraft pulp. Appl Environ Microbiol **61:**4110-3.

182. **Savchenko, A., T. Skarina, E. Evdokimova, J. D. Watson, R. Laskowski, C. H. Arrowsmith, A. M. Edwards, A. Joachimiak, and R. G. Zhang.** 2004. X-ray crystal structure of CutA from *Thermotoga maritima* at 1.4 A resolution. Proteins **54:**162-5.

183. **Schlosser, A., T. Kampers, and H. Schrempf.** 1997. The *Streptomyces* ATP-binding component MsiK assists in cellobiose and maltose transport. J Bacteriol **179:**2092-5.

184. **Schneider, E.** 2001. ABC transporters catalyzing carbohydrate uptake. Res Microbiol **152:**303-10.

185. **Schouls, L. M., S. Reulen, B. Duim, J. A. Wagenaar, R. J. Willems, K. E. Dingle, F. M. Colles, and J. D. Van Embden.** 2003. Comparative genotyping of *Campylobacter jejuni* by amplified fragment length polymorphism, multilocus sequence typing, and short repeat sequencing: strain diversity, host range, and recombination. J Clin Microbiol **41:**15-26.

186. **Schulze-Gahmen, U., J. Pelaschier, H. Yokota, R. Kim, and S. H. Kim.** 2003. Crystal structure of a hypothetical protein, TM841 of *Thermotoga maritima*, reveals its function as a fatty acid-binding protein. Proteins **50:**526-30.

187. **Schumann, J., A. Wrba, R. Jaenicke, and K. O. Stetter.** 1991. Topographical and enzymatic characterization of amylases from the extremely thermophilic eubacterium *Thermotoga maritima*. FEBS Lett **282:**122-6.

188. **Schwarzenbacher, R., J. M. Canaves, L. S. Brinen, X. Dai, A. M. Deacon, M. A. Elsliger, S. Eshaghi, R. Floyd, A. Godzik, C. Grittini, S. K. Grzechnik, C. Guda, L. Jaroszewski, C. Karlak, H. E. Klock, E. Koesema, J. S. Kovarik, A. Kreusch, P. Kuhn, S. A. Lesley, D. McMullan, T. M. McPhillips, M. A. Miller, M. D. Miller, A. Morse, K. Moy, J. Ouyang, A. Robb, K. Rodrigues, T. L. Selby, G. Spraggon, R. C. Stevens, H. van den Bedem, J. Velasquez, J. Vincent, X. Wang, B. West, G. Wolf, K. O. Hodgson, J. Wooley, and I. A. Wilson.** 2003. Crystal structure of uronate isomerase (TM0064) from *Thermotoga maritima* at 2.85 A resolution. Proteins **53:**142-5.

189. **Schwarzenbacher, R., A. M. Deacon, L. Jaroszewski, L. S. Brinen, J. M. Canaves, X. Dai, M. A. Elsliger, R. Floyd, A. Godzik, C. Grittini, S. K. Grzechnik, H. E. Klock, E. Koesema, J. S. Kovarik, A. Kreusch, P. Kuhn, S. A. Lesley, D. McMullan, T. M. McPhillips, M. D. Miller, A. Morse, K. Moy, M. S. Nelson, J. Ouyang, R. Page, A. Robb, K. Quijano, G. Spraggon, R. C. Stevens, H. van den Bedem, J. Velasquez, J. Vincent, F. von Delft, X. Wang, B. West, G. Wolf, K. O. Hodgson, J. Wooley, and I. A. Wilson.** 2004. Crystal structure of a putative glutamine amido transferase (TM1158) from *Thermotoga maritima* at 1.7 A resolution. Proteins **54:**801-5.

190. **Schwarzenbacher, R., L. Jaroszewski, F. von Delft, P. Abdubek, E. Ambing, T. Biorac, L. S. Brinen, J. M. Canaves, J. Cambell, H. J. Chiu, X. Dai, A. M. Deacon, M. DiDonato, M. A. Elsliger, S. Eshagi, R. Floyd, A. Godzik, C. Grittini, S. K. Grzechnik, E. Hampton, C. Karlak, H. E. Klock, E. Koesema, J. S. Kovarik, A. Kreusch, P. Kuhn, S. A. Lesley, I. Levin, D. McMullan, T.**

M. McPhillips, M. D. Miller, A. Morse, K. Moy, J. Ouyang, R. Page, K. Quijano, A. Robb, G. Spraggon, R. C. Stevens, H. van den Bedem, J. Velasquez, J. Vincent, X. Wang, B. West, G. Wolf, Q. Xu, K. O. Hodgson, J. Wooley, and I. A. Wilson. 2004. Crystal structure of a phosphoribosylaminoimidazole mutase PurE (TM0446) from *Thermotoga maritima* at 1.77-A resolution. Proteins **55:**474-8.

191. **Schwarzenbacher, R., L. Jaroszewski, F. von Delft, P. Abdubek, E. Ambing, T. Biorac, L. S. Brinen, J. M. Canaves, J. Cambell, H. J. Chiu, X. Dai, A. M. Deacon, M. DiDonato, M. A. Elsliger, S. Eshagi, R. Floyd, A. Godzik, C. Grittini, S. K. Grzechnik, E. Hampton, C. Karlak, H. E. Klock, E. Koesema, J. S. Kovarik, A. Kreusch, P. Kuhn, S. A. Lesley, I. Levin, D. McMullan, T. M. McPhillips, M. D. Miller, A. Morse, K. Moy, J. Ouyang, R. Page, K. Quijano, A. Robb, G. Spraggon, R. C. Stevens, H. van den Bedem, J. Velasquez, J. Vincent, X. Wang, B. West, G. Wolf, Q. Xu, K. O. Hodgson, J. Wooley, and I. A. Wilson.** 2004. Crystal structure of a type II quinolic acid phosphoribosyltransferase (TM1645) from *Thermotoga maritima* at 2.50 A resolution. Proteins **55:**768-71.

192. **Schwarzenbacher, R., L. Jaroszewski, F. von Delft, P. Abdubek, E. Ambing, T. Biorac, L. S. Brinen, J. M. Canaves, J. Cambell, H. J. Chiu, X. Dai, A. M. Deacon, M. DiDonato, M. A. Elsliger, S. Eshagi, R. Floyd, A. Godzik, C. Grittini, S. K. Grzechnik, E. Hampton, C. Karlak, H. E. Klock, E. Koesema, J. S. Kovarik, A. Kreusch, P. Kuhn, S. A. Lesley, I. Levin, D. McMullan, T. M. McPhillips, M. D. Miller, A. Morse, K. Moy, J. Ouyang, R. Page, K.**

Quijano, A. Robb, G. Spraggon, R. C. Stevens, H. van den Bedem, J. Velasquez, J. Vincent, X. Wang, B. West, G. Wolf, Q. Xu, K. O. Hodgson, J. Wooley, and I. A. Wilson. 2004. Crystal structure of an aspartate aminotransferase (TM1255) from *Thermotoga maritima* at 1.90 A resolution. Proteins **55:**759-63.

193. **Schwarzenbacher, R., F. von Delft, P. Abdubek, E. Ambing, T. Biorac, L. S. Brinen, J. M. Canaves, J. Cambell, H. J. Chiu, X. Dai, A. M. Deacon, M. DiDonato, M. A. Elsliger, S. Eshagi, R. Floyd, A. Godzik, C. Grittini, S. K. Grzechnik, E. Hampton, L. Jaroszewski, C. Karlak, H. E. Klock, E. Koesema, J. S. Kovarik, A. Kreusch, P. Kuhn, S. A. Lesley, I. Levin, D. McMullan, T. M. McPhillips, M. D. Miller, A. Morse, K. Moy, J. Ouyang, R. Page, K. Quijano, A. Robb, G. Spraggon, R. C. Stevens, H. van den Bedem, J. Velasquez, J. Vincent, X. Wang, B. West, G. Wolf, Q. Xu, K. O. Hodgson, J. Wooley, and I. A. Wilson.** 2004. Crystal structure of a putative PII-like signaling protein (TM0021) from *Thermotoga maritima* at 2.5 A resolution. Proteins **54:**810-3.

194. **Schwarzenbacher, R., F. von Delft, J. M. Canaves, L. S. Brinen, X. Dai, A. M. Deacon, M. A. Elsliger, S. Eshaghi, R. Floyd, A. Godzik, C. Grittini, S. K. Grzechnik, C. Guda, L. Jaroszewski, C. Karlak, H. E. Klock, E. Koesema, J. S. Kovarik, A. Kreusch, P. Kuhn, S. A. Lesley, D. McMullan, T. M. McPhillips, M. A. Miller, M. D. Miller, A. Morse, K. Moy, J. Ouyang, R. Page, A. Robb, K. Rodrigues, T. L. Selby, G. Spraggon, R. C. Stevens, H. van den Bedem, J. Velasquez, J. Vincent, X. Wang, B. West, G. Wolf, K. O.**

Hodgson, J. Wooley, and I. A. Wilson. 2004. Crystal structure of an iron-containing 1,3-propanediol dehydrogenase (TM0920) from *Thermotoga maritima* at 1.3 A resolution. Proteins **54:**174-7.

195. **Schwarzenbacher, R., F. von Delft, L. Jaroszewski, P. Abdubek, E. Ambing, T. Biorac, L. S. Brinen, J. M. Canaves, J. Cambell, H. J. Chiu, X. Dai, A. M. Deacon, M. DiDonato, M. A. Elsliger, S. Eshagi, R. Floyd, A. Godzik, C. Grittini, S. K. Grzechnik, E. Hampton, C. Karlak, H. E. Klock, E. Koesema, J. S. Kovarik, A. Kreusch, P. Kuhn, S. A. Lesley, I. Levin, D. McMullan, T. M. McPhillips, M. D. Miller, A. Morse, K. Moy, J. Ouyang, R. Page, K. Quijano, A. Robb, G. Spraggon, R. C. Stevens, H. van den Bedem, J. Velasquez, J. Vincent, X. Wang, B. West, G. Wolf, Q. Xu, K. O. Hodgson, J. Wooley, and I. A. Wilson.** 2004. Crystal structure of a putative oxalate decarboxylase (TM1287) from *Thermotoga maritima* at 1.95 A resolution. Proteins **56:**392-5.

196. **Selig, M., K. B. Xavier, H. Santos, and P. Schonheit.** 1997. Comparative analysis of Embden-Meyerhof and Entner-Doudoroff glycolytic pathways in hyperthermophilic archaea and the bacterium *Thermotoga*. Arch Microbiol **167:**217-32.

197. **Selmer, M., S. Al-Karadaghi, G. Hirokawa, A. Kaji, and A. Liljas.** 1999. Crystal structure of *Thermotoga maritima* ribosome recycling factor: a tRNA mimic. Science **286:**2349-52.

198. **Selmer, M., S. Al-Karadaghi, G. Hirokawa, A. Kaji, and A. Liljas.** 1999. Crystallization and preliminary X-ray analysis of *Thermotoga maritima* ribosome recycling factor. Acta Crystallogr D Biol Crystallogr **55 (Pt 12):**2049-50.

199. **Shin, D. H., J. Brandsen, J. Jancarik, H. Yokota, R. Kim, and S. H. Kim.** 2004. Structural analyses of peptide release factor 1 from *Thermotoga maritima* reveal domain flexibility required for its interaction with the ribosome. J Mol Biol **341:**227-39.

200. **Shin, D. H., Y. Lou, J. Jancarik, H. Yokota, R. Kim, and S. H. Kim.** 2004. Crystal structure of YjeQ from *Thermotoga maritima* contains a circularly permuted GTPase domain. Proc Natl Acad Sci U S A **101:**13198-203.

201. **Shin, D. H., H. H. Nguyen, J. Jancarik, H. Yokota, R. Kim, and S. H. Kim.** 2003. Crystal structure of NusA from *Thermotoga maritima* and functional implication of the N-terminal domain. Biochemistry **42:**13429-37.

202. **Shin, D. H., A. Roberts, J. Jancarik, H. Yokota, R. Kim, D. E. Wemmer, and S. H. Kim.** 2003. Crystal structure of a phosphatase with a unique substrate binding domain from *Thermotoga maritima*. Protein Sci **12:**1464-72.

203. **Shockley, K. R., K. L. Scott, M. A. Pysz, S. B. Conners, M. R. Johnson, C. I. Montero, R. D. Wolfinger, and R. M. Kelly.** 2005. Genome-wide transcriptional variation within and between steady states for continuous growth of the hyperthermophile *Thermotoga maritima*. Appl Environ Microbiol **71:**5572-6.

204. **Shumilin, I. A., R. Bauerle, J. Wu, R. W. Woodard, and R. H. Kretsinger.** 2004. Crystal structure of the reaction complex of 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase from *Thermotoga maritima* refines the

catalytic mechanism and indicates a new mechanism of allosteric regulation. J Mol Biol **341:**455-66.

205. **Simpson, H. D., U. R. Haufler, and R. M. Daniel.** 1991. An extremely thermostable xylanase from the thermophilic eubacterium *Thermotoga*. Biochem J **277 (Pt 2):**413-7.

206. **Skinner, M. M., J. M. Puvathingal, R. L. Walter, and A. M. Friedman.** 2000. Crystal structure of protein isoaspartyl methyltransferase: a catalyst for protein repair. Structure Fold Des **8:**1189-201.

207. **Snel, B., G. Lehmann, P. Bork, and M. A. Huynen.** 2000. STRING: a web-server to retrieve and display the repeatedly occurring neighbourhood of a gene. Nucleic Acids Res **28:**3442-4.

208. **Song, H. K., M. Bochtler, M. K. Azim, C. Hartmann, R. Huber, and R. Ramachandran.** 2003. Isolation and characterization of the prokaryotic proteasome homolog HslVU (ClpQY) from *Thermotoga maritima* and the crystal structure of HslV. Biophys Chem **100:**437-52.

209. **Spraggon, G., D. Pantazatos, H. E. Klock, I. A. Wilson, V. L. Woods, Jr., and S. A. Lesley.** 2004. On the use of DXMS to produce more crystallizable proteins: structures of the *T. maritima* proteins TM0160 and TM1171. Protein Sci **13:**3187-99.

210. **Spraggon, G., R. Schwarzenbacher, A. Kreusch, C. C. Lee, P. Abdubek, E. Ambing, T. Biorac, L. S. Brinen, J. M. Canaves, J. Cambell, H. J. Chiu, X. Dai, A. M. Deacon, M. DiDonato, M. A. Elsliger, S. Eshagi, R. Floyd, A. Godzik, C. Grittini, S. K. Grzechnik, E. Hampton, L. Jaroszewski, C.**

Karlak, H. E. Klock, E. Koesema, J. S. Kovarik, P. Kuhn, I. Levin, D. McMullan, T. M. McPhillips, M. D. Miller, A. Morse, K. Moy, J. Ouyang, R. Page, K. Quijano, A. Robb, R. C. Stevens, H. van den Bedem, J. Velasquez, J. Vincent, F. von Delft, X. Wang, B. West, G. Wolf, Q. Xu, K. O. Hodgson, J. Wooley, S. A. Lesley, and I. A. Wilson. 2004. Crystal structure of an Udp-n-acetylmuramate-alanine ligase MurC (TM0231) from *Thermotoga maritima* at 2.3 A resolution. Proteins **55:**1078-81.

211. **Spraggon, G., R. Schwarzenbacher, A. Kreusch, D. McMullan, L. S. Brinen, J. M. Canaves, X. Dai, A. M. Deacon, M. A. Elsliger, S. Eshagi, R. Floyd, A. Godzik, C. Grittini, S. K. Grzechnik, L. Jaroszewski, C. Karlak, H. E. Klock, E. Koesema, J. S. Kovarik, P. Kuhn, T. M. McPhillips, M. D. Miller, A. Morse, K. Moy, J. Ouyang, R. Page, K. Quijano, F. Rezezadeh, A. Robb, E. Sims, R. C. Stevens, H. van den Bedem, J. Velasquez, J. Vincent, F. von Delft, X. Wang, B. West, G. Wolf, Q. Xu, K. O. Hodgson, J. Wooley, S. A. Lesley, and I. A. Wilson.** 2004. Crystal structure of a methionine aminopeptidase (TM1478) from *Thermotoga maritima* at 1.9 A resolution. Proteins **56:**396-400.

212. **Stoodley, P., S. Wilson, L. Hall-Stoodley, J. D. Boyle, H. M. Lappin-Scott, and J. W. Costerton.** 2001. Growth and detachment of cell clusters from mature mixed-species biofilms. Appl Environ Microbiol **67:**5608-13.

213. **Stoop, J. M., and D. M. Pharr.** 1992. Partial purification and characterization of mannitol: mannose 1-oxidoreductase from celeriac (*Apium graveolens* var. rapaceum) roots. Arch Biochem Biophys **298:**612-9.

214. **Sulzenbacher, G., C. Bignon, T. Nishimura, C. A. Tarling, S. G. Withers, B. Henrissat, and Y. Bourne.** 2004. Crystal structure of *Thermotoga maritima* alpha-L-fucosidase. Insights into the catalytic mechanism and the molecular basis for fucosidosis. J Biol Chem **279:**13119-28.

215. **Suresh, C., M. Kitaoka, and K. Hayashi.** 2003. A thermostable non-xylanolytic alpha-glucuronidase of *Thermotoga maritima* MSB8. Biosci Biotechnol Biochem **67:**2359-64.

216. **Suresh, C., A. A. Rus'd, M. Kitaoka, and K. Hayashi.** 2002. Evidence that the putative alpha-glucosidase of *Thermotoga maritima* MSB8 is a pNP alpha-D-glucuronopyranoside hydrolyzing alpha-glucuronidase. FEBS Lett **517:**159-62.

217. **Susin, M. F., H. R. Perez, R. L. Baldini, and S. L. Gomes.** 2004. Functional and structural analysis of HrcA repressor protein from *Caulobacter crescentus*. J Bacteriol **186:**6759-67.

218. **Takahata, Y., M. Nishijima, T. Hoaki, and T. Maruyama.** 2001. *Thermotoga petrophila* sp. nov. and *Thermotoga naphthophila* sp. nov., two hyperthermophilic bacteria from the Kubiki oil reservoir in Niigata, Japan. Int J Syst Evol Microbiol **51:**1901-9.

219. **Tanaka, T., T. Fukui, S. Fujiwara, H. Atomi, and T. Imanaka.** 2004. Concerted action of diacetylchitobiose deacetylase and Exo-ss-D-glucosaminidase in a novel chitinolytic pathway in the hyperthermophilic archaeon *Thermococcus kodakaraensis* KOD1. J Biol Chem.

220. **Tarling, C. A., S. He, G. Sulzenbacher, C. Bignon, Y. Bourne, B. Henrissat, and S. G. Withers.** 2003. Identification of the catalytic nucleophile of the family

29 alpha-L-fucosidase from *Thermotoga maritima* through trapping of a covalent glycosyl-enzyme intermediate and mutagenesis. J Biol Chem **278:**47394-9.

221. **Torres-Larios, A., K. K. Swinger, A. S. Krasilnikov, T. Pan, and A. Mondragon.** 2005. Crystal structure of the RNA component of bacterial ribonuclease P. Nature **437:**584-7.

222. **Toth, E. A., and T. O. Yeates.** 2000. The structure of adenylosuccinate lyase, an enzyme with dual activity in the de novo purine biosynthetic pathway. Structure Fold Des **8:**163-74.

223. **van den Ent, F., and J. Lowe.** 2000. Crystal structure of the cell division protein FtsA from *Thermotoga maritima*. Embo J **19:**5300-7.

224. **Van Ooteghem, S. A., S. K. Beer, and P. C. Yue.** 2002. Hydrogen production by the thermophilic bacterium *Thermotoga neapolitana*. Appl Biochem Biotechnol **98-100:**177-89.

225. **Van Ooteghem, S. A., A. Jones, D. Van Der Lelie, B. Dong, and D. Mahajan.** 2004. H(2) production and carbon utilization by *Thermotoga neapolitana* under anaerobic and microaerobic growth conditions. Biotechnol Lett **26:**1223-32.

226. **Vargas, M., and K. M. Noll.** 1996. Catabolite repression in the hyperthermophilic bacterium *Thermotoga neapolitana* is independent of cAMP. Microbiology **142 (Pt 1):**139-44.

227. **Varrot, A., V. L. Yip, Y. Li, S. S. Rajan, X. Yang, W. F. Anderson, J. Thompson, S. G. Withers, and G. J. Davies.** 2005. NAD+ and metal-ion dependent hydrolysis by family 4 glycosidases: structural insight into specificity for phospho-beta-D-glucosides. J Mol Biol **346:**423-35.

228. **Velikodvorskaya, T. V., I. Volkov, V. T. Vasilevko, V. V. Zverlov, and E. S. Piruzian.** 1997. Purification and some properties of *Thermotoga neapolitana* thermostable xylanase B expressed in *E. coli* cells. Biochemistry (Mosc) **62:**66-70.

229. **Wagner, E., S. Marcandier, O. Egeter, J. Deutscher, F. Gotz, and R. Bruckner.** 1995. Glucose kinase-dependent catabolite repression in *Staphylococcus xylosus*. J Bacteriol **177:**6144-52.

230. **Wahl, M. C., G. P. Bourenkov, H. D. Bartunik, and R. Huber.** 2000. Flexibility, conformational diversity and two dimerization modes in complexes of ribosomal protein L12. Embo J **19:**174-86.

231. **Wahl, M. C., R. Huber, S. Marinkovic, E. Weyher-Stingl, and S. Ehlert.** 2000. Structural investigations of the highly flexible recombinant ribosomal protein L12 from *Thermotoga maritima*. Biol Chem **381:**221-9.

232. **Wassenberg, D., W. Liebl, and R. Jaenicke.** 2000. Maltose-binding protein from the hyperthermophilic bacterium *Thermotoga maritima*: stability and binding properties. J Mol Biol **295:**279-88.

233. **Wassenberg, D., H. Schurig, W. Liebl, and R. Jaenicke.** 1997. Xylanase XynA from the hyperthermophilic bacterium *Thermotoga maritima*: structure and stability of the recombinant enzyme and its isolated cellulose-binding domain. Protein Sci **6:**1718-26.

234. **Watanabe, S., M. Hamano, H. Kakeshita, K. Bunai, S. Tojo, H. Yamaguchi, Y. Fujita, S. L. Wong, and K. Yamane.** 2003. Mannitol-1-phosphate dehydrogenase (MtlD) is required for mannitol and glucitol assimilation in

Bacillus subtilis: possible cooperation of *mtl* and *gut* operons. J Bacteriol **185:**4816-24.

235. **Wiegert, T., K. Hagmaier, and W. Schumann.** 2004. Analysis of orthologous *hrcA* genes in *Escherichia coli* and *Bacillus subtilis*. FEMS Microbiol Lett **234:**9-17.

236. **Winterhalter, C., P. Heinrich, A. Candussio, G. Wich, and W. Liebl.** 1995. Identification of a novel cellulose-binding domain within the multidomain 120 kDa xylanase XynA of the hyperthermophilic bacterium *Thermotoga maritima*. Mol Microbiol **15:**431-44.

237. **Woodward, J., N. I. Heyer, J. P. Getty, H. M. O'Neill, E. Pinkhassik, and B. R. Evans.** 2002. Presented at the Proceedings of the 2002 U.S. DOE Hydrogen Program Review.

238. **Worbs, M., R. Huber, and M. C. Wahl.** 2000. Crystal structure of ribosomal protein L4 shows RNA-binding sites for ribosome incorporation and feedback control of the S10 operon. Embo J **19:**807-18.

239. **Worning, P., L. J. Jensen, K. E. Nelson, S. Brunak, and D. W. Ussery.** 2000. Structural analysis of DNA sequence: evidence for lateral gene transfer in *Thermotoga maritima*. Nucleic Acids Res **28:**706-9.

240. **Wu, S. P., S. S. Mansy, and J. A. Cowan.** 2005. Iron-sulfur cluster biosynthesis. Molecular chaperone DnaK promotes IscU-bound [2Fe-2S] cluster stability and inhibits cluster transfer activity. Biochemistry **44:**4284-93.

241. **Xu, Q., R. Schwarzenbacher, D. McMullan, P. Abdubek, E. Ambing, T. Biorac, J. M. Canaves, H. J. Chiu, X. Dai, A. M. Deacon, M. DiDonato, M. A.**

Elsliger, A. Godzik, C. Grittini, S. K. Grzechnik, E. Hampton, M. Hornsby, L. Jaroszewski, H. E. Klock, E. Koesema, A. Kreusch, P. Kuhn, S. A. Lesley, I. Levin, M. D. Miller, A. Morse, K. Moy, J. Ouyang, R. Page, K. Quijano, R. Reyes, A. Robb, E. Sims, G. Spraggon, R. C. Stevens, H. van den Bedem, J. Velasquez, J. Vincent, F. von Delft, X. Wang, B. West, A. White, G. Wolf, O. Zagnitko, K. O. Hodgson, J. Wooley, and I. A. Wilson. 2005. Crystal structure of a formiminotetrahydrofolate cyclodeaminase (TM1560) from *Thermotoga maritima* at 2.80 A resolution reveals a new fold. Proteins **58:**976-81.

242. **Xu, Q., R. Schwarzenbacher, D. McMullan, F. von Delft, L. S. Brinen, J. M. Canaves, X. Dai, A. M. Deacon, M. A. Elsliger, S. Eshagi, R. Floyd, A. Godzik, C. Grittini, S. K. Grzechnik, L. Jaroszewski, C. Karlak, H. E. Klock, E. Koesema, J. S. Kovarik, A. Kreusch, P. Kuhn, S. A. Lesley, I. Levin, T. M. McPhillips, M. D. Miller, A. Morse, K. Moy, J. Ouyang, R. Page, K. Quijano, A. Robb, G. Spraggon, R. C. Stevens, H. van den Bedem, J. Velasquez, J. Vincent, X. Wang, B. West, G. Wolf, K. O. Hodgson, J. Wooley, and I. A. Wilson.** 2004. Crystal structure of a ribose-5-phosphate isomerase RpiB (TM1080) from *Thermotoga maritima* at 1.90 A resolution. Proteins **56:**171-5.

243. **Xue, Y., and W. Shao.** 2004. Expression and characterization of a thermostable beta-xylosidase from the hyperthermophile, *Thermotoga maritima*. Biotechnol Lett **26:**1511-5.

244. **Yang, Z., A. Savchenko, A. Yakunin, R. Zhang, A. Edwards, C. Arrowsmith, and L. Tong.** 2003. Aspartate dehydrogenase, a novel enzyme identified from structural and functional studies of TM1643. J Biol Chem **278:**8804-8.

245. **Yernool, D. A., J. K. McCarthy, D. E. Eveleigh, and J. D. Bok.** 2000. Cloning and characterization of the glucooligosaccharide catabolic pathway beta-glucan glucohydrolase and cellobiose phosphorylase in the marine hyperthermophile *Thermotoga neapolitana*. Journal of Bacteriology **182:**5172-5179.

246. **Yip, V. L., A. Varrot, G. J. Davies, S. S. Rajan, X. Yang, J. Thompson, W. F. Anderson, and S. G. Withers.** 2004. An unusual mechanism of glycoside hydrolysis involving redox and elimination steps by a family 4 beta-glycosidase from *Thermotoga maritima*. J Am Chem Soc **126:**8354-5.

247. **Yu, J. S., M. Vargas, C. Mityas, and K. M. Noll.** 2001. Liposome-mediated DNA uptake and transient expression in *Thermotoga*. Extremophiles **5:**53-60.

248. **Zhang, J., Y. Zhang, and M. Inouye.** 2003. *Thermotoga maritima* MazG protein has both nucleoside triphosphate pyrophosphohydrolase and pyrophosphatase activities. J Biol Chem **278:**21408-14.

249. **Zhang, R. G., Y. Kim, T. Skarina, S. Beasley, R. Laskowski, C. Arrowsmith, A. Edwards, A. Joachimiak, and A. Savchenko.** 2002. Crystal structure of *Thermotoga maritima* 0065, a member of the IclR transcriptional factor family. J Biol Chem **277:**19183-90.

250. **Zuber, U., and W. Schumann.** 1994. CIRCE, a novel heat shock element involved in regulation of heat shock operon *dnaK* of *Bacillus subtilis*. J Bacteriol **176:**1359-63.

251. **Zverlov, V., K. Piotukh, O. Dakhova, G. Velikodvorskaya, and R. Borriss.** 1996. The multidomain xylanase A of the hyperthermophilic bacterium

Thermotoga neapolitana is extremely thermoresistant. Appl Microbiol Biotechnol **45:**245-7.

252. **Zverlov, V. V., I. Y. Volkov, G. A. Velikodvorskaya, and W. H. Schwarz.** 2001. The binding pattern of two carbohydrate-binding modules of laminarinase Lam16A from *Thermotoga neapolitana*: differences in beta-glucan binding within family CBM4. Microbiology **147:**621-9.

253. **Zverlov, V. V., I. Y. Volkov, T. V. Velikodvorskaya, and W. H. Schwarz.** 1997. Highly thermostable endo-1,3-beta-glucanase (laminarinase) LamA from *Thermotoga neapolitana*: nucleotide sequence of the gene and characterization of the recombinant gene product. Microbiology **143 (Pt 5):**1701-8.

254. **Zverlov, V. V., I. Y. Volkov, T. V. Velikodvorskaya, and W. H. Schwarz.** 1997. *Thermotoga neapolitana bglB* gene, upstream of *lamA*, encodes a highly thermostable beta-glucosidase that is a laminaribiase. Microbiology **143 (Pt 11):**3537-42.

TABLE 1.1. Expression-based functional genomics studies completed or in progress for *T. maritima*

Experiment	Size of cDNA microarray	Growth conditions examined	Reference(s)
Carbohydrate utilization	Targeted	Batch growth, ten carbohydrates (barley glucan, carboxymethyl cellulose, galactomannan, glucose, glucomannan, laminarin, mannose, starch, β-xylan, xylose)	(21)
Heat shock	Targeted	Batch growth, at baseline and five subsequent time points (0, 5, 30, 60, 90)	(159)
Biofilm growth	Full genome	Continuous culture in a chemostat, biofilm and planktonic cells	(158)
Co-culture, with *M. jannaschii*	Full genome	Batch growth, high-density co-culture (*T. maritima* and *M. jannaschii*), and pure *T. maritima* culture	(71)
Carbohydrate utilization	Full genome	Continuous culture in a chemostat, glucose-, maltose-, and lactose-grown cells	(136)
Biological variability in a steady state	Full genome	Continuous culture in a chemostat, examined variability in repeated samples taken at two temperature (80 and 85 degrees), and two dilution rates ($D=0.17\ h^{-1}$, $D=0.25h^{-1}$)	(203)
Carbohydrate specificity of ABC transporters	Full genome	Batch growth, fourteen carbohydrates (barley glucan, carboxymethyl cellulose, galactomannan, glucose, glucomannan, laminarin, mannose, starch, β-xylan, xylose)	(27)
Growth phase variation	Full genome	Batch growth, different growth phases (mid-log, early stationary, late stationary) for pure *T. maritima* and cells co-cultured with *M. jannaschii*	Johnson et. al, in press.
Glucan disaccharides	Full genome	Batch growth, maltose and cellobiose, with and without sulfur; Continuous culture in a chemostat, cellobiose, maltose, and cellobiose-maltose mixture	Unpublished data
Heat shock	Full genome	Batch growth, at baseline and five subsequent time points (0, 5, 30, 60, 90)	Unpublished data
Antibiotic resistance	Full genome	Batch growth, chloramphenicol-resistant mutant during growth phases; continuous culture of mutant compared to wild type.	Unpublished data

TABLE 1.2. Selected *T. maritima* genes and proteins characterized by functional genomics or biochemisty since 1999.

Gene identifier	Gene name	Original annotation	Current annotation	Family	Crystal Structure	Reference(s)
TM0076	*xyl3*	xylosidase	β-xylosidase	GH3		(243)
TM0209	*pfp*	6-PFK phosphofructokinase	PPi-dependent 6-PFK			(30)
TM0281		α-L-arabinofuranosidase	α-L-arabinofuranosidase	GH51		(123)
TM0289	*pfk*		6-PFK phosphofructokinase, ATP-dependent			(30)
TM0306		α-L-fucosidase, putative	α-L-fucosidase	GH29		(214, 220)
TM0434	*agu4A*	α-glucosidase	α-glucuronidase	GH 4		(216)
TM0504		hypothetical protein	putative signaling peptide			(71)
TM0539		tryptophan synthase, beta subunit	Indole rescue protein			(54)
TM0653	*miaB*	conserved hypothetical protein	t-RNA methylthiotransferase			(151)
TM0752	*agu4B*	α -glucosidase	α -glucuronidase	GH 4	1VJTA	(215)
TM0841		S-layer like array protein	fatty-acid binding protein	DegV/ COG1307		(186)
TM0875		hypothetical protein	YggU-like protein	YggU-like	1O22	(3)
TM0913	*mazG*	mazG protein	Pyrophosphatase, nucleotide triphosphate pyrophosphorylase,	MazG		(248)
TM1068		α-glucosidase	α-glucuronidase (>99% id/466 aa with TM0434)	GH4		(216)
TM1067	*rtpA*	oligopeptide ABC transporter, periplasmic oligopeptide-binding protein	Rhamnose ABC transporter, periplasmic rhamnose binding protein, putative	Opp/Dpp		(27)
TM1192	*galA*		α-galactosidase			(103)
TM1193	*lacZ*	β-galactosidase	β-galactosidase	GH2		(75)
TM1223	*cbtA*	oligopeptide ABC transporter, periplasmic oligopeptide-binding protein	Cellobiose ABC transporter, periplasmic cellobiose-binding protein, putative	Opp/Dpp	1VR5	(21, 27)
TM1226	*mbtA*	oligopeptide ABC transporter, periplasmic oligopeptide-binding protein	Mannan ABC transporter, periplasmic mannan-binding protein, putative; possibly interacts with cellobiose ABC transporter subunits	Opp/Dpp		(27)

(Table 1.2, continued)

TM1267	*hydG*	thiH protein, putative	Putative hydrogenase maturation protein			(175)
TM1269	*hydE*	biotin synthetase, putative	Putative hydrogenase maturation protein			(175)
TM1281	*bglT*	6-P-β-glucosidase	6-P-β-glucosidase	GH 4		(227, 246)
TM1371	*iscU*					(9, 112)
TM1420		hypothetical protein	[2Fe-2S] protein, possibly involved in hydrogenase metal cluster assembly			(141)
TM1643		conserved hypothetical protein	aspartate dehydrogenase		1J5P, 1H2H	(244)
TM1667	*xylA*	xylose isomerase	xylose isomerase			(5)
TM1834	*aglA*	alpha-glucosidase	alpha-glucosidase			(162)
TM1848	*cepA*	cellobiose phosphorylase	cellobiose phosphorylase			(163)
TM1851		α-mannosidase	cytosolic α-mannosidase			(128)

* proposed designation

TABLE 1.3. ABC transport systems of *T. maritima*

Family and predicted substrate	Genes	Genomic location	Source(s) or reference(s)
Opp/Dpp ABC family			
Carbohydrates			
β-glucans	*bgtpABCDF* [a]	TM0027-TM0031	(27)
Xylan, xylose	*xtpHJLMG* [a]	TM0056-TM0060	(27)
Xylan, xylose	*xtpABCDF* [a]	TM0071-TM0075	(27),GI:23270356
Xylan	*xtpN* [a]	TM0309	(27)
Rhamnose	*rtpABCDF* [a]	TM1063-TM1067	(27)
Lactose	*ltpABCDF* [c]	TM1194, TM1196-9	(136)
Cellobiose, barley	*cbtABCDF* [a,d]	TM1219-TM1223	(21, 27)
β-mannans, mannose	*mbtA* [a]	TM1226	(27)
β-mannans	*mtpABCDF* [a,d]	TM1746-TM1750	(21, 27)
Quorum sensing			
Possible peptide export		TM0500-3	(71)
Unknown			
	dppABCDF [b]	TM0300-TM0304	Unknown
	dppA[b]	TM0460	Unknown
	dppFABC [b]	TM0530-TM0533	Unknown
	dppCDFAB [b]	TM1149-TM1153	Unknown
Sugar ABC family			
CUT1 subfamily			
Arabinose?		TM0277-9, TM0287-8	Possibly nonfunctional; frameshift binding protein
	malE3F3ugpE [e] *malK2*	TM0418-TM0421	Unknown
Uronic acid Polysaccharides	*ugpBAE*[e]	TM0430-TM0432	(27)
	ugpB, *ugpA*,,*ugpE* [e]	TM0595, TM0596, TM0598	Unknown
N-acetylglucosamine or Nag polysaccharides	*ugpBAE* [e]	TM0810-TM0812	(27)
Maltose, maltotriose, β-(1→4) mannotetraose	*malE1F1G1*[f]	TM1202-TM1204	(130)
	ugpBAE [e]	TM1232-TM1235	Unknown
Maltose	*malK1*[a]	TM1276	(27)
Maltose, maltotriose, Trehalose	*malE2F2* [f]	TM1836, TM1839	(130, 232)
	ugpBAE [e]	TM1853-TM1855	Unknown
CUT2 subfamily			
Monosaccharides?	*rbsC2A2B2*[a]	TM0112, TM0114, TM0115	(27)
Ribose, arabinose, xylose	*rbsDB1A1C1* [a]	TM0955-6, TM0958-9	(27)
Unknown		TM0102-TM0105	Unknown
Metal ABC family			
Cobalamin/Fe(III)-siderophores?	*fecBCD*	TM0078-80	Unknown

(Table 1.3, continued)

Zinc or manganese transporter?		TM0122	Unknown
Cobalamin/Fe(III)-siderophores?	*fecBCD*	TM0189-191	Unknown
Lipoprotein release family			
Unknown		TM0194 (ATPase)	Unknown
Unknown		TM0351-2	Unknown
Unknown		TM0705-6	Unknown
Sulfonate or taurine family			
Unknown	*tauABC*	TM0202-4	
Unknown	*tauABC*	TM0483-5	
Cobalt family			
Unknown	*cbiO*	TM0222	
Unknown	*cbiO*	TM1663	
Unknown			
Multidrug transport family			
Unknown		TM0287-8	
Unknown		TM0351-2	
Unknown		TM0387-9	
Unknown		TM0765	
Unknown		TM0793-4	
Unknown		TM1100	
Unknown		TM1302, 1304	
Unknown		TM1303, TM1306	
Unknown		TM1326, 7	
Unknown		TM1403-4	
Unknown		TM1638	
Amino acid families			
Glutamine?		TM0591-3	
		TM0827	
Branched-chain		TM1136-9	
Proteolytic domain family			
		TM0043	
Phosphate family			
		TM1261-4	
Spermidine/putrescine family			
		TM1377	

TABLE 1.4. Publications for *T. maritima* proteins characterized structurally by JCSG and others since 1999

Gene identifier	Gene name	Original annotation	Current annotation	Family	Crystal Structure	Reference(s)
TM0021	*glnB*	conserved hypothetical protein	PII-like signaling protein, potentially involved in signaling in response to cellular nitrogen status	P-II	1O51A	(193)
TM0064	*uxaC*	uronate isomerase, putative	uronate isomerase		1J5SA	(188)
TM0065	*kdgR**	transcriptional regulator, IclR family	KdgR transcriptional regulator	IclR	1MKMA	(249)
TM0096		conserved hypothetical protein	flavin mononucleotide binding protein; possible role in nitrogen metabolism	NIFR3	1VHNA	(144)
TM0126	*drrB*	response regulator	response regulator		1P2F	(7, 170)
TM0140	*igps*	indoleglycerol-phosphate synthase	indoleglycerol-phosphate synthase		1I4N	(79)
TM0156	*cutA*	alkaline phosphatase	alkaline phosphatase		1KR4	(182)
TM0160		conserved hypothetical protein	conserved hypothetical protein		1VJLB	(209)
TM0211	*gcvT*		T-protein of the glycine cleavage system		1WOS	(90)
TM0220	*fliG*	flagellar rotor protein	flagellar rotor protein		1QC7	(16, 108)
TM0231	*murC*	UDP-N-acetylmuramate-alanine ligase	UDP-N-acetylmuramate-alanine ligase; involved in peptidoglycan synthesis		1J6UA	(210)
TM0266		DNA-binding protein HU	DNA-binding protein HU		1B8Z	(23)
TM0293	*proA*	γ-glutamyl phosphate reductase	γ-glutamyl phosphate reductase; second step of proline biosynthesis	Aldehyde dehydrogenase	1O20A	(139)
TM0306		α-L-fucosidase, putative	α-L-fucosidase	GH29	1ODU	(214, 220)
TM0343	*dahps*	chorismate mutase, putative	3-deoxy-d-arabino-heptulosonate-7-phosphate synthase		1RZM	(204)
TM0364		4-α-glucanotransferase	4-α-glucanotransferase		1LWJ	(173)
TM0423		glycerol dehydrogenase	glycerol dehydrogenase		1KQ3	(14, 96)
TM0446	*purE*	phosphoribosylaminoimidazole carboxylase, catalytic subunit	phosphoribosylaminoimidazole carboxylase		1O4VA	(190)
TM0454		Ribosomal Protein L11	Ribosomal Protein L11			(61)

(Table 1.4, continued)

TM0457		ribosomal protein L12	ribosomal protein L12		1DD3	(230, 231)
TM0521	*hslV*	heat shock protein HslV	heat shock protein HslV	HslV	1M4Y	(208)
TM0574	*queA*	S-adenosylmethionine tRNA ribosyltransferase	S-Adenosylmethionine Trna Ribosyltransferase		1VKY	(114)
TM0613		conserved hypothetical protein	HEPN domain protein		1O3U	(35)
TM0651		conserved hypothetical protein	putative phosphorylated carbohydrate phosphatase	HAD	1NF2A	(202)
TM0653	*miaB*	conserved hypothetical protein	t-RNA methylthiotransferase			(151)
TM0654	*papt*	spermidine synthase	spermidine synthase		1INL	(81)
TM0665		cysteine synthase	O-Acetylserine Sulfhydrylase		1O58A	(53)
TM0680a	*fliN*	Fli-Y2 chemotaxis protein	FliN flagellar rotor protein		1YAB	(17)
TM0689	*tim*	triosephosphate isomerase	triosephosphate isomerase		1B9BB	(110)
TM0702	*cheA*	chemotaxis sensor histidine kinase CheA	chemotaxis sensor histidine kinase CheA		1TQG	(11, 146, 160)
TM0704	*pimt*	L-isoaspartate(D-aspartate) O-methyltransferase	protein isoaspartyl methyltransferase		1DL5	(206)
TM0727	*pmb*	pmbA-related protein	putative modulator of DNA gyrase		1VL4	(168)
TM0752	*agu4B*	α -glucosidase	α-glucuronidase	GH 4	1VJTA	(215)
TM0767	*mmtA*	maltosyltransferase	maltosyltransferase		1GJW	(172)
TM0835	*ftsA*	cell division protein FtsA, putative	cell division protein FtsA, putative		1E4G	(223)
TM0841		S-layer like array protein	fatty-acid binding protein	DegV/ COG1307		(186)
TM0851	*hrcA*	HrcA CIRCE-binding transcriptional regulator	HrcA CIRCE-binding transcriptional regulator	HrcA	1STZ	(105)
TM0864	*tRNase Z*	conserved hypothetical protein	Ribonuclease Z		1WW1	(62)
TM0872	*mraW*	conserved hypothetical protein	S-adenosyl-L-methionine dependent methyltransferase	COG0275	1N2XA	(121)
TM0875		hypothetical protein	YggU-like protein	YggU-like	1O22	(3)
TM0904	*cheC*	chemotaxis protein CheC	chemotaxis protein CheC		1XKR	(145)
TM0920		alcohol dehydrogenase, iron-containing	predicted iron-containing 1,3-propanediol dehydrogenase		1O2D	(194)
TM0935		conserved hypothetical protein	tandem cystathione-β-synthase (CBS) domain protein	CBS	1O50	(122)
TM0979	*dsrH*	conserved hypothetical protein	conserved hypothetical protein	YchN	1X9A	(45, 149)

(Table 1.4, continued)

TM1040		histidinol-phosphate aminotransferase	histidinol-phosphate aminotransferase		1UU0	(40)
TM1080	*rpiB*	ribose-5-phosphate isomerase	ribose-5-phosphate isomerase	RpiB	1O1X	(242)
TM1083		hypothetical protein	archaease, possible chaperonin	archaease	1J5U	(20)
TM1095	*purB*		adenylosuccinate lyase		1C3C	(222)
TM1112		conserved hypothetical protein	conserved hypothetical protein	cupin	1LKN, 1O5U	(117)
TM1158		conserved hypothetical protein	Glutamine amidotransferase, putative			(189)
TM1171		transcriptional regulator, crp family	transcriptional regulator, crp family	Crp	1O5L	(209)
TM1255		aspartate aminotransferase	Aspartate aminotransferase		1O4S	(192)
TM1287		conserved hypothetical protein	Oxalate decarboxylase, putative		1O4T	(195)
TM1290		conserved hypothetical protein	hypothetical protein		1RDU	(36, 37)
TM1363	*prfA*	peptide chain release factor 1	peptide chain release factor 1		1RQ0	(199)
TM1394		hypothetical protein	Hsp33 chaperone	Hsp 33	1VQ0	(65)
TM1399	*frr*	ribosome recycling factor	ribosome recycling factor		1T1M	(197, 198)
TM1414	*bfr*	Invertase (β-fructosidase)	Invertase (β-fructosidase)	GH32	1UYP	(1)
TM1442		anti-sigma factor antagonist, putative	anti sigma-factor antagonist		1SBO, 1T6R	(38, 50, 92)
TM1459		hypothetical protein	manganese-containing cupin protein		1VJ2	(66)
TM1463	*rnpA*	ribonuclease P protein component	bacterial ribonuclease P protein component		1NZ0	(221)
TM1464		conserved hypothetical protein	Conserved hypothetical protein possibly involved in carbohydrate metabolism			(97)
TM1478		methionine aminopeptidase	methionine aminopeptidase		1O0X	(211)
TM1499		ribosomal protein L4	ribosomal protein L4		1DMG	(238)
TM1509		conserved hypothetical protein	conserved hypothetical protein		1TVI	(148)
TM1535	*oppS*	octoprenyl-diphosphate synthase, putative	octaprenyl pyrophosphate synthase		1V4E	(49)
TM1560		serine cycle enzyme, putative	formiminotetrahydrofolate cyclodeaminase		1O5H	(241)
TM1618	*cheX*	chemotaxis protein CheX	chemotaxis protein CheX		1XKO	(145)
TM1621	*gdpd*	hypothetical protein	glycerophosphodiester phosphodiesterase		1O1Z	(179)
TM1641	*dyrA*	dihydrofolate reductase	dihydrofolate reductase		1D1G	(28)

(Table 1.4, continued)

TM1643	*nadX*	conserved hypothetical protein	aspartate dehydrogenase		1J5P, 1H2H	(244)
TM1645		nicotinate-nucleotide pyrophosphorylase	type II quinolic acid phosphoribosyltransferase		1O4U	(191)
TM1662	*surE*	5'-nucleotidase surE	5'-nucleotidase surE		1J9L	(93)
TM1692	*nifS*	aminotransferase, class V	NifS-Like Protein		1ECX	(72)
TM1717	*yjeQ*	conserved hypothetical protein	RNA-binding protein	YjeQ	1U0L	(200)
TM1730	*ruvB*	Holliday junction DNA helicase ruvB	Holliday junction DNA helicase ruvB		1IN5	(157)
TM1734	*phoU*	phosphate transport system protein phoU homolog 2	metalloprotein with multinuclear iron clusters		1SUM	(106)
TM1765	*nusB*	N utilization substance protein B	NusB antitermination factor		1TZX	(12)
TM1777	*nusA*	N utilization substance protein A	N utilization substance protein A		1L2F	(201)
TM1816		conserved hypothetical protein	conserved hypothetical protein		1T3V	(25)
TM1834	*aglA*	Alpha-glucosidase (Maltase)	alpha-glucosidase			(109)

FIGURE LEGENDS

FIG. 1.1. Genomic diversity across the *Thermotogales* assessed by comparative genomic hybridization (CGH) (adapted from (127). A.) The outer and second circles represents the predicted coding regions on the plus strand and minus strand, respectively, color-coded by role categories: violet, amino acid biosynthesis; light blue, biosynthesis of cofactors, prosthetic groups, and carriers; light green, cell envelope; red, cellular processes; brown, central intermediary metabolism; yellow, DNA metabolism; light gray, energy metabolism; magenta, fatty acid and phospholipid metabolism; pink, protein synthesis and fate; orange, purines, pyrimidines, nucleosides, and nucleotides; olive, regulatory functions and signal transduction; dark green, transcription; teal, transport and binding proteins; gray, unknown function; salmon, other categories; blue, hypothetical proteins. The third circle displays the atypical nucleotide composition curve. The fourth circle displays the "archaeal islands" as predicted by Nelson et al. (). The fifth circle displays the regions, divergent in *Thermotoga* strain RQ2 compared to MSB8, that were analyzed in detail by Mongodin et al. (127). The circles 6 to 14 show the CGH hybridization ratios R (comparison to MSB8) for strains RQ2 (circle 6), S1/L12B (circle 7), PB1platt (circle 8), NE7/L9B (circle 9), NE2x (circle 10), LA10 (circle 11), LA4 (circle 12), RQ7 (circle 13) and VMA1/L12B (circle 14): gray, $R<3$ (genes shared between MSB8 and test strain); yellow, $3<R<5$; orange, $5<R<7$; red, $7<R<10$; brown, $R>10$ (genes divergent between MSB8 and test strain). B.) CGH results and hierarchical clustering based on the CGH data.

FIG. 1.2. Predicted pathway for the utilization of α- and β-glucan polysaccharides by *T. maritima*. Extracellular and intracellular hydrolases, as well as predicted transporters are shown. Abbreviations include: 4-α-GTase, 4-alpha-glucanotransferase. ATP, adenosine tri-phosphate, EMP, Embden-Meyerhoff-Parnas glycolytic pathway, ED, Entner-Doudoroff glycolytic pathway.

FIG. 1.3. Predicted pathway for the utilization of β-mannan polysaccharides by *T. maritima*. Extracellular and intracellular hydrolases, as well as predicted transporters are shown. Abbreviations include: DH, dehydrogenase, ATP, adenosine tri-phosphate.

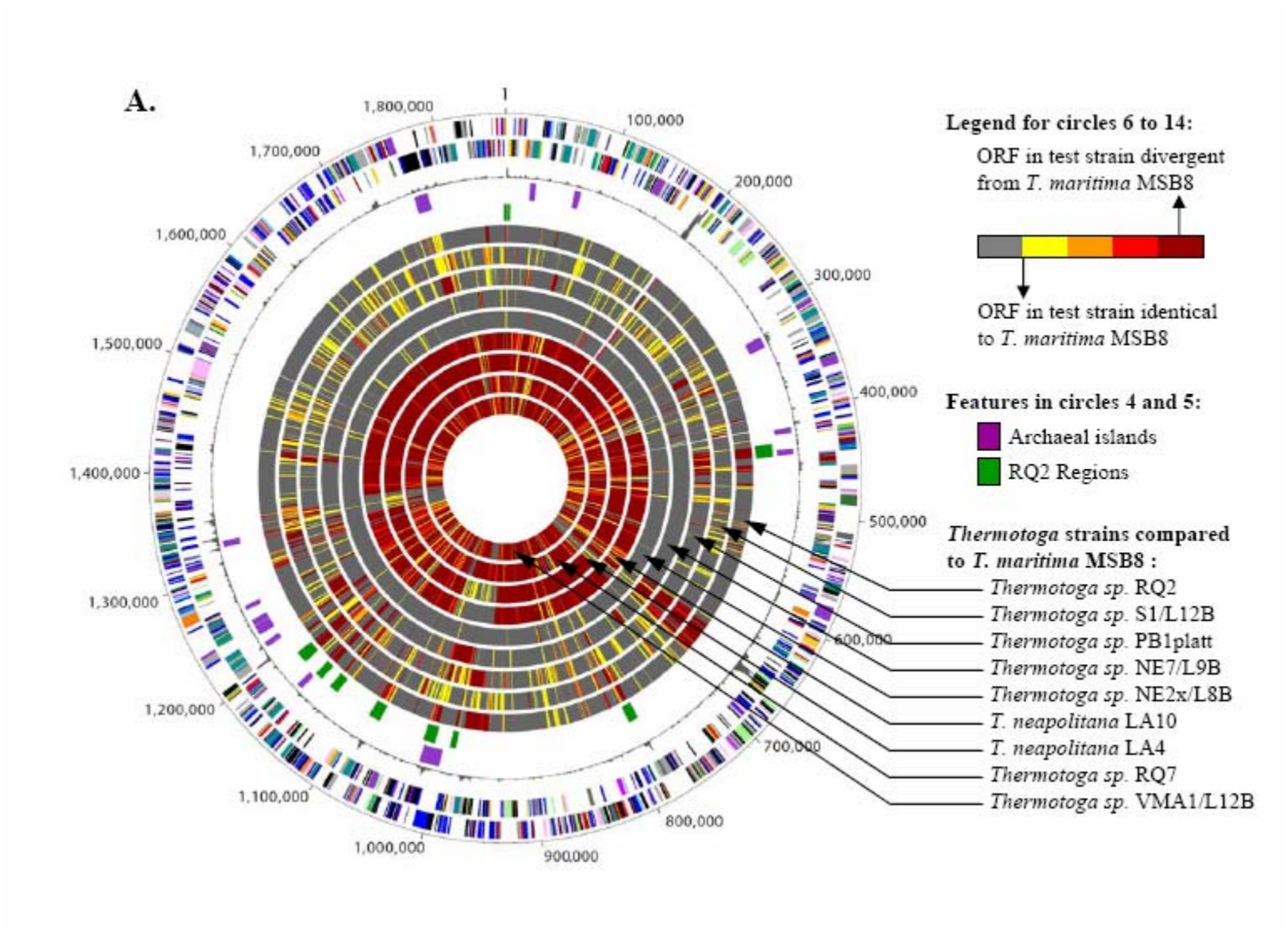
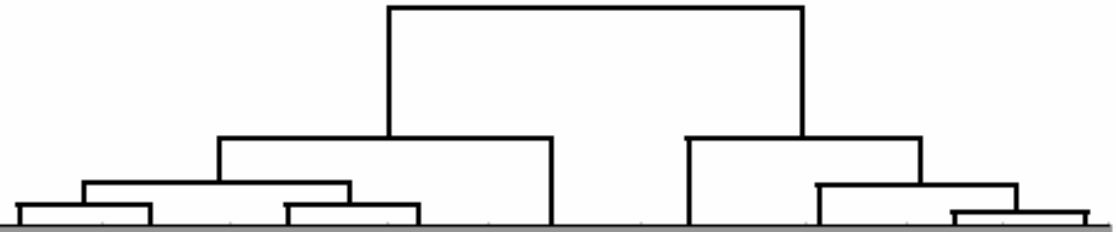

A.
1
100,000
200,000
300,000
400,000
500,000
600,000
700,000
800,000
900,000
1,000,000
1,100,000
1,200,000
1,300,000
1,400,000
1,500,000
1,600,000
1,700,000
1,800,000
Legend for circles 6 to 14:
ORF in test strain divergent from T. maritima MSB8
ORF in test strain identical to T. maritima MSB8
Features in circles 4 and 5:
Archaeal islands
RQ2 Regions
Thermotoga strains compared to T. maritima MSB8 :
Thermotoga sp. RQ2
Thermotoga sp. S1/L12B
Thermotoga sp. PB1platt
Thermotoga sp. NE7/L9B
Thermotoga sp. NE2x/L8B
T. neapolitana LA10
T. neapolitana LA4
Thermotoga sp. RQ7
Thermotoga sp. VMA1/L12B

Figure 1.1A

B.



Test Strains compared to MSB8	*NE_7/L9B*	*S1/L12B*	*NE2x_L8B*	*RQ2*	*PB1platt*	*VMA1/L12B*	*RQ7*	*LA4*	*LA10*
Number of genes with R < 3	1648	1280	1784	1558	1284	233	141	129	206
% of genes with R < 3	88.36	68.63	95.66	83.54	68.85	12.49	7.56	6.92	11.05
Number of genes with 3 < R < 10	98	436	78	178	395	338	339	180	231
% of genes with 3 < R < 10	5.25	23.38	4.18	9.54	21.18	18.12	18.18	9.65	12.39
Number of genes with R > 10	119	149	3	129	186	1294	1385	1556	1428
% of genes with R > 10	6.38	7.99	0.16	6.92	9.97	69.38	74.26	83.43	76.57

Figure 1.1B

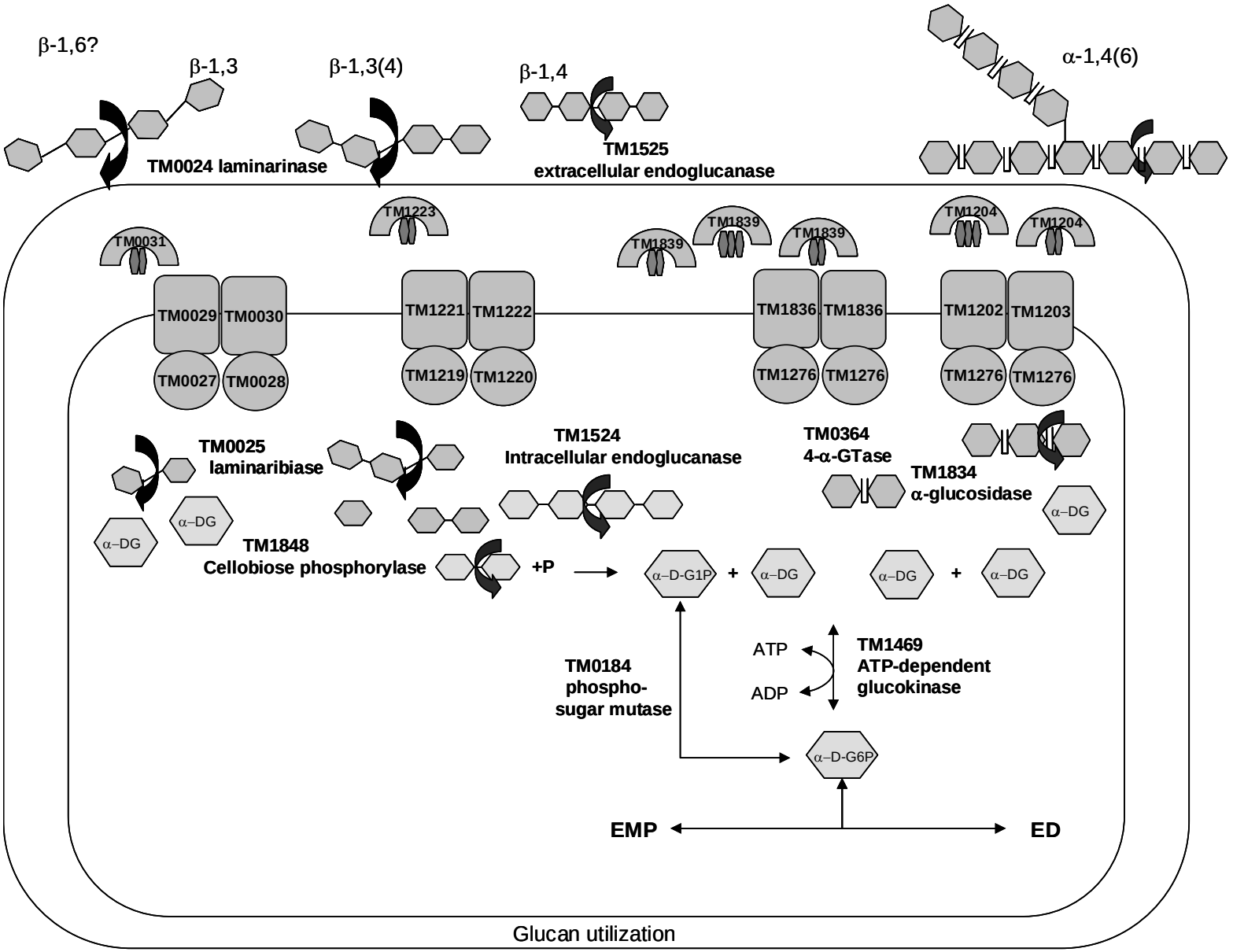

β-1,6?
β-1,3
β-1,3(4)
β-1,4
α-1,4(6)
TM0024 laminarinase
TM1525
extracellular endoglucanase
TM1223
TM0031
TM1839
TM1839
TM1839
TM1204
TM1204
TM0029
TM0030
TM1221
TM1222
TM1836
TM1836
TM1202
TM1203
TM0027
TM0028
TM1219
TM1220
TM1276
TM1276
TM1276
TM1276
TM0025
laminaribiase
TM1524
Intracellular endoglucanase
TM0364
4-α-GTase
TM1834
α-glucosidase
α–DG
α–DG
α–DG
TM1848
Cellobiose phosphorylase
+P
α–D-G1P
+
α–DG
α–DG
+
α–DG
ATP
ADP
TM1469
ATP-dependent
glucokinase
TM0184
phospho-
sugar mutase
α–D-G6P
EMP
ED
Glucan utilization

Figure 1.2

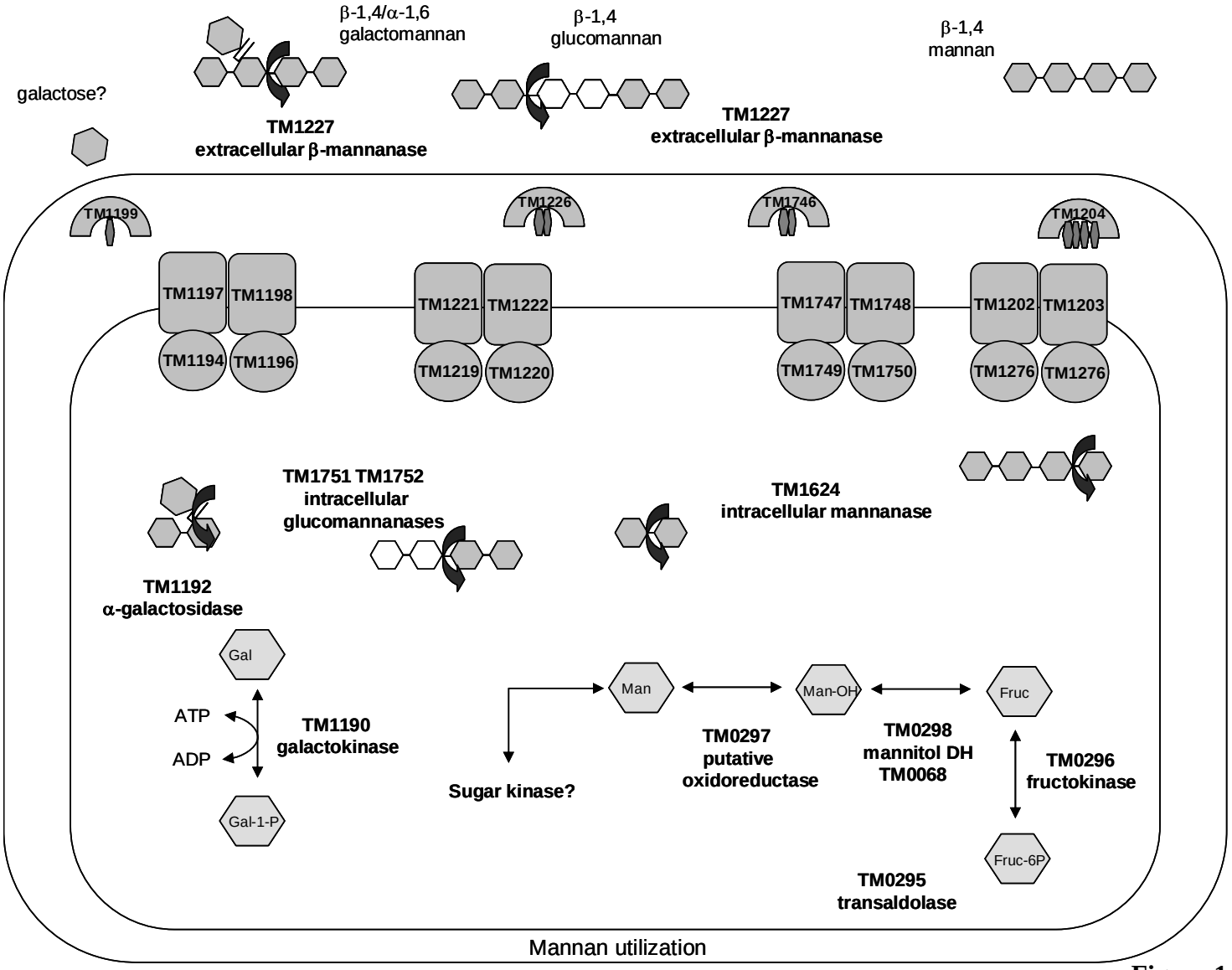

β-1,4/α-1,6
galactomannan
β-1,4
glucomannan
β-1,4
mannan
galactose?
TM1227
extracellular β-mannanase
TM1227
extracellular β-mannanase
TM1199
TM1226
TM1746
TM1204
TM1197
TM1198
TM1221
TM1222
TM1747
TM1748
TM1202
TM1203
TM1194
TM1196
TM1219
TM1220
TM1749
TM1750
TM1276
TM1276
TM1751 TM1752
intracellular
glucomannanases
TM1624
intracellular mannanase
TM1192
α-galactosidase
Gal
ATP
ADP
TM1190
galactokinase
Gal-1-P
Sugar kinase?
Man
Man-OH
Fruc
TM0297
putative
oxidoreductase
TM0298
mannitol DH
TM0068
TM0296
fructokinase
Fruc-6P
TM0295
transaldolase
Mannan utilization

Figure 1.3

Chapter 2:

Transcriptional analysis of biofilm formation processes in the anaerobic hyperthermophilic bacterium *Thermotoga maritima*

Marybeth A. Pysz[1], Shannon B. Conners, Clemente I. Montero,
Keith R. Shockley[3], Matthew R. Johnson, Donald E. Ward[2], and Robert M. Kelly

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

[1] Current address: Roswell Park Cancer Institute
Department of Pharmacology and Therapeutic
Elm and Carlton Streets
Buffalo, NY 14263

[2] Current address: Genencor International, Inc.
925 Page Mill Road
Palo Alto, CA 94304

[3] Current address: The Jackson Laboratory
600 Main Street
Bar Harbor, ME 04609

Pysz, M. A., S. B. Conners, C. I. Montero, K. R. Shockley, M. R. Johnson, D. E. Ward, and R. M. Kelly. 2004. Transcriptional analysis of biofilm formation processes in the anaerobic, hyperthermophilic bacterium *Thermotoga maritima*. Appl Environ Microbiol **70:**6098-112.

ABSTRACT

***Thermotoga maritima*, a fermentative, anaerobic, hyperthermophilic bacterium, was found to attach to bioreactor glass walls, nylon mesh and polycarbonate filters during chemostat cultivation on maltose-based media at 80°C. A whole genome cDNA microarray was used to examine differential expression patterns between biofilm and planktonic populations. Mixed model statistical analysis revealed differential expression (two-fold or more) of 114 ORFs in sessile cells (6% of the genome), over a third of which were initially annotated as hypothetical proteins in the *T. maritima* genome. Among the previously annotated genes in the *T. maritima* genome, which showed expression changes during biofilm growth, were several that corresponded to biofilm-formation genes identified in mesophilic bacteria (i.e., *Pseudomonas* species, *Escherichia coli*, and *Staphylococcus epidermidis*). Most notably, *T. maritima* biofilm-bound cells exhibited increased transcription of genes involved in iron and sulfur transport, as well as in biosynthesis of cysteine, thiamine, NAD, and isoprenoid side chains of quinones. These findings were all consistent with the up-regulation of iron-sulfur cluster assembly and repair functions in biofilm cells. Significant up-regulation of several β-specific glycosidases was also noted in biofilm cells, despite the fact that maltose was the primary carbon source fed to the chemostat. The reasons for increased β-glycosidase levels are unclear but are likely related to the processing of biofilm-based polysaccharides. In addition to revealing insights into the phenotype of sessile *T. maritima* communities, the methodology developed here can be extended to study other anaerobic biofilm formation processes as well as to examine aspects of microbial ecology in hydrothermal environments.**

INTRODUCTION

Mesophilic bacteria in natural and pathogenic environments are often associated with biofilms. This localization facilitates interactions and coexistence in an optimized microenvironment while at the same time limiting the adverse consequences of competition and selectivity (12). The establishment of a sessile community of cells encapsulated by a polysaccharide matrix on a surface involves a complex series of steps: initial attachment, production of exopolysaccharides, early biofilm development, mature biofilm formation, and detachment of cells, perhaps as communities (50, 98). These steps have been investigated for several mesophilic bacteria, including *Pseudomonas aeruginosa* (20, 107), *Bacillus cereus* (73), *Vibrio cholerae* (129), and a *Streptococcus* sp. (102). Biofilm formation apparently requires expression of a distinct set of genes that differentiate sessile from planktonic cells, including those related to chemotaxis, motility, exopolysaccharide biosynthesis, and stress response (90). However, this set of genes may only comprise about 1% of the total genome such that differences between planktonic and sessile cells may be subtle (28, 123). This is not surprising since biofilm-bound populations likely include newly recruited cells that have a planktonic phenotype as well as cells that represent various stages of biofilm formation (92, 121). Furthermore, even interactions between planktonic cells and surfaces can affect gene expression. For *Bacillus cereus*, planktonic cells grown in the presence of biofilm substratum (glass wool) shared common differentially expressed genes with biofilm bound cells (73). Thus, composite planktonic and sessile communities likely contain a continuous distribution of distinct phenotypes that have temporal and spatial signatures (98).

The capacity to form biofilms is not limited to aerobic, mesophilic bacteria. Biofilms are also evident in high temperature environments, such as terrestrial geothermal settings and hydrothermal vents (85). Several anaerobic hyperthermophiles (microorganisms with optimal growth temperatures at or above 80°C) have been shown to produce exopolysaccharides (3, 31, 69). These exopolysaccharides form the basis for biofilms, which have been observed in pure cultures of *Archaeoglobus fulgidus* (53) *Thermotoga maritima* (86), and *Thermococcus litoralis* (89), as well as in co-cultures of *T. maritima* and *Methanococcus jannaschii* (65, 86). Biofilm formation was induced by elevated pH, decreased and increased growth temperature, high salt, and exposure to ultraviolet light, oxygen, or antibiotic in *A. fulgidus* (53) and by ammonium chloride in *T. litoralis* (89).

A key challenge that must be addressed to further explore biofilm formation processes in hyperthermophilic anaerobes is the experimental complexity associated with the growth of these organisms. This problem was addressed with a high temperature, anaerobic chemostat that was used to generate biofilms in cultures of *T. maritima* that could be sampled and examined for differential gene expression patterns by whole genome cDNA microarrays comparing planktonic to sessile cells. Transcriptional patterns related to the biofilm phenotype in this hyperthermophilic microorganism were then determined and compared to biofilm formation in less thermophilic microorganisms. Such information concerning biofilm formation mechanisms in hyperthermophiles is needed to develop a better understanding of the microbial ecology in hydrothermal habitats, particularly in regard to surface colonization.

MATERIALS AND METHODS

Microorganism and growth conditions. *Thermotoga maritima* (DSM 3109) was grown anaerobically on Sea Salts Medium (SSM) containing 40 g/l sea salts (Sigma, St. Louis, MO), 1 g/l yeast extract (Fisher Scientific, Pittsburgh, PA), 3.1 g/l PIPES buffer (Sigma Chemical, St. Louis, MO), 2 g/l tryptone, 2 ml/l of 0.05% Resazurin, and 10 ml/l 10X Wolin minerals (119). Growth medium was adjusted to pH 6.8 with KOH (Fisher Scientific, Pittsburgh, PA), and autoclaved prior to use. Batch cultures (50 ml) were inoculated under N_2 (high purity nitrogen; National Welders, Raleigh, NC) headspace, as previously described (79, 87), and were grown at 80°C for 8 to 10 hours in oil baths. Maltose (Sigma Chemical, St. Louis) was added to SSM (final concentration 5 g/l) as a carbon source prior to inoculation. Continuous cultivation of T. maritima was performed in a 2-L five-neck, round-bottom flask, as previously described (79, 87). A 50 ml batch culture was used to inoculate 1 L of SSM medium supplemented with 5 g/l maltose in the flask. This seed culture was grown at 80°C for 8.5 hours under continuous nitrogen sparging, after which medium was fed at a dilution rate of 0.25 h^{-1}. Media for continuous cultivation was prepared in 9 L batches at 1.2X concentration as mentioned above, to which 1 L of a filter-sterilized maltose solution (50 g) was added immediately after autoclaving. The pH of the culture was continuously monitored with a Chemcadet pH controller (Cole Parmer, Vernon Hills, IL) and adjusted by the addition of 1 M NaOH. Temperature control was effected using a Digi-Sense controller (Cole-Parmer, Vernon Hills, IL) such that variations were typically ± 0.8°C, and verified by a mercury glass thermometer inserted into the culture. Steady-state conditions were monitored by following cell counts (see below) and optical densities at 600 nm. All planktonic cell

samples were collected from the outlet line into sterile pyrex bottles (see below), from which 1 ml of cells was fixed in glutaraldehyde for cell counting.

Biofilm substrata and collection. Nylon mesh (Sefar America, Hamden, CT) and polycarbonate filters (Poretics 0.22 μm, Fisher Scientific, Pittsburgh, PA) were used as substrata for biofilm formation. Twelve squares of mesh (13.3 cm × 9.8 cm) were cut, rolled tightly, and tied with polycarbonate string. Three rolled mesh squares were tied to one another at the ends. Polycarbonate filters were tied to the center of each set of mesh squares to be used for biofilm imaging, while the mesh itself provided biomass for RNA samples within the biofilm. The mesh and loose polycarbonate filters were placed in the reactor and autoclaved prior to startup. The strings were suspended in the growing culture until the sample was collected, whereby they were pulled quickly through one of the 5 necks of the reactor. The mesh samples and polycarbonate filters were rinsed twice in sterile media while on ice to remove loosely adhered planktonic cells. Polycarbonate filters were removed from each tube and placed in 2.5% glutaraldehyde (Sigma, St. Louis, MO) to fix the biofilm cells for examination and imaging under the microscope (see below). The mesh squares were separated from the strings and submerged in 50 mL conical tubes containing 300 mM NaCl (Fisher Scientific, Pittsburgh, PA). The conical tubes were vortexed vigorously (4°C) to remove biofilm material from the mesh, after which the mesh was removed and the resulting suspension centrifuged (10000×g, 20 minutes, 4°C) to pellet the biofilm cell material. Rinse media and 300 mM NaCl were chilled at 4°C prior to use. RNA was extracted as described below.

Imaging and microscopy methods. Epifluorescent micrographs were taken with a SPOT digital camera (Southern Micro Instruments, Atlanta, GA) attached to a Nikon (Labophot-2) microscope (Southern Micro Instruments, Atlanta, GA) with 100X oil-immersion lens. Planktonic cell suspensions were fixed in 2.5% glutaraldehyde and stained with acridine orange (1 g/l; Fisher Scientific, Pittsburgh, PA) to determine cell densities (79, 87). Biofilm cells on polycarbonate filters were fixed as described above, stained in acridine orange (1 g/l), and dried briefly under vacuum prior to imaging. A scanning electron microscope (JEOL JSM-35CF Microscope, North Carolina State University, Department of Veterinary Medicine) was also used to image biofilm cells on polycarbonate filters. Filters were fixed in 2.5% glutaraldehyde, and critically point dried in CO_2. Images of reactor walls of the continuous culture were also taken regularly with a Nikon Coolpix 950 digital camera.

RNA sample collection. Approximately 200 ml samples of planktonic cells were withdrawn through culture outlet (79) into sterile pyrex bottles on ice. Fifty ml of cells were collected prior to sampling to eliminate existing fluid in the lines. The 200 ml-samples were used for RNA extractions and processed as described previously (14). Biofilm pellets were rinsed once after being extracted in 300 mM NaCl (4°C) and used immediately for RNA isolation (i.e. RNase inhibiting buffers were added directly after rinsing step). Total RNA from planktonic cells was extracted from 3 different time point samples (Figure 2.1) during the steady-state operation; approximately 1 mg of RNA was obtained from each sampling time from which 100 μg was pooled. Similarly, total RNA from the 3 rolls of mesh from one point in the middle of the chemostat run was pooled to

produce a biofilm sample. The cDNA generated from the planktonic and biofilm cells was hybridized to glass slides containing the targeted microarray, scanned and analyzed, as described previously (15).

Construction of the whole genome cDNA microarray. Open reading frames were identified from the *T. maritima* MSB8 genome available at http://www.tigr.org/tigrscripts/CMR2/GenomePage3.spl?database=btm. DNA primers were designed with similar annealing temperatures and minimal hairpin formation using Genomax (Informax, Bethesda, MD). Probes were PCR-amplified in a PTC-100 Thermocycler (MJ Research, Inc., Waltham, MA) using *Taq* polymerase (Boehringer, Indianapolis, IN) and *T. maritima* genomic DNA, isolated as described previously (14). Purification of PCR products and microarray construction were performed using protocols described elsewhere (14). PCR products were randomized within plates before printing using a random number generator, and each gene was spotted six times on each array.

Labeling and hybridization. The whole genome microarray was interrogated using methods previously described (14). Briefly, first-strand cDNA was prepared from *T. maritima* total RNA using Stratascript (Stratagene, La Jolla, CA) and random hexamer primers (Invitrogen Life Technologies, Carlsbad, CA); 5-[3-Aminoallyl]-2'-deoxyuridine-5'-triphosphate (Sigma) was used for dye incorporation, as described elsewhere (35). Each biofilm or planktonic cDNA sample was labeled with Cyanine-3

and Cyanine-5 and the samples hybridized to different arrays. The slides were scanned and processed using a Perkin-Elmer Scanarray Express Lite.

Statistical analyses and determination of differential gene expression. Replication of treatments, arrays, dyes, and cDNA spots allowed the use of analysis of variance (ANOVA) models for data analysis (124). Cyanine-3 labeled biofilm cDNA and Cyanine-5 labeled planktonic cDNA were hybridized on one chip, and Cyanine-3 labeled planktonic cDNA and Cyanine-5 labeled biofilm cDNA were hybridized on another chip. The data import code and statistical analysis procedures reported previously (14) were used to analyze spot intensities obtained from Quantarray. Briefly, a linear normalization ANOVA model (124) was used to estimate global variation in the form of fixed (dye (D), treatment (T)) and random (array (A), spot (A(BS)), block (A(B)) effects and random error using the model $\log_2(y_{ijklmn}) = \mu + A_i + D_j + T_k + A_i(B_l) + A_i(S_m B_l) + \varepsilon_{ijklm}$. A gene-specific ANOVA model was then used to partition the remaining variation into gene-specific effects using the model $r_{ijklm} = \mu + A_i + D_j + T_k + A_i(B_l) + A_i(S_m B_l) + \varepsilon_{ijklm}$. Volcano plots were used to visualize interesting contrasts, or comparisons between two treatments (124). A statistical test with the null hypothesis of no differential expression was performed for each of the 1880 ORFs included on the array. A Bonferroni correction was used to adjust alpha (α) for the expected increase in false positives due to multiple tests (124). The Bonferroni correction, calculated by dividing 0.05 by 1880, yielded a corrected alpha of 0.00003, equivalent to a $-\log_{10}$(p-value) = 4.5. Genes meeting this significance criterion, and showing fold changes of ± 2.0 or greater, were selected for further examination.

For complete information on significance of expression changes and fold changes, see our website (to be included after publication on our microarray data page at http://www.che.ncsu.edu/extremophiles/microarray/index.html).

RT-PCR confirmation of gene expression. Real time PCR was used to confirm the microarray results of four up-regulated genes (TM0851 (2.3-fold), TM1645 (8.1-fold), TM1848 (6.9-fold), TM1867 (5.1-fold)) and one unchanged, control gene (TM0403 (1.2-fold)). Primers were designed using Genomax software: TM0851: GCATAACCGTCAGGATAGGAAG and TTCGACGTGAAGAGGTACACAC; TM1645: TGTCATGCTGGACAATCTCTCT and ACTTCCACGATCACGTTAGGAT; TM1848: ATGGAAGCACTTACCACCAGTT and CCAGTCACCTGTCTCTTTGATG; TM1867: GGAGAACATGGAGATTCAGAGG and ATCGCACTTCTGACAAATCTGA; and TM0403 AGGTGATGCTTCTCATAGCGGT and ATCCTAATGCAATCCAGCAGATCCA. Reverse transcription of RNA to cDNA was performed as described above. RT-PCR was performed using the SYBRGREEN kit and iCycler iQ Real-time PCR detection system (Bio-Rad Laboratories, Hercules, CA) according to manufacturer protocols. Briefly, reactions for 10 ng of samples were carried out for the 5 genes at three different temperatures to determine the optimum S-curves. Optimization indicated that all reactions could be performed at 55°C. Standard curves (20 ng, 4 ng, 0.8 ng, and 0.16 ng) for biofilm samples were run along with 10 ng of planktonic and biofilm samples for each gene. Quantitative results were calculated using vendor-provided software (Bio-Rad Laboratories, Hercules, CA). In all cases, RT-PCR results exhibited the same patterns as those obtained from cDNA microarray analysis.

Fold changes calculated from real time PCR were as follows: TM0851, 6.7-fold; TM1645, 10.0-fold; TM1867, 51.1-fold; TM1848, 17.4-fold; TM0403, 1.3-fold. In all cases of differentially expressed genes, the microarray tended to underestimate the fold changes calculated by real-time PCR, which is not surprising given the smaller dynamic range of microarray scanners when compared to real time PCR.

RESULTS AND DISCUSSION

***T. maritima* growth in continuous culture and biofilm formation.** Because efforts with batch culture were unsuccessful in generating sufficient attached cellular material for transcriptional analysis, a high temperature anaerobic chemostat was operated to collect *T. maritima* biofilm formed on removable nylon mesh. The mesh was used to create a compact, high surface area substratum for biofilm attachment; materials like this have been used successfully to study *P. aeruginosa* biofilms (19). *T. maritima* (T_{opt} 80°C) (38) was grown in continuous culture (dilution rate, D = 0.25 hr^{-1}) for over 300 hours (Figure 2.1). Figure 2.2a shows the formation of substantial wall growth in the 80°C reactor (88). Epifluorescent micrographs of polycarbonate filters placed in the chemostat showed significant cell attachment at 80°C (Figure 2.2B); this was supported by SEM analysis of biofilm cells on the filters which showed cells associated with rope-like structures, consistent with SEM analysis of mesophilic biofilms (Figure 2.2C) (22, 24, 41).

Whole-genome cDNA microarray analysis of differential gene expression of sessile and planktonic *T. maritima*. Despite the inherent heterogeneity of the biofilm state, planktonic and sessile *T. maritima* cells could be differentiated by transcriptional response patterns as determined by cDNA microarray analysis (see Figure 2.3). Table 2.1

lists genes exhibiting significant expression changes (≥ 2-fold, –log10(p-value) ≥ 4.5) for biofilm cells as compared to planktonic cells. The cDNA microarray results were confirmed by real-time RT-PCR (see Materials and Methods). For *T. maritima*, approximately 114 genes of the entire genome were differentially expressed two-fold or higher at this significance level; 43 of these 114 genes were originally annotated as hypothetical proteins (67). *T. maritima* gene expression patterns were further analyzed according to function, genomic location, and in comparison to biofilm gene expression profiles in mesophilic bacteria. A complete list of expression changes for predicted operons responding significantly between biofilm and planktonic cells is shown in Table 2.2. Where appropriate, gene annotations are updated with information from comparative genomics and functional studies subsequent to the publication of the *T. maritima* genome sequence (67).

Oxidative and thermal stress response. Biofilm formation has been observed as a response to oxidative stress in the hyperthermophilic archaeon, *A. fulgidus* (53), and certain aerobic mesophilic biofilms showed increased expression of oxidative stress genes (29, 93). Increased protein levels of superoxide dismutase and alkyl hydroperoxide reductase in aerobic mesophile biofilms have also been reported (92). Here, the observed down-regulation of predicted operon members rubrerythin (TM0657, -3.7-fold), superoxide reductase (SOR, TM0658, -3.4-fold), and rubredoxin (TM0659, -3.1-fold) in biofilm cells, along with an ahpC-related alkylhydroperoxide reductase (TM0807, -6.0-fold), was somewhat unexpected. All four proteins share high identity (57-67%) to homologs in *Pyrococcus* species (9, 16, 39, 45). The *Pyrococcus* homologs of rubredoxin

and SOR are known to be involved in the NADPH-dependent detoxification of dioxygen to H_2O_2 (39), while a *P. horikoshii* AhpC homolog (69% id/214 aa with TM0807) participates in a pathway responsible for the reduction of H_2O_2 to alcohol (45). However, an AhpC-related gene (TM0780) encoding a putative thioredoxin peroxidase/bacterioferritin comigratory protein (Bcp) was up-regulated 2.6-fold in biofilm cells (40). A second related gene (TM0386, 3.0-fold) containing an apparently unique combination of a Bcp thioredoxin peroxidase domain and a nitroreductase domain was also up-regulated.

Several possible explanations exist for the down-regulation of the SOR gene cluster. Lower expression of these genes has been observed during stationary phase (Johnson and Kelly, unpublished data); therefore, decreases in expression here may reflect similarities between stationary phase and biofilm cells. Alternatively, down-regulation of these genes may suggest lower residual oxygen exposure for cells trapped within a biofilm matrix, although both *T. maritima* planktonic and biofilm cells were cultured in the same anaerobic chemostat. Finally, the down-regulation may reduce exposure of biofilm cells to hydrogen peroxide during oxygen detoxification. Work in *E. coli* K12 has implicated cysteine as the reductant responsible for rapid recycling of iron (II) to iron (III), allowing reactions between hydrogen peroxide and iron (III) to drive the formation of DNA-damaging hydroxyl radicals (77). Indications of DNA damage in *T. maritima* biofilm cells were observed in the up-regulation of genes encoding a homolog of Sms/RadA (TM0199, 4.6-fold) involved in recombination and repair in *E. coli* K12 (6), a putative endonuclease specific to archaea (TM0664, 2.3-fold), and a predicted endonuclease (TM1545, 2.0-fold) related to proteins involved in recombination events.

DNA protection and repair proteins were also induced in *Listeria monocytogenes* biofilms (109).

Gene expression analysis of biofilm-bound *T. maritima* cells revealed the induction of genes implicated in thermal stress response. Here, biofilm cells displayed 2.3-fold higher expression of the CIRCE (controlling inverted repeat of chaperone expression) -binding HrcA repressor (TM0851), which controls expression of major heat shock operons in a number of species. We have previously noted the conservation of the CIRCE element upstream of the *T. maritima hrcA-dnaJ-grpE* and *groESL* operons (TM0849-TM0851) but not upstream of *dnaK-smHSP* (TM0373-TM0374) (80). Thermal stress genes have been shown to be up-regulated in biofilms of *P. aeruginosa* (*groES*, *dnaK*) (123) and *S. mutans* (*grpE* and *dnaK*) (102). *T. maritima grpE* (TM0850, 1.7-fold), *dnaJ* (TM0849, 1.4-fold), *dnaK* (TM0373, 1.9-fold), heat shock protein class I gene (TM0374, 2.3-fold), and *groES* (TM0373, 1.4-fold) were all up-regulated in biofilm cells. It was interesting that a cold shock protein (TM1874, –2.8 fold) recently shown to act as a functional homolog of the RNA-binding *E. coli* K12 *cspA* was down-regulated (78). Schembri et al. (93) also observed down-regulation of the cold shock protein encoded by *cspA* in a microarray experiment comparing *E. coli* K12 biofilm and exponential phase planktonic cells.

Exopolysaccharide biosynthesis and degradation. A 2.0-fold upregulation of TM1535 (octaprenyl pyrophosphate synthase) involved in isoprenoid chain synthesis was observed in biofilm cells. Isoprenoid chains serve as scaffolds or lipid carriers for the assembly of monosaccharides into linear and/or branched polysaccharide chains via glycosyl

transferases (114). Despite the presence in the *T. maritima* genome of a large cluster of glycosyltransferases, very few were differentially regulated between biofilm and planktonic cells. In fact, a number were expressed 1.5-1.7 fold higher in planktonic cells, including TM0630, a NDP-sugar epimerase related to UDP-glucose-4-epimerases, TM0627, a putative NDP-linked sugar glycosyltransferase, and TM0818, a techoic acid biosynthesis protein related to GumM in *Xanthomonas campestris pv. gum* (data not shown). A glycosyltransferase (TM0392, -2.2) predicted to be involved in the synthesis of NDP linked sugars was also down-regulated in biofilm cells, while a homolog of *E. coli* K12 *ushA* which encodes a periplasmic protein with UDP-sugar hydrolase activity (11) was up-regulated (TM1878, 2.3 fold). It is possible that exopolysaccharide synthesis occurs both in biofilm and planktonic cells, since commonalities in expression patterns have been observed between biofilm cells and planktonic cells in the presence of biofilm substratum (73).

Glycosyl hydrolases may also be involved in exopolysacchride synthesis and/or degradation. Induction of an NAD^+-dependent family 4 α-glucuronidase (TM0752, 3.1-fold) (101), and a β-galactosidase (TM0310, 2.9-fold) was observed in the apparent absence of growth substrates related to these enzymes. Both proteins have been observed to be up-regulated during early stationary phase in *T. maritima-M. jannaschii* co-culture experiments when the formation of biofilm material is observed (Johnson and Kelly, unpublished). Cellobiose phosphorylase (CepA) (TM1848) (67) exhibited a 6.9-fold expression increase in biofilm cells compared to planktonic cells at 80°C. CepA from *Thermotoga neopolitana* has sole substrate specificity for cellobiose (128) which it converts to D-glucose and glucose-1-phosphate (70). Characterization of the *T. maritima*

homolog revealed substrate specificity for cellobiose in the hydrolysis reaction but relaxed synthetic specificity for the reverse reaction, allowing mannose, xylose, glucosamine, 2- and 6-deoxy-D-glucose, and β-D-glucoside to act as glucosyl acceptors for glucose-1-P (81). The strong up-regulation of this gene was unexpected since maltose (α-1,6) and not cellobiose (β-1,4) was used as the primary carbon source in the growth medium. The up-regulation of the operon (TM1524-1536) containing β-endoglucanases Cel12B (TM1524) and Cel12A (TM1525) (18, 58), previously shown to be up-regulated on carboxymethylcellulose, barley, and konjac glucomannan (14), was also noted. Further work will be necessary to determine whether the induction of glycoside hydrolases in biofilm cells is related to the synthesis or breakdown of exopolysaccharide-based biofilm material or the sloughing of biofilm.

ABC transporters. Several ABC transporter genes were differentially expressed in biofilm cells. Despite the fact that maltose was the primary carbohydrate in the growth medium, genes within a maltose utilization and transport operon (TM1834-1839) were down-regulated in biofilm cells. On the contrary, genes predicted to encode an uncharacterized multiple-sugar transport system (TM0418-0421) downstream of the FTR1-related iron transporter (TM0417) were up-regulated in biofilm cells along with a gene sharing domain similarity with sugar phosphate isomerases (TM0422, 2.8-fold).

Additional homologs to ABC transporters were up-regulated during biofilm growth (Table 2.2). It was particularly intriguing to note the up-regulation of two genes, which bear similarity to genes encoding antimicrobial peptide exporters. TM0352 (2.1-fold) is predicted to encode an ATP-binding ABC subunit (COG1136), while TM0351

(2.0-fold) possibly encodes an ABC-associated permease component (COG0577). Three additional upstream genes encode a putative membrane fusion protein (TM0353), outer membrane protein (TM0354), and TolC protein (TM0355). These genes (TM0353-TM0355) were not expressed differentially between biofilm and planktonic conditions. A distantly related, though not well conserved, multi-protein system is essential for biofilm adhesion in *Pseudomonas fluorescens* WCS365, consisting of an ABC ATPase, ABC permease, outer membrane protein and large adhesion protein with repetitive domains which is secreted via the transporter (36). A glycerol uptake facilitator protein (TM1429) (2.5-fold) was also up-regulated in *T. maritima* biofilm cells; ferric iron and glycerol may be required for antimicrobial peptide release as shown during biofilm growth of *Bacilli* (125).

Response of iron/sulfur uptake and utilization genes in biofilm cells. Biofilm cells showed increased expression of iron and sulfur uptake systems, consistent with up-regulation of genes encoding iron-sulfur cluster-containing proteins and components of a chaperone system involved in iron-sulfur cluster formation and repair. Predicted operons containing these genes are present in a number of distinct regions of the *T. maritima* genome. Known Fe-S clusters or cysteine-rich sequence motifs in the corresponding proteins are noted in Table 2.2.

Up-regulation of iron uptake is important in mesophile biofilm formation processes (10) which likely relates to the observed induction of genes encoding iron acquisition proteins in *T. maritima* biofilm cells. Increased expression was noted for genes encoding homologs of FeoB (TM0051, 5.4-fold), which is a G protein-like iron (II)

transport system characterized in several species (2), and FeoA (TM0050, 4.4-fold), also presumed to be involved in iron transport (34). A second putative transporter gene (TM0417, 4.4-fold) related to yeast FTR1 high-affinity Fe^{2+} permeases (97), and the ATP-binding subunit of a putative iron (III) ABC transporter, FepC (TM0191, 2.1-fold), were also induced in biofilm cells. A protein distantly related to bacterial ferritins (TM0560, COG2406) was the most highly downregulated gene in biofilm cells (-11.6-fold), presumably reflecting a decreased need for iron sequestration (5). Iron uptake regulation mechanisms have not been determined experimentally for *T. maritima*, but a small, statistically significant increase in the expression of a ferric uptake regulator (*fur*) homolog was noted (TM0122, 1.5-fold). Sequences resembling Fur binding sites are found upstream of the predicted iron transporter TM0417, and also upstream of TM0122, which precedes a similarly regulated set of ABC transporter components related to metal uptake systems (data not shown).

Genes (TM0483-TM0485) homologous to two *E. coli* K12 ABC transporter systems for sulfonates (25, 113) were preferentially induced in biofilm cells. *E. coli* K12 and *Rhodobacter capsulatus* (63) *tauABC* encode taurine uptake ABC transporters, while the *ssuABC* operon encodes an alkanesulfonate transport system in *E. coli* K12 and *B. subtilis* (25, 112). Although the natural substrates of the two *tauABC*-related systems in *T. maritima* have not been determined, sulfates and cysteine are present in the growth medium. Imported taurine and sulfates are typically incorporated via the cysteine biosynthesis pathway, but no recognizable homolog to the *E. coli* TauD desulfonation enzyme is apparent in the *T. maritima* genome. However, homologs to the uncharacterized conserved ORF (TM0486) are found upstream of *tauABC* homologs in

two *Streptococcus pneumoniae* strains, *Clostridium acetobutylicum*, and *Corynebacterium glutamicum* (96). Crystal structures of two proteins related to TM0486 (pfam01910) suggest a ferredoxin-like fold and a possible role in protein-protein interaction regulated by the binding of sulfate ions (105). Several genes encoding predicted serine and cysteine biosynthesis enzymes were up-regulated here, including cysteine synthase (TM0665, 3.9-fold), serine acetyltransferase (TM0666, 3.9-fold), and a cystathione beta-lyase/ cystathione gamma synthase homolog predicted to be involved in cysteine degradation (TM1270, 2.7 fold) (Figure 2.4).

The up-regulation of genes encoding members of a predicted iron-sulfur cluster chaperone complex offers insight into the apparent need of biofilm cells to acquire iron and sulfur from the environment and increase synthesis of cysteine. Iron-sulfur cluster synthesis and repair in biofilms may be a more general phenomenon, as a recent report indicates the upregulation of the iron-sulfur chaperones *nifSU* in mature biofilms of *E. coli* K12 (7). Three paralogous cysteine desulfurases-IscS, NifS, and SufS-have been characterized in *E. coli* K12 (103). While TM1371 and TM1372 have been referred to as *iscS* and *iscU* in characterization efforts, the lack of other *isc* genes in this genomic region has been noted (8, 59). Recent characterization of the SufABCDES iron-sulfur cluster assembly complex in *E. coli* K12 (103) suggests a more appropriate designation of *sufS* (TM1371) and *sufA* (TM1372), given the co-localization with *sufBCD* homologs and the known role of SufABCDS in iron-sulfur cluster assembly under conditions of iron limitation and oxidative stress in *E. coli* K12 (99) and *Erwinia chrysanthemi* (66). A homolog to SufE, which stimulates the cysteine desulfurase activity of SufS in *E. coli* K12 (72), is not identifiable in *T. maritima*. The proteins encoded by *sufC* (TM1368) and

sufB (TM1369) have been shown to interact in *T. maritima* cells (82); despite the lack of differential expression of the *sufC* homolog, the distantly related *sufB* and *sufD* (TM1370) are both expressed 6.3-fold higher in biofilm. Structural characterization of SufS/IscS (TM1371) has revealed conformational flexibility consistent with a role in iron-sulfur cluster donation to a variety of proteins, while SufA (TM1372) may act as a scaffold for iron-sulfur cluster assembly . A second SufS/IscS homolog in *T. maritima*, previously designated NifS (TM1692) (42), was not differentially expressed here.

Two putative regulators found in the *T. maritima* genome (TM0567, TM1527) bear sequence similarity to IscR, a negative regulator of the Isc "housekeeping" iron-sulfur cluster assembly complex in *E. coli* K12 (94). Three cysteine residues in *E. coli* IscR coordinate a [2Fe-2S] cluster which, when destabilized, disrupts DNA binding to IscR and allows transcription of the Isc operon (94). All three conserved cysteine residues are present in TM1527, located within a biofilm up-regulated gene string encoding FixABCX homologs and a hypothetical protein (TM1534, 2.2-fold) with a conserved CXXC{12}CXXC motif. While the FixABCX proteins of *T. maritima* have not yet been characterized, homologous proteins function in electron transfer chains in other bacteria, including *Rhizobium meliloti* (23), *Rhizobium leguminosarum* (30), *Azorhizobium caulinodans* (43), and *E. coli* K12 (118).

Additional plausible targets for Fe-S cluster assembly complexes are suggested by differential expression data. TM0034 (2.1-fold) contains two cysteine-rich sequence motifs, which are predicted to bind iron-sulfur clusters (see Figure 2.4). Up-regulated genes in a glutamate synthesis operon (TM0394-TM0398) encode a putative NADH oxidase (TM0395) and three domains of glutamate synthase, a multiple iron-sulfur cluster

binding complex (84). Also up-regulated is an iron-sulfur cluster binding protein (TM0396, 3.2-fold) that shares identity (44% id/143 aa) with a carbon monoxide dehydrogenase from *A. fulgidus*. O'Toole and Kolter (75) have shown that glutamate and/or iron containing medium can restore the ability of some biofilm-defective *P. fluorescens* strains to form biofilm.

Two separate predicted operons encoding a number of cofactor biosynthesis enzymes (TM1266-TM1270, TM0787-TM0789) were overexpressed in *T. maritima* biofilm cells. Expression changes for the putative *thi1-thiC* homologs (TM0787-TM0788), which are most closely related to archaeal thiamine biosynthesis enzymes but largely absent from other eubacteria (91), were considerably lower (<2.0-fold) than those of *thiH-bioB-metC* (TM1267, TM1269-TM1270). *E. coli* K12 *iscS*- mutants have been shown to be deficient in thiamine biosynthesis (54), likely as a result of degradation of an iron-sulfur cluster in the ThiH protein (57). *E. coli* K12 ThiH is involved in biosynthesis of the thiamine thiazole ring, a process which requires sulfur donation from cysteine to ThiS via IscS (54, 106). The iron-sulfur cluster binding motif found in *E. coli* K12 ThiH is conserved in the *T. maritima* ThiH homolog (TM1267, 8.9-fold).

A connection to iron-sulfur cluster assembly is also apparent in the up-regulation of genes encoding enzymes involved in nicotinate biosynthesis (TM1643-1645). *E. coli iscS*- mutants have been shown to require NAD as well as thiamine, presumably due to defects in assembly of the iron-sulfur cluster of quinolinate synthetase, NadA (54). A NifS cysteine desulfurase homolog has also been shown to be required for NAD biosynthesis in *B. subtilis* (100). Expression increases were observed here for genes encoding NadA (TM1644, 7.1-fold), NadC (TM1645, 8.1-fold), and an $NADP^+$-

dependent L-aspartate dehydrogenase (TM1643, 27.9-fold) recently shown to convert L-aspartate to iminoaspartate (126) as an alternative to an NadB-type L-aspartate oxidase in the *T. maritima* NAD biosynthesis pathway. Increased NAD and/or NADH pools in sessile *T. maritima* may relate to the up-regulation of genes encoding L-lactate dehydrogenase (TM1867, 5.1-fold) (74), putative NADH oxidases (TM0379, 2.3-fold; TM0395, 3.5-fold), and the predicted dihydrolipoamide dehydrogenase (TM0381, 1.7-fold). L-lactate dehydrogenase induction has also been observed in *B. cereus* biofilms (73), and may be involved in regenerating NAD^+ (68) in conjunction with NADH oxidases (39, 119).

Regulation of biofilm formation and maintenance. The genome sequence of *T. maritima* reveals the apparent lack of orthologs to a number of known biofilm-induced regulators including RpoS, BmrAB, and CpsR. However, putative transcriptional regulators induced in sessile *T. maritima* cells included the sensor histidine kinase TM0187 (2.2-fold) and the response regulator TM1360 (1.6-fold). While the role of these proteins have not yet been determined, the importance of a variety of related proteins in signaling processes during mesophilic biofilm formation has been well-established (56). Small but statistically-significant expression changes (1.8-fold) were also observed for TM0842, a CheY-related response regulator, and TM0841, a similarly regulated S-layer-like array protein sharing 35% id/460 aa with *Thermus thermophilus* SlpM (71), an activator of bacterial cell surface-layer protein synthesis.

Regulation of sigma factor expression influences biofilm formation in a number of species (1, 130). Little functional information is available for *T. maritima* sigma

factors (13); however, the up-regulation of homologs to *sig*A (TM1451, 2.0-fold) and *sig*E (TM1598, 1.6-fold) in biofilm cells hints at a possible role for these proteins as global regulators during *T. maritima* biofilm formation. In contrast, the only two other *T. maritima* sigma factor homologs, *sigH* (TM0534) and *fli*A (TM0902), showed little fluctuation in expression levels for the two conditions compared here. Putative regulatory proteins induced in biofilm cells included members of the LytR (TM1866, 2.3-fold), biotin repressor (TM1602, 2.4-fold), and AcrR/TetR (TM0823, 2.7-fold) families. None of these proteins have been characterized in *T. maritima*, although regulators with TetR DNA binding domains have previously been shown to be important in biofilm formation in mesophiles (17, 49).

No function is known for the predicted transcriptional regulator TM1602 (2.4-fold); however, the major facilitator superfamily permease (TM1603) bears sequence similarity to a *B. subtilis* transporter conferring resistance to the toxic oxyanion tellurite ($TeO_3(^{2-})$) (51) (Table 2.1). The IscS cysteine desulfurase and the CysK cysteine synthase of *Geobacillus stearothermophilis* also confer tellurite resistance on *E. coli K12*, presumably protecting cells from superoxide-mediated iron-sulfur cluster degradation (104, 115). Complementation of a tellurite hypersentitive *E. coli* K12 *iscS*- mutant with *G. stearothermophilus iscS* confers tellurite resistance and relieves a growth requirement for thiamine but not nicotinic acid (104). The isolation of a number of tellurite- and selenite-resistant strains of bacteria from hydrothermal vents near sulfide rocks and bacterial biofilms suggests that the expression changes observed here may indicate an adaptive response to an iron-sulfur cluster degradation stimulus in the natural environment of *T. maritima* (83).

Summary. The complex nature of biofilm formation processes, the dynamic physical and chemical characteristics of these microenvironments, and the likely heterogeneity of cellular states comprising biofilm populations make assigning a definitive biofilm phenotype difficult for *T. maritima*. Nonetheless, clear transcriptional differences were ascertained here that relate to cells involved in surface colonization. There is still much to be understood about biofilm formation and dynamics for *T. maritima*, but this work provides evidence for biofilm formation by *T. maritima*; a methodology for generating sufficient biofilm populations on nylon mesh in a high-temperature anaerobic chemostat for subsequent investigation of transcriptional response comparing planktonic and sessile cells; as well as a list of candidate genes whose expression patterns suggest a role in this process.

ACKNOWLEDGMENTS

This work was supported in part by grants from the National Science Foundation, NASA Exobiology Program and the Department of Energy (Energy Biosciences Program). SBC acknowledges support from an NIEHS Traineeship in Bioinformatics. The authors wish to thank M. Dykstra at the Electron Microscopy Center, NCSU School of Veterinary Medicine, for assistance with electron microscopy, and Stephanie Bridger, Ubie Sullivan, Jennifer Strayhorn and Leon Kluskens for their assistance in generation of the PCR products used to construct the array. The authors also thank R. Wolfinger and K. Scott, SAS Institute, Cary, NC, for help with implementing the mixed model analysis and the NCSU Genome Research Laboratory for assistance with microarray development and use.

CHAPTER 2 REFERENCES

1. **Adams, J. L., and R. J. McLean.** 1999. Impact of *rpoS* deletion on *Escherichia coli* biofilms. Appl Environ Microbiol **65:**4285-4287.
2. **Andrews, S. C., A. K. Robinson, and F. Rodriguez-Quinones.** 2003. Bacterial iron homeostasis. FEMS Microbiol Rev **27:**215-237.
3. **Anton, J., I. Meseguer, and F. Rodriguezvalera.** 1988. Production of an extracellular polysaccharide by *Haloferax mediterranei*. Appl Environ Microbiol **54:**2381-2386.
4. **Arigoni, F., P. A. Kaminski, H. Hennecke, and C. Elmerich.** 1991. Nucleotide sequence of the *fixABC* region of *Azorhizobium caulinodans* ORS571: similarity of the *fixB* product with eukaryotic flavoproteins, characterization of *fixX*, and identification of *nifW*. Mol Gen Genet **225:**514-520.
5. **Baaghil, S., A. Lewin, G. R. Moore, and N. E. Le Brun.** 2003. Core Formation in *Escherichia coli* bacterioferritin requires a functional ferroxidase center. Biochemistry **42:**14047-14056.
6. **Beam, C. E., C. J. Saveson, and S. T. Lovett.** 2002. Role for *radA/sms* in recombination intermediate processing in *Escherichia coli*. J Bacteriol **184:**6836-6844.
7. **Beloin, C., J. Valle, P. Latour-Lambert, P. Faure, M. Kzreminski, D. Balestrino, J. A. Haagensen, S. Molin, G. Prensier, B. Arbeille, and J. M. Ghigo.** 2004. Global impact of mature biofilm lifestyle on *Escherichia coli* K-12 gene expression. Mol Microbiol **51:**659-674.

8. **Bertini, I., J. A. Cowan, C. Del Bianco, C. Luchinat, and S. S. Mansy.** 2003. *Thermotoga maritima* IscU. Structural characterization and dynamics of a new class of metallochaperone. J Mol Biol **331:**907-924.

9. **Blake, P. R., J. B. Park, F. O. Bryant, S. Aono, J. K. Magnuson, E. Eccleston, J. B. Howard, M. F. Summers, and M. W. Adams.** 1991. Determinants of protein hyperthermostability: purification and amino acid sequence of rubredoxin from the hyperthermophilic archaebacterium *Pyrococcus furiosus* and secondary structure of the zinc adduct by NMR. Biochemistry **30:**10885-10895.

10. **Bollinger, N., D. J. Hassett, B. H. Iglewski, J. W. Costerton, and T. R. McDermott.** 2001. Gene expression in *Pseudomonas aeruginosa*: Evidence of iron override effects on quorum sensing and biofilm-specific gene regulation. J Bacteriol **183:**1990-1996.

11. **Burns, D. M., and I. R. Beacham.** 1986. Nucleotide sequence and transcriptional analysis of the *E. coli ushA* gene, encoding periplasmic UDP-sugar hydrolase (5'-nucleotidase): regulation of the *ushA* gene, and the signal sequence of its encoded protein product. Nucleic Acids Res **14:**4325-4342.

12. **Caldwell, D. E., and J. W. Costerton.** 1996. Are bacterial biofilms constrained to Darwin's concept of evolution through natural selection? Microbiologia SEM **12:**347-358.

13. **Camarero, J. A., A. Shekhtman, E. A. Campbell, M. Chlenov, T. M. Gruber, D. A. Bryant, S. A. Darst, D. Cowburn, and T. W. Muir.** 2002. Autoregulation of a bacterial sigma factor explored by using segmental isotopic labeling and NMR. Proc Natl Acad Sci U S A **99:**8536-8541.

14. **Chhabra, S. R., K. R. Shockley, S. B. Conners, K. L. Scott, R. D. Wolfinger, and R. M. Kelly.** 2003. Mixed model analysis of carbohydrate-induced differential expression patterns in the hyperthermophilic bacterium *Thermotoga maritima*. J Biol Chem **278:**7540-7552.

15. **Chhabra, S. R., K. R. Shockley, D. E. Ward, and R. M. Kelly.** 2002. Regulation of endo-acting glycosyl hydrolases in the hyperthermophilic bacterium *Thermotoga maritima* grown on glucan- and mannan-based polysaccharides. Appl Environ Microbiol **68:**545-554.

16. **Clay, M. D., C. A. Cosper, F. E. Jenney, Jr., M. W. Adams, and M. K. Johnson.** 2003. Nitric oxide binding at the mononuclear active site of reduced *Pyrococcus furiosus* superoxide reductase. Proc Natl Acad Sci U S A **100:**3796-3801.

17. **Conlon, K. M., H. Humphreys, and J. P. O'Gara.** 2002. *icaR* encodes a transcriptional repressor involved in environmental regulation of *ica* operon expression and biofilm formation in *Staphylococcus epidermidis*. J Bacteriol **184:**4400-4408.

18. **Dakhova, O. N., N. E. Kurepina, V. V. Zverlov, V. A. Svetlichnyi, and G. A. Velikodvorskaya.** 1993. Cloning and expression in *Escherichia coli* of *Thermotoga neapolitana* genes coding for enzymes of carbohydrate substrate degradation. Biochem Biophys Res Commun **194:**1359-1364.

19. **Davies, D. G., A. M. Chakrabarty, and G. G. Geesey.** 1993. Exopolysaccharide production in biofilms: substratum activation of alginate gene expression by *Pseudomonas aeruginosa*. Appl Environ Microbiol **59:**1181-1186.

20. **Davies, D. G., and G. G. Geesey.** 1995. Regulation of the alginate biosynthesis gene *algC* in *Pseudomonas aeruginosa* during biofilm development in continuous culture. Appl Environ Microbiol **61:**860-867.

21. **Day, M. W., B. T. Hsu, L. Joshua-Tor, J. B. Park, Z. H. Zhou, M. W. Adams, and D. C. Rees.** 1992. X-ray crystal structures of the oxidized and reduced forms of the rubredoxin from the marine hyperthermophilic archaebacterium *Pyrococcus furiosus*. Protein Sci **1:**1494-1507.

22. **Domingues, M. R., J. C. Araujo, M. B. A. Varesche, and R. F. Vazoller.** 2002. Evaluation of thermophilic anaerobic microbial consortia using fluorescence *in situ* hybridization (FISH). Water Sci Technol **45:**27-33.

23. **Donald, R. G., D. W. Nees, C. K. Raymond, A. I. Loroch, and R. A. Ludwig.** 1986. Characterization of three genomic loci encoding *Rhizobium sp.* strain ORS571 N2 fixation genes. J Bacteriol **165:**72-81.

24. **Donlan, R. M.** 2000. Biofilm control in industrial water systems: Approaching an old problem in new ways, p. 333-360. *In* L. V. Evans (ed.), Biofilms: Recent advances in their study and control. Harwood Academic Publishers, Singapore.

25. **Eichhorn, E., J. R. van der Ploeg, and T. Leisinger.** 2000. Deletion analysis of the *Escherichia coli* taurine and alkanesulfonate transport systems. J Bacteriol **182:**2687-2695.

26. **Fiegler, H., and R. Bruckner.** 1997. Identification of the serine acetyltransferase gene of *Staphylococcus xylosus*. FEMS Microbiol Lett **148:**181-187.

27. **Frazao, C., G. Silva, C. M. Gomes, P. Matias, R. Coelho, L. Sieker, S. Macedo, M. Y. Liu, S. Oliveira, M. Teixeira, A. V. Xavier, C. Rodrigues-**

Pousada, M. A. Carrondo, and J. Le Gall. 2000. Structure of a dioxygen reduction enzyme from *Desulfovibrio gigas*. Nat Struct Biol **7:**1041-1045.

28. **Ghigo, J. M.** 2003. Are there biofilm-specific physiological pathways beyond a reasonable doubt? Res Microbiol **154:**1-8.

29. **Golovlev, E. L.** 2002. The mechanism of formation of *Pseudomonas aeruginosa* biofilm, a type of structured population. Microbiology **71:**249-254.

30. **Gronger, P., S. S. Manian, H. Reilander, M. O'Connell, U. B. Priefer, and A. Puhler.** 1987. Organization and partial sequence of a DNA region of the *Rhizobium leguminosarum* symbiotic plasmid pRL6JI containing the genes *fixABC*, *nifA*, *nifB* and a novel open reading frame. Nucleic Acids Res **15:**31-49.

31. **Guezennec, J.** 1999. Microbial exopolysaccharides from extreme environments. Agro Food Industry Hi-Tech **10:**34-35.

32. **Guo, R. T., C. J. Kuo, C. C. Chou, T. P. Ko, H. L. Shr, P. H. Liang, and A. H. Wang.** 2003. Crystal structure of octaprenyl pyrophosphate synthase from hyperthermophilic *Thermotoga maritima* and mechanism of product chain length determination. J Biol Chem.

33. **Guo, R. T., C. J. Kuo, C. C. Chou, T. P. Ko, H. L. Shr, P. H. Liang, and A. H. Wang.** 2004. Crystal structure of octaprenyl pyrophosphate synthase from hyperthermophilic *Thermotoga maritima* and mechanism of product chain length determination. J Biol Chem **279:**4903-4912.

34. **Hantke, K.** 2003. Is the bacterial ferrous iron transporter FeoB a living fossil? Trends Microbiol **11:**192-195.

35. **Hasseman, J.** 2001, posting date. TIGR microarray protocols. [Online.]

36. **Hinsa, S. M., M. Espinosa-Urgel, J. L. Ramos, and G. A. O'Toole.** 2003. Transition from reversible to irreversible attachment during biofilm formation by *Pseudomonas fluorescens* WCS365 requires an ABC transporter and a large secreted protein. Mol Microbiol **49:**905-918.

37. **Hirata, H., T. Fukazawa, S. Negoro, and H. Okada.** 1986. Structure of a beta-galactosidase gene of *Bacillus stearothermophilus*. J Bacteriol **166:**722-727.

38. **Huber, R., T. A. Langworthy, H. Konig, M. Thomm, C. R. Woese, U. B. Sleytr, and K. O. Stetter.** 1986. *Thermotoga maritima* sp. nov. represents a new genus of unique extremely thermophilic eubacteria growing up to 90 DegreesC. Arch Microbiol **144:**324-333.

39. **Jenney, F., M. Verhage, X. Cui, and M. W. W. Adams.** 1999. Anaerobic microbes: oxygen detoxification without superoxide dismutase. Science **286:**306-309.

40. **Jeong, W., M. K. Cha, and I. H. Kim.** 2000. Thioredoxin-dependent hydroperoxide peroxidase activity of bacterioferritin comigratory protein (BCP) as a new member of the thiol-specific antioxidant protein (TSA)/Alkyl hydroperoxide peroxidase C (AhpC) family. J Biol Chem **275:**2924-2930.

41. **Kacklany, S. C., S. B. Levery, J. S. Kim, B. L. Reuhs, L. W. Lion, and W. C. Ghiorse.** 2001. Structure and carbohydrate analysis of the exopolysaccharide capsule of *Pseudomonas putida* G7. Environ Microbiol **3:**774-784.

42. **Kaiser, J. T., T. Clausen, G. P. Bourenkow, H. D. Bartunik, S. Steinbacher, and R. Huber.** 2000. Crystal structure of a NifS-like protein from *Thermotoga maritima*: implications for iron sulphur cluster assembly. J Mol Biol **297:**451-464.

43. **Kaminski, P. A., F. Norel, N. Desnoues, A. Kush, G. Salzano, and C. Elmerich.** 1988. Characterization of the *fixABC* region of *Azorhizobium caulinodans* ORS571 and identification of a new nitrogen fixation gene. Mol Gen Genet **214:**496-502.

44. **Kammler, M., C. Schon, and K. Hantke.** 1993. Characterization of the ferrous iron uptake system of *Escherichia coli*. J Bacteriol **175:**6212-6219.

45. **Kashima, Y., and K. Ishikawa.** 2003. Alkyl hydroperoxide reductase dependent on thioredoxin-like protein from *Pyrococcus horikoshii*. J Biochem (Tokyo) **134:**25-29.

46. **Kashiwagi, K., S. Miyamoto, E. Nukui, H. Kobayashi, and K. Igarashi.** 1993. Functions of PotA and PotD proteins in spermidine-preferential uptake system in *Escherichia coli*. J Biol Chem **268:**19358-19363.

47. **Kim, D. Y., D. R. Kim, S. C. Ha, N. K. Lokanath, C. J. Lee, H. Y. Hwang, and K. K. Kim.** 2003. Crystal structure of the protease domain of a heat-shock protein HtrA from *Thermotoga maritima*. J Biol Chem **278:**6543-6551.

48. **Kim, D. Y., and K. K. Kim.** 2002. Crystallization and preliminary X-ray studies of the protease domain of the heat-shock protein HtrA from *Thermotoga maritima*. Acta Crystallogr D Biol Crystallogr **58:**170-172.

49. **Kojic, M., and V. Venturi.** 2001. Regulation of *rpoS* gene expression in *Pseudomonas*: Involvement of a TetR family regulator. J Bacteriol **183:**3712-3720.

50. **Kolter, R., and R. Losick.** 1998. One for all and all for one. Science **280:**226-227.

51. **Kumano, M., A. Tamakoshi, and K. Yamane.** 1997. A 32 kb nucleotide sequence from the region of the lincomycin-resistance gene (22 degrees-25 degrees) of the *Bacillus subtilis* chromosome and identification of the site of the lin-2 mutation. Microbiology **143 (Pt 8):**2775-2782.

52. **Kuo, T. H., and P. H. Liang.** 2002. Reaction kinetic pathway of the recombinant octaprenyl pyrophosphate synthase from *Thermotoga maritima*: how is it different from that of the mesophilic enzyme. Biochim Biophys Acta **1599:**125-133.

53. **LaPaglia, C., and P. L. Hartzell.** 1997. Stress-induced production of biofilm in the hyperthermophile *Archaeoglobus fulgidus*. Appl Environ Microbiol **63:**3158-3163.

54. **Lauhon, C. T., and R. Kambampati.** 2000. The *iscS* gene in *Escherichia coli* is required for the biosynthesis of 4-thiouridine, thiamin, and NAD. J Biol Chem **275:**20096-20103.

55. **Lee, M. H., Y. W. Kim, T. J. Kim, C. S. Park, J. W. Kim, T. W. Moon, and K. H. Park.** 2002. A novel amylolytic enzyme from *Thermotoga maritima*, resembling cyclodextrinase and alpha-glucosidase, that liberates glucose from the reducing end of the substrates. Biochem Biophys Res Commun **295:**818-825.

56. **Lejeune, P.** 2003. Contamination of abiotic surfaces: what a colonizing bacterium sees and how to blur it. Trends Microbiol **11:**179-184.

57. **Leonardi, R., S. A. Fairhurst, M. Kriek, D. J. Lowe, and P. L. Roach.** 2003. Thiamine biosynthesis in *Escherichia coli*: isolation and initial characterisation of the ThiGH complex. FEBS Lett **539:**95-99.

58. **Liebl, W., P. Ruile, K. Bronnenmeier, K. Riedel, F. Lottspeich, and I. Greif.** 1996. Analysis of a *Thermotoga maritima* DNA fragment encoding two similar thermostable cellulases, CelA and CelB, and characterization of the recombinant enzymes. Microbiology **142 (Pt 9):**2533-2542.

59. **Mansy, S. S., G. Wu, K. K. Surerus, and J. A. Cowan.** 2002. Iron-sulfur cluster biosynthesis. *Thermotoga maritima* IscU is a structured iron-sulfur cluster assembly protein. J Biol Chem **277:**21397-21404.

60. **Mansy, S. S., S. P. Wu, and J. A. Cowan.** 2003. Iron-sulfur cluster biosynthesis. Biochemical characterization of the conformational dynamics of *Thermotoga maritima* IscU and the relevance for cellular cluster assembly. J Biol Chem.

61. **Mansy, S. S., S. P. Wu, and J. A. Cowan.** 2004. Iron-sulfur cluster biosynthesis: biochemical characterization of the conformational dynamics of *Thermotoga maritima* IscU and the relevance for cellular cluster assembly. J Biol Chem **279:**10469-10475.

62. **Marlovits, T. C., W. Haase, C. Herrmann, S. G. Aller, and V. M. Unger.** 2002. The membrane protein FeoB contains an intramolecular G protein essential for Fe(II) uptake in bacteria. Proc Natl Acad Sci U S A **99:**16243-16248.

63. **Masepohl, B., F. Fuhrer, and W. Klipp.** 2001. Genetic analysis of a *Rhodobacter capsulatus* gene region involved in utilization of taurine as a sulfur source. FEMS Microbiol Lett **205:**105-111.

64. **Michelini, E. T., and G. C. Flynn.** 1999. The unique chaperone operon of *Thermotoga maritima*: cloning and initial characterization of a functional Hsp70 and small heat shock protein. J Bacteriol **181:**4237-4244.

65. **Muralidharan, V., K. D. Rinker, I. S. Hirsh, E. J. Bouwer, and R. M. Kelly.** 1997. Hydrogen transfer between methanogens and fermentative heterotrophs in hyperthermophilic cocultures. Biotechnol Bioeng **56:**268-278.

66. **Nachin, L., L. Loiseau, D. Expert, and F. Barras.** 2003. SufC: an unorthodox cytoplasmic ABC/ATPase required for [Fe-S] biogenesis under oxidative stress. EMBO J **22:**427-437.

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

68. **Neves, A. R., R. Ventura, N. Mansour, C. Shearman, M. J. Gasson, C. Maycock, A. Ramos, and H. Santos.** 2002. Is the glycolytic flux in *Lactococcus lactis* primarily controlled by the redox charge? Kinetics of NAD+ and NADH pools determined by 13C NMR. J Biol Chem **277:**28088-28098.

69. **Nicolaus, B., M. C. Manca, I. Romano, and L. Lama.** 1993. Production of an exopolysaccharide from two thermophilic archaea belonging to the genus *Sulfolobus*. FEMS Microbiol Lett **109:**203-206.

70. **Nidetzky, B., C. Eis, and M. Albert.** 2000. Role of non-covalent enzyme-substrate interactions in the reaction catalysed by cellobiose phosphorylase from *Cellulomonas uda*. Biochem J **351:**649-659.

71. **Olabarria, G., L. A. Fernandez-Herrero, J. L. Carrascosa, and J. Berenguer.** 1996. *slpM*, a gene coding for an "S-layer-like array" overexpressed in S-layer mutants of *Thermus thermophilus* HB8. J Bacteriol **178:**357-365.

72. **Ollagnier-de-Choudens, S., D. Lascoux, L. Loiseau, F. Barras, E. Forest, and M. Fontecave.** 2003. Mechanistic studies of the SufS-SufE cysteine desulfurase: evidence for sulfur transfer from SufS to SufE. FEBS Lett **555:**263-267.

73. **Oosthuizen, M. C., B. Steyn, J. Theron, P. Cosette, D. Lindsay, A. von Holy, and V. S. Brozel.** 2002. Proteomic analysis reveals differential protein expression by *Bacillus cereus* during biofilm formation. Appl Environ Microbiol **68:**2770-2780.

74. **Ostendorp, R., W. Liebl, H. Schurig, and R. Jaenicke.** 1993. The L-lactate dehydrogenase gene of the hyperthermophilic bacterium *Thermotoga maritima* cloned by complementation in *Escherichia coli*. Eur J Biochem **216:**709-715.

75. **O'Toole, G. A., and R. Kolter.** 1998. Initiation of biofilm formation in *Pseudomonas fluorescens* WCS365 proceeds via multiple, convergent signalling pathways: a genetic analysis. Mol Microbiol **28:**449-461.

76. **Pan, G., A. L. Menon, and M. W. Adams.** 2003. Characterization of a [2Fe-2S] protein encoded in the iron-hydrogenase operon of *Thermotoga maritima*. J Biol Inorg Chem **8:**469-474.

77. **Park, S., and J. A. Imlay.** 2003. High levels of intracellular cysteine promote oxidative DNA damage by driving the fenton reaction. J Bacteriol **185:**1942-1950.

78. **Phadtare, S., J. Hwang, K. Severinov, and M. Inouye.** 2003. CspB and CspL, thermostable cold-shock proteins from *Thermotoga maritima*. Genes Cells **8:**801-810.

79. **Pysz, M. A., K. D. Rinker, K. R. Shockley, and R. M. Kelly.** 2001. Continuous cultivation of hyperthermophiles, p. 31-40, Hyperthermophilic Enzymes, Pt a, vol. 330.

80. **Pysz, M. A., D. E. Ward, K. R. Shockley, C. I. Montero, S. B. Conners, M. R. Johnson, and R. M. Kelly.** 2004. Transcriptional analysis of dynamic heat-shock response by the hyperthermophilic bacterium *Thermotoga maritima*. Extremophiles (EPub Feb 27).

81. **Rajashekhara, E., M. Kitaoka, Y. K. Kim, and K. Hayashi.** 2002. Characterization of a cellobiose phosphorylase from a hyperthermophilic eubacterium, *Thermotoga maritima* MSB8. Biosci Biotechnol Biochem **66:**2578-2586.

82. **Rangachari, K., C. T. Davis, J. F. Eccleston, E. M. Hirst, J. W. Saldanha, M. Strath, and R. J. Wilson.** 2002. SufC hydrolyzes ATP and interacts with SufB from *Thermotoga maritima*. FEBS Lett **514:**225-228.

83. **Rathgeber, C., N. Yurkova, E. Stackebrandt, J. T. Beatty, and V. Yurkov.** 2002. Isolation of tellurite- and selenite-resistant bacteria from hydrothermal

vents of the Juan de Fuca Ridge in the Pacific Ocean. Appl Environ Microbiol **68:**4613-4622.

84. **Ravasio, S., B. Curti, and M. A. Vanoni.** 2001. Determination of the midpoint potential of the FAD and FMN flavin cofactors and of the 3Fe-4S cluster of glutamate synthase. Biochemistry **40:**5533-5541.

85. **Reysenbach, A. L., and E. Shock.** 2002. Merging genomes with geochemistry in hydrothermal ecosystems. Science **296:**1077-1082.

86. **Rinker, K. D.** 1998. Growth Physiology and Bioenergetics of the Hyperthermophilic Archaeon *Thermococcus litoralis* and Bacterium *Thermotoga maritima*. PhD. North Carolina State University, Raleigh.

87. **Rinker, K. D., C. J. Han, and R. M. Kelly.** 1999. Continuous culture as a tool for investigating the growth physiology of heterotrophic hyperthermophiles and extreme thermoacidophiles. J Appl Microbiol **85:**118-127.

88. **Rinker, K. D., and R. M. Kelly.** 2000. Effect of carbon and nitrogen sources on growth dynamics and exopolysaccharide production for the hyperthermophilic archaeon *Thermococcus litoralis* and bacterium *Thermotoga maritima*. Biotech Bioeng **69:**537-547.

89. **Rinker, K. D., and R. M. Kelly.** 1996. Growth physiology of the hyperthermophilic archaeon *Thermococcus litoralis*: Development of a sulfur-free defined medium, characterization of an exopolysaccharide, and evidence of biofilm formation. Appl Environ Microbiol **62:**4478-4485.

90. **Sauer, K.** 2003. The genomics and proteomics of biofilm formation. Genome Biol **4:**219.

91. **Sauer, K., and A. K. Camper.** 2001. Characterization of phenotypic changes in *Pseudomonas putida* in response to surface-associated growth. J Bacteriol **183:**6579-6589.

92. **Sauer, K., A. K. Camper, G. D. Ehrlich, J. W. Costerton, and D. G. Davies.** 2002. *Pseudomonas aeruginosa* displays multiple phenotypes during development as a biofilm. J Bacteriol **184:**1140-1154.

93. **Schembri, M. A., K. Kjaergaard, and P. Klemm.** 2003. Global gene expression in *Escherichia coli* biofilms. Mol Microbiol **48:**253-267.

94. **Schwartz, C. J., J. L. Giel, T. Patschkowski, C. Luther, F. J. Ruzicka, H. Beinert, and P. J. Kiley.** 2001. IscR, an Fe-S cluster-containing transcription factor, represses expression of *Escherichia coli* genes encoding Fe-S cluster assembly proteins. Proc Natl Acad Sci U S A **98:**14895-14900.

95. **Sharma, R., C. Rensing, B. P. Rosen, and B. Mitra.** 2000. The ATP hydrolytic activity of purified ZntA, a Pb(II)/Cd(II)/Zn(II)-translocating ATPase from *Escherichia coli*. J Biol Chem **275:**3873-3878.

96. **Snel, B., G. Lehmann, P. Bork, and M. A. Huynen.** 2000. STRING: a web-server to retrieve and display the repeatedly occurring neighbourhood of a gene. Nucleic Acids Res **28:**3442-3444.

97. **Stearman, R., D. S. Yuan, Y. Yamaguchi-Iwai, R. D. Klausner, and A. Dancis.** 1996. A permease-oxidase complex involved in high-affinity iron uptake in yeast. Science **271:**1552-1557.

98. **Stoodley, P., K. Sauer, D. G. Davies, and J. W. Costerton.** 2002. Biofilms as complex differentiated communities. Ann Rev Microbiology **56:**187-209.

99. **Storz, G., and J. A. Imlay.** 1999. Oxidative stress. Curr Opin Microbiol **2:**188-194.

100. **Sun, D., and P. Setlow.** 1993. Cloning, nucleotide sequence, and regulation of the *Bacillus subtilis nadB* gene and a *nifS*-like gene, both of which are essential for NAD biosynthesis. J Bacteriol **175:**1423-1432.

101. **Suresh, C., M. Kitaoka, and K. Hayashi.** 2003. A thermostable non-xylanolytic alpha-glucuronidase of *Thermotoga maritima* MSB8. Biosci Biotechnol Biochem **67:**2359-2364.

102. **Svensater, G., J. Welin, J. C. Wilkins, D. Beighton, and I. R. Hamilton.** 2001. Protein expression by planktonic and biofilm cells of *Streptococcus mutans*. FEMS Microbiol Lett **205:**139-146.

103. **Takahashi, Y., and U. Tokumoto.** 2002. A third bacterial system for the assembly of iron-sulfur clusters with homologs in archaea and plastids. J Biol Chem **277:**28380-28383.

104. **Tantalean, J. C., M. A. Araya, C. P. Saavedra, D. E. Fuentes, J. M. Perez, I. L. Calderon, P. Youderian, and C. C. Vasquez.** 2003. The *Geobacillus stearothermophilus* V *iscS* gene, encoding cysteine desulfurase, confers resistance to potassium tellurite in *Escherichia coli* K-12. J Bacteriol **185:**5831-5837.

105. **Tao, X., R. Khayat, D. Christendat, A. Savchenko, X. Xu, S. Goldsmith-Fischman, B. Honig, A. Edwards, C. H. Arrowsmith, and L. Tong.** 2003. Crystal structures of MTH1187 and its yeast ortholog YBL001c. Proteins **52:**478-480.

106. **Taylor, S. V., N. L. Kelleher, C. Kinsland, H. J. Chiu, C. A. Costello, A. D. Backstrom, F. W. McLafferty, and T. P. Begley.** 1998. Thiamin biosynthesis in *Escherichia coli*. Identification of this thiocarboxylate as the immediate sulfur donor in the thiazole formation. J Biol Chem **273:**16555-16560.

107. **Tolker-Nielsen, T., U. C. Brinch, P. C. Ragas, J. B. Andersen, C. S. Jacobsen, and S. Molin.** 2000. Development and dynamics of *Pseudomonas* sp. biofilms. J Bacteriol **182:**6482-6489.

108. **Tremoulet, F., O. Duche, A. Namane, B. Martinie, and J. C. Labadie.** 2002. A proteomic study of *Escherichia coli* O157:H7 NCTC 12900 cultivated in biofilm or in planktonic growth mode. FEMS Microbiol Lett **215:**7-14.

109. **Tremoulet, F., O. Duche, A. Namane, B. Martinie, The European *Listeria* Genome Consortium, and J. C. Labadie.** 2002. Comparison of protein patterns of *Listeria monocytogenes* grown in biofilm or in planktonic mode by proteome analysis. FEMS Microbiol Lett **210:**25-31.

110. **Ugulava, N. B., B. R. Gibney, and J. T. Jarrett.** 2001. Biotin synthase contains two distinct iron-sulfur cluster binding sites: chemical and spectroelectrochemical analysis of iron-sulfur cluster interconversions. Biochemistry **40:**8343-8351.

111. **Valdes-Stauber, N., and S. Scherer.** 1996. Nucleotide sequence and taxonomical distribution of the bacteriocin gene *lin* cloned from *Brevibacterium linens* M18. Appl Environ Microbiol **62:**1283-1286.

112. **van der Ploeg, J. R., R. Iwanicka-Nowicka, T. Bykowski, M. M. Hryniewicz, and T. Leisinger.** 1999. The *Escherichia coli ssuEADCB* gene cluster is required

for the utilization of sulfur from aliphatic sulfonates and is regulated by the transcriptional activator Cbl. J Biol Chem **274:**29358-29365.

113. **van der Ploeg, J. R., M. A. Weiss, E. Saller, H. Nashimoto, N. Saito, M. A. Kertesz, and T. Leisinger.** 1996. Identification of sulfate starvation-regulated genes in *Escherichia coli*: a gene cluster involved in the utilization of taurine as a sulfur source. J Bacteriol **178:**5438-5446.

114. **van Kranenburg, R., H. R. Vos, S. van, II, M. Kleerebezem, and W. M. de Vos.** 1999. Functional analysis of glycosyltransferase genes from *Lactococcus lactis* and other gram-positive cocci: complementation, expression, and diversity. J Bacteriol **181:**6347-6353.

115. **Vasquez, C. C., C. P. Saavedra, C. A. Loyola, M. A. Araya, and S. Pichuantes.** 2001. The product of the *cysK* gene of *Bacillus stearothermophilus* V mediates potassium tellurite resistance in *Escherichia coli*. Curr Microbiol **43:**418-423.

116. **Verhagen, M. F., T. O'Rourke, and M. W. Adams.** 1999. The hyperthermophilic bacterium, *Thermotoga maritima*, contains an unusually complex iron-hydrogenase: amino acid sequence analyses versus biochemical characterization. Biochim Biophys Acta **1412:**212-229.

117. **Wakagi, T.** 2003. Sulerythrin, the smallest member of the rubrerythrin family, from a strictly aerobic and thermoacidophilic archaeon, *Sulfolobus tokodaii* strain 7. FEMS Microbiol Lett **222:**33-37.

118. **Walt, A., and M. L. Kahn.** 2002. The *fixA* and *fixB* genes are necessary for anaerobic carnitine reduction in *Escherichia coli*. J Bacteriol **184:**4044-4047.

119. **Ward, D. E., C. J. Donnelly, M. E. Mullendore, J. van der Oost, W. M. de Vos, and E. J. Crane.** 2001. The NADH oxidase from Pyrococcus furiosus. Implications for the protection of anaerobic hyperthermophiles against oxidative stress. Eur J Biochem **268:**5816-5823.

120. **Wassenberg, D., W. Liebl, and R. Jaenicke.** 2000. Maltose-binding protein from the hyperthermophilic bacterium *Thermotoga maritima*: stability and binding properties. J Mol Biol **295:**279-288.

121. **Watnick, P., and R. Kolter.** 2000. Biofilm, city of microbes. J Bacteriol **182:**2675-2679.

122. **Welker, C., G. Bohm, H. Schurig, and R. Jaenicke.** 1999. Cloning, overexpression, purification, and physicochemical characterization of a cold shock protein homolog from the hyperthermophilic bacterium *Thermotoga maritima*. Protein Sci **8:**394-403.

123. **Whiteley, M., M. G. Bangera, R. E. Bumgarner, M. R. Parsek, G. M. Teitzel, S. Lory, and E. P. Greenberg.** 2001. Gene expression in *Pseudomonas aeruginosa* biofilms. Nature **413:**860-864.

124. **Wolfinger, R. D., G. Gibson, E. D. Wolfinger, L. Bennett, H. Hamadeh, P. Bushel, C. Afshari, and R. S. Paules.** 2001. Assessing gene significance from cDNA microarray expression data via mixed models. J Comput Biol **8:**625-637.

125. **Yan, L., K. G. Boyd, D. R. Adams, and J. G. Burgess.** 2003. Biofilm-specific cross-species induction of antimicrobial compounds in bacilli. Appl Environ Microbiol **69:**3719-3727.

126. **Yang, Z., A. Savchenko, A. Yakunin, R. Zhang, A. Edwards, C. Arrowsmith, and L. Tong.** 2003. Aspartate dehydrogenase, a novel enzyme identified from structural and functional studies of TM1643. J Biol Chem **278:**8804-8808.

127. **Yeh, A. P., Y. Hu, F. E. Jenney, Jr., M. W. Adams, and D. C. Rees.** 2000. Structures of the superoxide reductase from *Pyrococcus furiosus* in the oxidized and reduced states. Biochemistry **39:**2499-2508.

128. **Yernool, D. A., J. K. McCarthy, D. E. Eveleigh, and J. D. Bok.** 2000. Cloning and characterization of the glucooligosaccharide catabolic pathway beta-glucan glucohydrolase and cellobiose phosphorylase in the marine hyperthermophile *Thermotoga neapolitana*. J Bacteriol **182:**5172-5179.

129. **Yildiz, F. H., and G. K. Schoolnik.** 1999. *Vibrio cholerae* O1 El Tor: identification of a gene cluster required for the rugose colony type, exopolysaccharide production, chlorine resistance, and biofilm formation. Proc Natl Acad Sci U S A **96:**4028-4033.

130. **Yu, H., M. J. Schurr, and V. Deretic.** 1995. Functional equivalence of *Escherichia coli* Sigma E and *Pseudomonas aeruginosa* AlgU: *E. coli rpoE* restores mucoidy and reduces sensitivity to reactive oxygen intermediates in *algU* mutants of *P. aeruginosa*. J Bacteriol **177:**3259-3268.

TABLE 2.1. *Thermotoga maritima* ORFs differentially regulated in biofilm[a]

Gene description	Source of gene description (reference)	Gene ID	Fold change	-$Log_{10}P$ value
Up-regulated in biofilm[b]				
Aspartate dehydrogenase	(126)	TM1643	27.9	16.3
ThiH protein, putative	28% id/347aa with EC ThiH (54, 106) conserved CX_3CX_2C [4Fe-4S] cluster binding motif	TM1267	8.9	15.4
Nicotinate-nucleotide pyrophosphorylase (NadC)	Crystallized (GI:34811257; 1O4U_A, 1O4U_B)	TM1645	8.1	9.1
		TM1266	7.2	10.0
Hypothetical protein	*(67)*			
Quinolinate synthetase A (NadA)	54% id/298 aa with Pab (PAB2345) and Pho (PH0013), 3 conserved cysteines	TM1644	7.1	11.9
Cellobiose phosphorylase	(81)	TM1848	6.9	12.9
Cation-transporting ATPase, P-type	COG2217, 30%id/724 aa with EC ZntA (95)	TM0317	6.6	14.1
SufD homolog, similar to ABC permease components	COG0719	TM1370	6.3	12.0
SufB homolog, similar to ABC ATP-binding components	(82)	TM1369	6.3	15.0
Iron(II) transport protein B (FeoB)	33% id/718aa with EC FeoB (44, 62)	TM0051	5.4	12.4
L-lactate dehydrogenase	(74)	TM1867	5.1	11.0
Ubiquinone/menaquinone methyltransferase	COG0500, SAM-dependent methyltransferases, 53% id/452aa with PF0738	TM0318	5.0	11.4
Biotin synthetase, putative (BioB homolog)	22% id/281 aa with EC BioB (110), conserved CX_3CX_2C motif, CX_2C motif	TM1269	5.0	8.6
IscS/SufS homolog, cysteine desulfurase	(60, 61)	TM1371	4.8	13.9
		TM0199	4.6	12.4
Sms/RadA homolog, DNA recombination/repair protein	39% id/452 aa with EC Sms/RadA (6), conserved $CX_2C\{X_{10}\}CX_2C$ motif is a zinc finger in EC Sms/RadA			
FTR1, predicted high affinity Fe^{2+}/Pb^{2+} permease	25% id/228 aa with Sac Ftr1p; two conserved REXXE iron-binding motifs (97)	TM0417	4.4	8.5
Esterase, putative	(67)	TM0053	4.1	8.8
Iron(II) transport protein A (FeoA)	42% id/57aa with EC FeoA (44)	TM0050	4.0	12.2
Cysteine synthase (CysK)	Crystal structure (GI:34810052, 1O58A-1O8D)	TM0665	3.9	12.0
Serine acetyltransferase (CysE)	43% id/195 aa with Stx CysE (26)	TM0666	3.9	9.3

(Table 2.1, continued)

FpaA family protein, contains flavodoxin domain and beta-metallo-lactamase domain	COG0426, 28% id/366 aa with Dvg ROO (27)	TM0755	3.8	8.5
Hypothetical protein	(67), 32% id/78 aa with TTE1569	TM1268	3.7	6.6
Uncharacterized homolog of gamma-carboxymucono-lactone decarboxylase	COG0599	TM0316	3.6	9.1
NADH oxidase	46% id/440 aa with Tne NADH: polysulfide oxidoreductase	TM0395	3.5	9.0
Ubiquinone/menaquinone biosynthesis methyltransferase, putative (UbiE homolog)	COG2226	TM0753	3.3	10.1
HycB domain containing protein, related to hydrogenase components	COG1142, 2 $CX_2CX_2CX_3C$ motifs, 1 $CX_2C\{X_7\}CX_2C$ motif, 1 $CX_2CX_4CX_3C$ motif	TM0396	3.2	10.8
Heavy metal binding protein	(67)	TM0320	3.2	9.2
Uncharacterized conserved protein	COG3862, CX_2CX_3C motif	TM1434	3.2	8.5
Predicted zinc-dependent protease	COG2738	TM1511	3.1	11.0
NAD+-dependent α-glucuronidase	(101)	TM0752	3.1	7.9
Protein containing Bcp domain and nitroreductase domain	COG1225 (Bcp domain), COG0778 (Nitroreductase domain)	TM0386	3.0	3.1
Hypothetical protein	(67)	TM0338	3.0	7.9
Hypothetical protein	(67)	TM0052	3.0	10.1
Predicted GTPase	COG1160	TM0445	2.9	10.8
Beta-D-galactosidase	40% id/670 aa with Bst β-galactosidase (37)	TM0310	2.9	9.5
IscU/SufA homolog, iron-sulfur cluster assembly scaffold	3 conserved cysteines bind a [2Fe-2S] cluster (8, 59)	TM1372	2.9	5.9
Uncharacterized conserved protein	COG0011, pfam01910, localized with TauABC (4 species)	TM0486	2.9	9.4
ABC transporter, permease protein, TauC family	COG0600	TM0485	2.9	11.1
ATP-dependent Clp protease, ATPase subunit, ClpA homolog	COG0542	TM0198	2.8	9.2
Glutamate synthase domain 3 (GltB)	COG0070	TM0394	2.8	6.5
Possible endonuclease or sugar phosphate isomerase	COG1082, pfam01261	TM0422	2.8	6.3
Hypothetical protein	(67)	TM0054	2.8	10.2
Predicted membrane protein	COG3462	TM0315	2.8	12.2
Uncharacterized conserved protein	COG0718	TM0687	2.8	7.9
Cystathionine gamma-synthase/beta lyase	COG0626	TM1270	2.7	9.3
Predicted transcriptional regulator, ACR/TetR related	pfam00440	TM0823	2.7	9.2
Sugar ABC transporter, permease protein, UpgA family	COG1175	TM0419	2.7	11.0
Sugar ABC transporter, periplasmic sugar-binding protein, MalE-related	COG2182	TM0418	2.7	6.9

(Table 2.1, continued)

ABC transporter, ATP-binding subunit, TauA family	COG0715	TM0484	2.6	5.9
Hypothetical protein	(67)	TM0714	2.6	9.2
Thioredoxin peroxidase	COG1225, 59% id/157 aa with Tal (Tal1368M)	TM0780	2.6	12.8
Hypothetical protein	(67)	TM0319	2.6	6.6
Predicted HHH nucleic acid binding protein	COG1623, pfam02457	TM0200	2.6	5.7
Glycerol uptake facilitator protein	COG0580	TM1429	2.5	7.6
FixX homolog, putative ferredoxin	COG2440, $CX_3CX_4CX_3CC$ motif, $CX_2CX_2CX_3C$ motif	TM1533	2.5	5.3
FixC homolog	42% id/438 aa with Azc FixC (4)	TM1532	2.5	7.9
PotD, spermidine/putrescine ABC transporter, PBP	39% id/320 aa with EC PotD (46)	TM1375	2.5	2.6
Transcriptional regulator, biotin repressor family	Crystallized (GI:22218828, 1J5YA); 36% id/409 aa with Bsu YceI	TM1602	2.4	9.7
Predicted CAAX amino terminal protease	pfam0257	TM1529	2.4	10.0
RAD55-related, RecA superfamily ATPase	COG0467, 78%id/235aa with PH0824, PAB2180, PH1931	TM0370	2.4	7.4
Pyridine nucleotide-disulphide oxidoreductase, putative	pfam00070	TM0754	2.4	5.9
Oxidoreductase, aldo/keto reductase family	COG0467	TM1743	2.4	7.2
NADH:polysulfide oxidoreductase	86% id/440 with Tne (GI:21702687)	TM0379	2.3	8.5
Hypothetical protein	(67)	TM0002	2.3	9.4
Heat shock operon repressor HrcA	(67, 80)	TM0851	2.3	7.3
Conserved hypothetical protein, GGDEF domain	COG2199, pfam00990	TM1588	2.3	7.5
Glutamate synthase domain 2, GltB	COG0069, pfam01645, CX_2CX_2C motif	TM0397	2.3	7.9
UDP-sugar disphosphatase precursor	38% id/504 aa with EC UshA (11)	TM1878	2.3	7.9
Putative endonuclease	COG1833, pfam01986	TM0664	2.3	10.5
Uncharacterized conserved protein	COG4198, pfam06245	TM1510	2.3	6.1
Membrane bound protein LytR, putative transcriptional regulator	COG1316, pfam03816	TM1866	2.3	9.8
FixA homolog, electron transfer flavoprotein	45% id/241 aa with Azc FixA (4)	TM1530	2.3	1.9
Sugar ABC transporter, ATP-binding protein, MalK homolog	COG3839	TM0421	2.3	11.3
Heat shock protein, class I	(64)	TM0374	2.3	8.5
Conserved hypothetical protein	(67)	TM0387	2.3	7.8
Predicted CoA-binding protein	COG1832	TM1435	2.2	7.3
Conserved hypothetical protein	pfam03706	TM1390	2.2	10.2
Hypothetical protein	(67)	TM1534	2.2	12.3
Frame shift	(67)	TM0621	2.2	5.2
Glutamate synthase domain 1, GltB	COG0067	TM0398	2.2	5.2
Putative sensor histidine kinase	(67)	TM0187	2.2	5.3
Conserved hypothetical protein	pfam01139	TM1357	2.1	8.8

(Table 2.1, continued)

ABC transporter, ATP-binding subunit, SalX domain	COG1136	TM0352	2.1	9.2
Endoglucanase (extracellular)	(15, 58)	TM1525	2.1	8.6
Iron(III) ABC transporter, ATP-binding protein, putative	COG1120, 50% id/242 aa with Pfu (PF0909)	TM0191	2.1	7.3
Putative transcriptional regulator	COG1318, 58% id/150 aa with Pho (PH0283)	TM0369	2.1	8.0
Iron-sulfur cluster-binding protein, putative	COG2768, 2 $CX_2CX_2CX_3C$ motifs (67)	TM0034	2.1	6.0
Heat shock serine protease, periplasmic	(47, 48)	TM0571	2.1	9.3
Hypothetical protein	(67)	TM0003	2.0	8.7
Permease, putative	COG0477, 36% id/409 aa with Bsu YceI (51)	TM1603	2.0	7.7
Octaprenyl pyrophosphate synthase	(32, 33, 52)	TM1535	2.0	7.6
Membrane protein, putative	COG3374	TM1536	2.0	8.5
ABC transporter, ATP-binding protein, TauB family	COG1116	TM0483	2.0	10.2
RNA polymerase σ^A factor	(13, 67)	TM1451	2.0	10.9
Putative Holliday junction resolvase	COG0816	TM1545	2.0	10.9
ABC transporter, permease subunit, SalY family	COG0577	TM0351	2.0	8.9
Down-regulated in biofilm[c]				
K+ channel, beta subunit	COG0667, pfam00248	TM0313	-2.0	9.3
Cyclomaltodextrinase	(55)	TM1835	-2.0	9.6
Putative regulator, XRE-family HTH	COG1917, pfam1381	TM0656	-2.0	6.7
Hypothetical protein	(67)	TM0794	-2.1	6.9
Predicted dehydrogenase	COG0673, 67% id/325 aa with Pfu (PF0554)	TM0312	-2.1	5.6
(3R)-hydroxymyristoyl-(acyl carrier protein) dehydratase	COG0764 (67)	TM0801	-2.1	10.1
Uncharacterized conserved protein	COG3906	TM0606	-2.1	8.5
Bacteriocin	33%id/251 aa to Brl linocin M18 (111)	TM0785	-2.2	11.4
Maltose ABC transporter, permease protein	(120)	TM1836	-2.2	2.0
Predicted glycosyltransferase	COG0438, pfam00534	TM0392	-2.2	7.3
Hypothetical protein	(67)	TM1241	-2.4	7.4
Uncharacterized conserved protein	COG3471	TM0786	-2.4	6.0
Ribosomal protein L7/L12	(67)	TM0457	-2.5	9.3
Ribosomal protein L10	(67)	TM0456	-2.6	9.0
Cold shock protein	(78, 122)	TM1874	-2.8	9.6
Rubredoxin	68% id/51aa with Pfu rubredoxin PF1282 (9, 21), 2 conserved CXXC motifs	TM0659	-3.1	9.9
Superoxide reductase (neelaredoxin)	57% id/128 aa with Pfu SOR (PF1281(39, 127)	TM0658	-3.5	11.4
Rubrerythrin	58% id/165 aa with Afu rubrerythrin (AF1640) (117)	TM0657	-3.7	7.8

(Table 2.1, continued)

NADPH-dependent alkyl hydroperoxide reductase	69% id/214 aa with Pho (PH1217) (45)	TM0807	-6.0	9.4
Protein distantly related to bacterial ferritins	COG2406, 82% id/183 aa with Mta (MA2882)	TM0560	-11.6	14.5

[a] Locus description based on conserved domain searches (CDD, NCBI) and similarity to characterized proteins.

Abbreviations used: Afu, *Archaeoglobus fulgidus*; Azc, *Azorhizobium caulinodans*; Bcp, bacterioferritin comigratory protein; Brl, *Brevibacterium linens*; Bst, *Bacillus stearothermophilus*; Bsu, *Bacillus subtilis*; Dvg, *Desulfovibrio gigas*; EC, *Escherichia coli* K12; Mta, *Methanosarcina acetivorans* str. C2A; Pfu, *Pyrococcus furiosus*; Pho, *Pyrococcus horikoshii*, Pab, *Pyrococcus abyssi*; Tal, *Thermoplasma acidophilum*; Tne, *Thermotoga neapolitana*; Stx, *Staphylococcus xylosus*.

[b] Genes up-regulated 2.0 fold or greater in biofilm. Significance based on Bonferroni-corrected significance criterion with $-\log_{10}$ (P-value) > 4.6.

[c] Genes down-regulated 2.0 fold or greater in biofilm Significance based on Bonferroni-corrected significance criterion with $-\log_{10}$ (P-value) > 4.6.

TABLE 2.2. Differential expression of genes in biofilm-bound cells as related to predicted *T. maritima* operons

Putative function	Gene ID	Gene Description	Fold-Change	-Log_{10}*P* value[c]
	TM0050	Iron(II) transport protein A (FeoA)	4.0	12.2
	TM0051	Iron(II) transport protein B (FeoB)	5.4	12.4
Iron transport[b]	TM0052	Hypothetical protein	3.0	10.1
	TM0053	Esterase, putative	4.1	8.8
	TM0054	Hypothetical protein	2.8	10.2
	TM0198	ATP-dependent Clp protease, ATPase subunit, ClpA family	2.8	9.2
DNA processing[b]	TM0199	Sms/RadA homolog, DNA recombination/repair protein	4.6	12.4
	TM0200	Predicted HHH nucleic acid binding protein	2.6	5.7
	TM0201	NADP-reducing hydrogenase, subunit D, putative	1.5	2.7
	TM0202	TauA homolog	1.7	5.5
	TM0315	Predicted membrane protein	2.8	12.2
Cation transport system[b]	TM0316	Uncharacterized homolog of gamma-carboxymuconolactone decarboxylase subunit	3.6	9.1
	TM0317	Cation-transporting ATPase, P-type	6.6	14.1
	TM0318	Ubiquinone/menaquinone methyltransferase	5.0	11.4
	TM0319	Hypothetical protein	2.6	6.6
	TM0320	Heavy metal binding protein	3.2	9.2
ABC transport[a]	TM0351	ABC transporter, permease subunit, SalY family	2.0	8.9
	TM0352	ABC transporter, ATP-binding subunit, SalX domain	2.1	9.2
	TM0353	Predicted membrane fusion protein (COG0845)	1.4	6.4
DNA Repair[a]	TM0369	Predicted transcriptional regulator	2.1	8.0
	TM0370	RAD55-related, RecA superfamily ATPase (COG0467)	2.4	7.4
Thermal Stress	TM0373	DnaK protein (64)	1.9	8.5
	TM0374	Heat shock protein, class I	2.3	8.5
	TM0394	Glutamate synthase domain 3 (GltB)	2.8	6.5
	TM0395	NADH oxidase	3.5	9.0
Glutamate synthesis[a]	TM0396	HycB domain containing protein, related to hydrogenase components	3.2	10.8
	TM0397	Glutamate synthase domain 2, GltB	2.3	7.9

(Table 2.2, continued)

	TM0398	Glutamate synthase domain 1, GltB	2.2	5.2
	TM0417	FTR1, predicted high affinity Fe^{2+}/Pb^{2+} permease	4.4	8.5
Sugar and iron transport[b]	TM0418	Sugar ABC transporter, periplasmic sugar-binding protein, MalE-related	2.7	6.9
	TM0419	Sugar ABC transporter, permease protein, UpgA family	2.7	11.0
	TM0420	Sugar ABC transporter, permease protein	1.4	7.6
	TM0421	Sugar ABC transporter, ATP-binding protein, MalK homolog	2.3	11.3
	TM0422	Possible endonuclease or sugar phosphate isomerase	2.8	6.3
	TM0483	ABC transporter, ATP-binding protein, TauB family	2.0	10.2
Sulfur compound transport systems[a]	TM0484	ABC transporter, ATP-binding subunit, TauA family	2.6	5.9
	TM0485	ABC transporter, permease protein, TauC family	2.9	11.1
	TM0486	Uncharacterized conserved protein	2.9	9.4
Oxygen detoxification[b]	TM0657	Rubrerythrin	-3.7	7.8
	TM0658	Superoxide reductase (neelaredoxin)	-3.5	11.4
	TM0659	Rubredoxin	-3.1	9.9
Amino acid metabolism[b]	TM0664	Putative endonuclease	2.3	10.5
	TM0665	Cysteine synthase	3.9	12.0
	TM0666	Serine acetyltransferase (CysE)	3.9	9.3
	TM0752	NAD+-dependent α-glucuronidase	3.1	7.9
Oxygen detoxification/ electron transfer[b]	TM0753	Ubiquinone/menaquinone biosynthesis methyltransferase, putative (UbiE homolog)	3.3	10.1
	TM0754	Oxidoreductase	2.4	5.9
	TM0755	FpaA family protein, contains flavodoxin domain and beta-metallo-lactamase domain (COG0426)	3.8	8.5
	TM1266	Hypothetical protein	7.2	10.0
Biotin/thiamine synthesis[b]	TM1267	ThiH protein, putative	8.9	15.4
	TM1268	Hypothetical protein	3.7	6.6
	TM1269	Biotin synthetase, putative (BioB homolog)	5.0	8.6
	TM1270	Cystathionine gamma-synthase/beta lyase	2.7	9.3
	TM1368	SufC homolog, similar to ABC ATP-binding components	1.1	2.0
Iron-sulfur cluster assembly[a]	TM1369	SufB homolog, similar to ABC permease components	6.3	15.0
	TM1370	SufD homolog, similar to ABC permease components	6.3	12.0
	TM1371	SufS/IscS homolog, cysteine desulfurase	4.8	13.9
	TM1372	SufA/IscU homolog, iron-sulfur cluster assembly scaffold	2.9	5.9
Sugar/electron transfer cascade[b]	TM1524	Endoglucanase (intracellular)	1.8	7.1
	TM1525	Endoglucanase (extracellular)	2.1	8.6

(Table 2.2, continued)

	TM1526	Hypothetical protein (COG1633)	1.7	6.3
	TM1527	IscR homolog, putative	1.8	6.0
	TM1528	1,4-dihydroxy-2-naphthoate octaprenyltransferase, MenA homolog	1.7	6.0
	TM1529	Predicted CAAX amino terminal protease	2.4	10.0
	TM1530	FixA homolog, electron transfer flavoprotein	2.3	1.9
	TM1531	FixB homolog, electron transfer flavoprotein	2.0	4.9
	TM1532	FixC homolog	2.5	7.9
	TM1533	FixX homolog	2.5	5.3
	TM1534	Hypothetical protein,	2.2	12.3
	TM1535	Octaprenyl pyrophosphate synthase (OPP)	2.0	7.6
	TM1536	Putative membrane protein	2.0	8.5
Nicotinate synthesis[b]	TM1643	Aspartate dehydrogenase	27.9	16.3
	TM1644	Quinolinate synthetase A (NadA)	7.1	11.9
	TM1645	Nicotinate-nucleotide pyrophosphorylase (NadC)	8.1	9.1
Maltose utilization and transport[b]	TM1834	α-glucosidase	-1.5	5.1
	TM1835	Cyclomaltodextrinase	-2.0	9.6
	TM1836	Maltose ABC transporter, permease protein	-2.2	2.0
	TM1837	Maltose transport system permease protein	-1.9	8.1
	TM1838	Hypothetical protein	-1.4	5.7
	TM1839	Maltose ABC transporter, periplasmic maltose-binding protein	1.0	0.4

[a] Complete operon predicted by www.tigr.org

[b] Partial operon predicted by www.tigr.org

[c] Note that some $-\log_{10}$p-values are below the Bonferroni significance criterion of 4.6.

FIGURE LEGENDS

FIG. 2.1. *Thermotoga maritima* growth in 1.5-L continuous culture at 80°C. Planktonic samples were collected at points designated by (Δ), and biofilm samples were collected at (×).

FIG. 2.2. *Thermotoga maritima* biofilm formation on (A) nylon mesh and reactor walls during continuous cultivation. Polycarbonate filters inserted into the culture showed microcolony formation by (B) epifluorescent microscopy and (C) scanning electron microscopy.

FIG. 2.3. Volcano plot showing differential gene expression in planktonic and biofilm *Thermotoga maritima* cells grown in chemostat culture at 80°C. Horizontal line indicates Bonferroni corrected α=0.05 significance level.

FIG. 2.4. A predicted pathway for iron-sulfur cluster biogenesis in *Thermotoga maritima* biofilm cells. Expression data suggest a number of known iron-sulfur cluster binding proteins and proteins with conserved cysteine-rich motifs as plausible targets for the SufABCDS iron-sulfur cluster assembly chaperone complex. Note that fold changes are listed after gene IDs; n.c. denotes no change in expression.

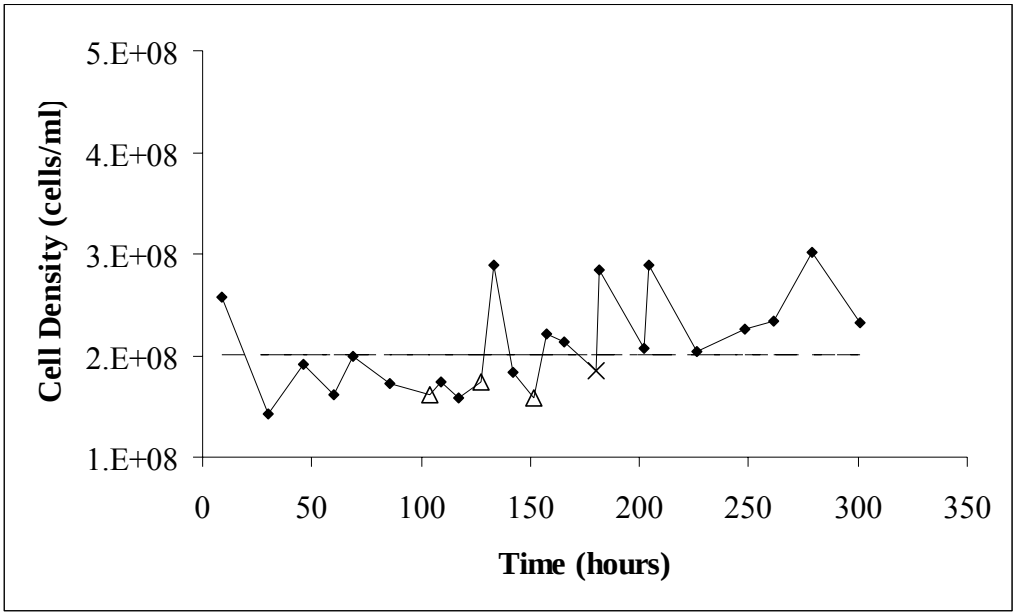
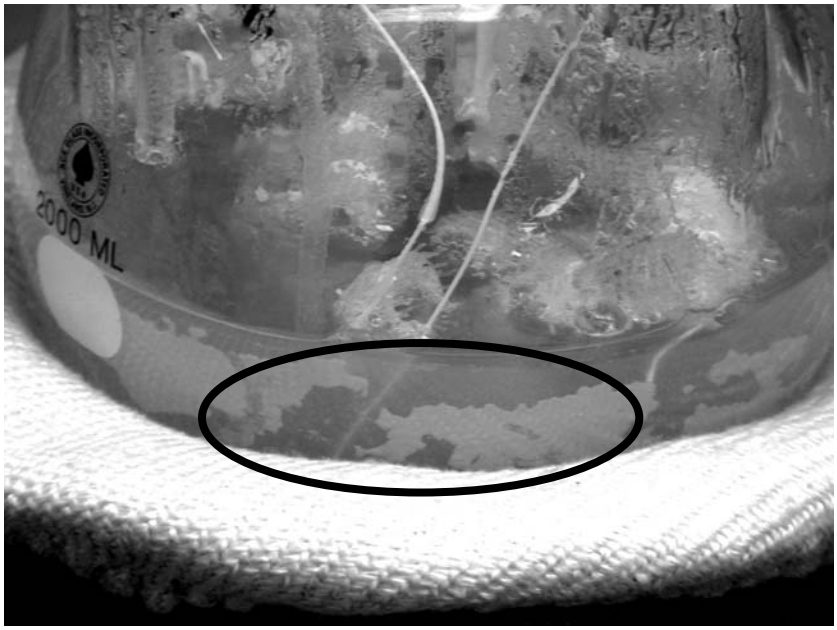
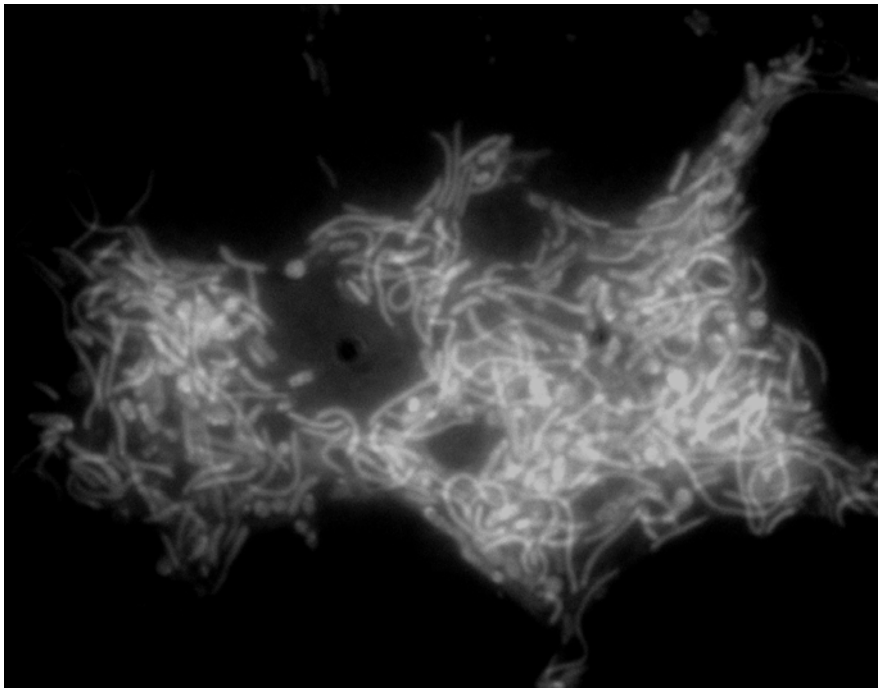

5.E+08
4.E+08
3.E+08
2.E+08
1.E+08
Cell Density (cells/ml)
0
50
100
150
200
250
300
350
Time (hours)

Figure 2.1

A



B



C

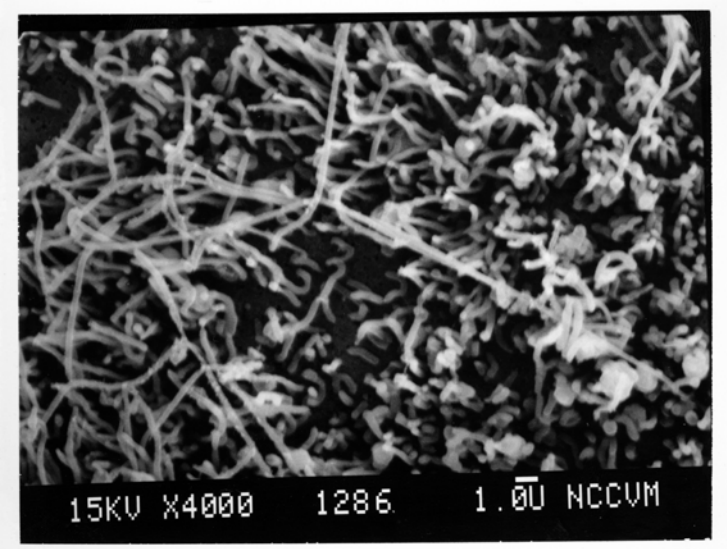


Figure 2.2

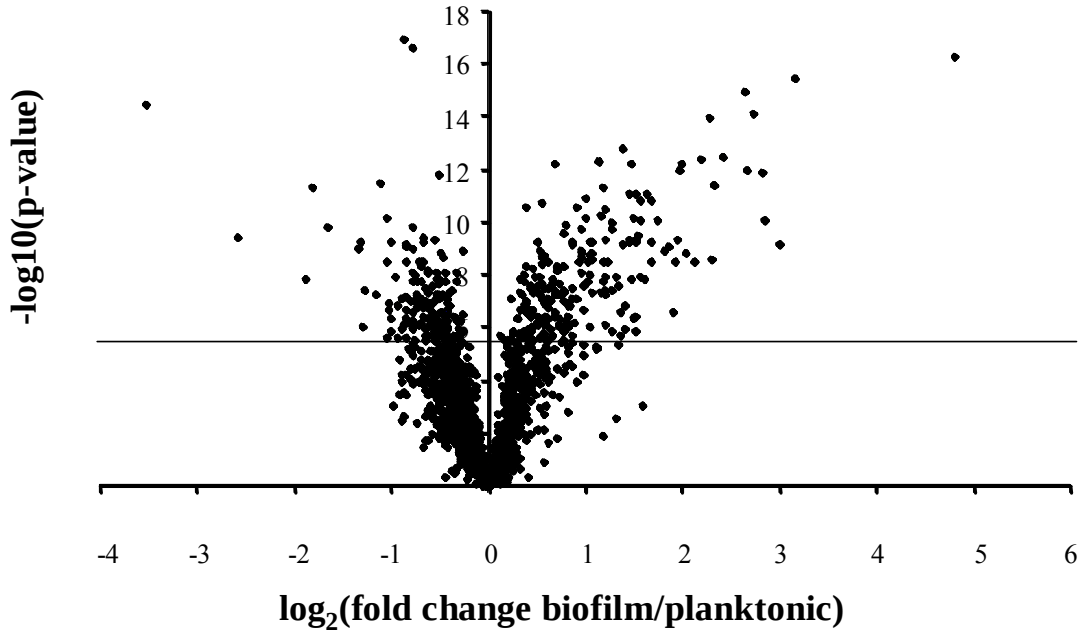
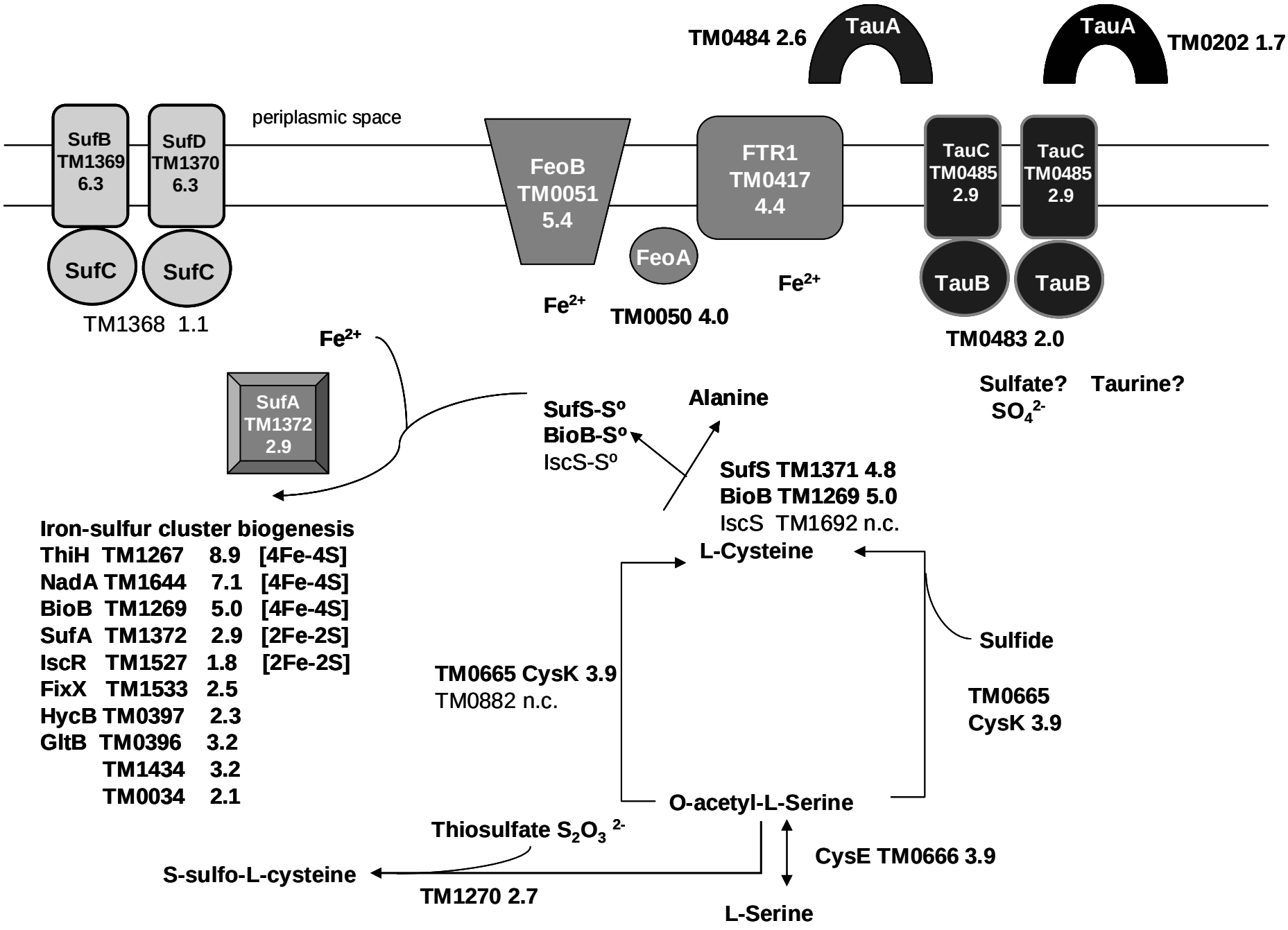

18
16
14
12
10
8
-log10(p-value)
-4
-3
-2
-1
0
1
2
3
4
5
6
log2(fold change biofilm/planktonic)

Figure 2.3


TM0484 2.6
TauA
TauA
TM0202 1.7
periplasmic space
SufB
TM1369
6.3
SufD
TM1370
6.3
SufC
SufC
TM1368 1.1
FeoB
TM0051
5.4
FeoA
TM0050 4.0
FTR1
TM0417
4.4
Fe2+
Fe2+
TauC
TM0485
2.9
TauC
TM0485
2.9
TauB
TauB
TM0483 2.0
Sulfate? Taurine?
SO4 2-
Fe2+
SufA
TM1372
2.9
SufS-So
BioB-So
IscS-So
Alanine
SufS TM1371 4.8
BioB TM1269 5.0
IscS TM1692 n.c.
Iron-sulfur cluster biogenesis
ThiH TM1267 8.9 [4Fe-4S]
NadA TM1644 7.1 [4Fe-4S]
BioB TM1269 5.0 [4Fe-4S]
SufA TM1372 2.9 [2Fe-2S]
IscR TM1527 1.8 [2Fe-2S]
FixX TM1533 2.5
HycB TM0397 2.3
GltB TM0396 3.2
TM1434 3.2
TM0034 2.1
L-Cysteine
Sulfide
TM0665 CysK 3.9
TM0882 n.c.
TM0665
CysK 3.9
O-acetyl-L-Serine
Thiosulfate S2O3 2-
CysE TM0666 3.9
S-sulfo-L-cysteine
TM1270 2.7
L-Serine

Chapter 3:

An expression-driven approach to the prediction of carbohydrate transport and utilization regulons in the hyperthermophilic bacterium *T. maritima*

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

Shannon B. Conners, Clemente I. Montero, Donald A. Comfort, Keith R. Shockley[2], Matthew R. Johnson, Swapnil R. Chhabra[1], and Robert M. Kelly

[2] Current Address: Biosystems Research Department
Sandia National Laboratories
Livermore, CA 94551

[3] Current Address: Jackson Lab
600 Main Street
Bar Harbor, ME 04609

Conners, S. B., C. I. Montero, D. A. Comfort, K. R. Shockley, M. R. Johnson, S. R. Chhabra, and R. M. Kelly. 2005. An expression-driven approach to the prediction of carbohydrate transport and utilization regulons in the hyperthermophilic bacterium *Thermotoga maritima*. J Bacteriol **187:**7267-82.

ABSTRACT

Comprehensive analysis of genome-wide expression patterns during growth of the hyperthermophilic bacterium *T. maritima maritima* on fourteen monosaccharide and polysaccharide substrates was undertaken with the goal of proposing carbohydrate specificities for transport systems and putative transcriptional regulators. Saccharide-induced regulons were predicted through complementary use of comparative genomics, mixed model analysis of genome-wide microarray expression data, and examination of upstream sequence patterns. Results indicate that *T. maritima* relies extensively on ABC transporters for carbohydrate uptake, many of which are likely controlled by local regulators responsive to either the transport substrate or a key metabolic degradation product. Roles in uptake of specific carbohydrates were suggested for members of the expanded Opp/Dpp family of ABC transporters. In this family, phylogenetic relationships among transport systems revealed patterns of possible duplication and divergence as a strategy for evolution of new uptake capabilities. The presence of GC-rich hairpin sequences between substrate-binding proteins and other components of Opp/Dpp family transporters offers a possible explanation for differential regulation of transporter subunit genes. Numerous improvements to *T. maritima* genome annotations were proposed, including the identification of ABC transport systems originally annotated as oligopeptide transporters as candidate transporters for rhamnose, xylose, β-xylan, and β- glucans, and identification of genes likely to encode proteins missing from current annotations of the pentose phosphate pathway. Beyond the information obtained for *T. maritima*, this work illustrates how expression-based strategies can be used for improving genome annotation in other microorganisms, especially those for which genetic systems are unavailable.

INTRODUCTION

T. maritima maritima, a hyperthermophilic anaerobe with an optimal growth temperature of 80°C, has been found in diverse high-temperature locations and is capable of using a wide variety of simple and complex carbohydrate substrates for growth. The complexity of its carbohydrate utilization strategies, revealed by genome sequencing (48) and through previous work (11, 12, 47, 51), is surprising, given the primitive features of this microorganism. Considerable genomic plasticity has been observed even within the *T. maritima* genus, with respect to gene content of carbohydrate active enzymes and transporter subunits, which may to some extent relate to lateral gene transfer events (48, 49). Despite the range of sugar-active enzymes found within *T. maritima* MSB8 genome (Table 2.S1), a PTS (phosphoenolpyruvate-dependent phosphotransferase system) similar to those used by other species for preferential uptake of selected sugars is apparently absent (48). No homologs of the PTS components EI and HPr (phosphocarrier proteins), nor the sugar-specific EII sugar transporter subunits have been identified in the *T. maritimales*. Homologs of PTS-associated transcriptional regulators are found in *T. maritima* MSB8 but have not been implicated in global transcriptional regulation of sugar uptake. While catabolite repression by glucose has been demonstrated for *T. maritima neapolitana* (78), a mechanism for the global regulation of sugar utilization remains to be identified within the *T. maritima* genus.

The importance of carbohydrates as carbon and energy sources for *T. maritima* is reflected by the disproportionate number of ABC (ATP-Binding Cassette) transporters which are found within *T. maritima* relative to its genome size (56). These ABC transporters can be classified into large families of sugar transporters and peptide (Opp, oligopeptide; Dpp, dipeptide) transporters, although it has been suggested that both types may participate in the

uptake of simple and complex sugars in *T. maritima* (11, 12, 28). Attempts to annotate the functional specificity of these transporters using computational tools have been largely unsuccessful (59) due to the phylogenetic distance between homologs in *T. maritima* and model bacteria. In fact, several sets of *T. maritima* "oligopeptide" transporters are more closely related to archaeal sugar transporters (15, 29) than characterized bacterial peptide transporters, and may have arrived in the *T. maritima* lineage through lateral gene transfer (48). Presumably, subsequent duplication and divergence events generated paralogous sets of transporter gene sub-families with different sugar-binding specificities. Determining the apparent specificities of each system, and associated transcriptional regulators or hydrolases, is a key step in testing this hypothesis. Most members of the LacI (Lactose Repressor) family of carbohydrate-responsive transcriptional regulators in *T. maritima* cannot be easily assigned into known functional classes using a subset of protein sites (44). Similarly, the specificities of the multiple *T. maritima* homologs of the XylR (Xylose Repressor) family regulators cannot be determined from sequence homology alone. The presence of these genes nearby sets of ABC transporters suggests that they may play a regulatory role in uptake and utilization of different carbohydrates. Genetic systems enabling knockouts or *in vivo* over-expression studies of genes are currently lacking for *T. maritima* as well as for the majority of sequenced bacterial genomes, which now number ~180 complete and >300 in progress (4). Clearly, alternative complementary methodologies are necessary for performing large-scale functional predictions for expanded protein families in organisms like *T. maritima* which lack genetic systems

Transcriptional analysis has proven to be a useful tool for the annotation of members of expanded gene families in a number of genomes. Such approaches have been instrumental in revealing biological pathways (41), and suggesting likely functions for individual genes,

operons, or multiple members of related families of glycoside hydrolases, transporters, and regulatory proteins (3, 5, 76). Previous studies in *T. maritima* to examine carbohydrate-related gene expression utilized Northern blots to examine transcription of selected hydrolases during growth on glucan- and mannan-based polysaccharides (12). Work with a full genome array comparing gene expression patterns of *T. maritima* on glucose, maltose, and lactose further underscored the relevance of this approach in the absence of a genetic system for this organism (51), and guided subsequent biochemical studies which suggested divergence of transporter substrates for two members of the maltose binding protein family (47). Efforts employing a targeted cDNA microarray demonstrated an expanded methodology for predicting carbohydrate-related gene expression in *T. maritima* (11). Here, a comprehensive analysis of genome-wide expression patterns during growth on fourteen monosaccharide and polysaccharide substrates (Table 3.1) was undertaken with the goal of suggesting sugar specificities for transport systems and putative regulators of unknown specificity found within the genome. Similar expression-based strategies could prove useful in improving genome annotation in other species of bacteria and archaea whose genomes have been sequenced but which also lack genetic systems.

MATERIALS AND METHODS

Growth of *T. maritima* and RNA Isolation. Cultures of *T. maritima* MSB8 were grown in artificial sea water using optical density measurements and epifluorescence microscopic cell density enumeration, as described previously (12). Growth substrates glucose, mannose, arabinose, rhamnose, ribose, xylose, β-xylan (birchwood), laminarin (*Laminaria digitata*), and starch (potato) were obtained from Sigma (St. Louis, MO. USA). Galactomannan (carob), glucomannan (konjac), and β-glucan (barley) were obtained from Megazyme (Wicklow,

Ireland), and pustulan (*U. papullosa*) was obtained from Calbiochem (San Diego, CA). These growth substrates were prepared as described previously (12). Substrate purities as provided by the manufacturers varied from 95 to 99%. All carbohydrate growth substrates were included in the medium at a final concentration of 0.25% (w/v). To ensure minimum carryover between substrates, cells were grown for at least 6 passes on each carbon source using a 0.5% (v/v) starting innoculum before obtaining the growth curves. Isolation of total RNA from *T. maritima* was performed on cells that were grown until early- to mid-exponential phase on the various growth substrates, using a protocol described previously (18).

Microarray protocols. A *T. maritima* cDNA microarray was constructed and utilized using methodologies discussed previously (11, 24). Hybridizations were carried out for 18 hours following modified TIGR protocols described elsewhere (11, 20, 21). Hybridized slides were scanned on a Perkin Elmer ExpressLite Scanner (Perkin Elmer) and quantitated using ScanArray 2.1 (Perkin Elmer).

Mixed model analyses of microarray data. Replication of treatments, arrays, dyes, and cDNA spots allowed the use of analysis of variance (26, 83) models for data analysis. A loop design was constructed (Figure 3.1) and reciprocal labeling utilized for all samples to estimate dye effects for each treatment. Scanarray spot intensities were imported into SAS (SAS Institute, Cary, NC) and flagged low intensity or low quality spots removed before further analysis. After local background subtraction and log transformation of spot intensities, a linear normalization ANOVA model (83) was used to estimate global variation in the form of fixed effects (dye (D), treatment (T)) random effects (array (A), spot A(S)) effects and random error using the model

$\log_2(y_{ijklmn}) = \mu + A_i + D_j + T_k + A_i(S_l) + \varepsilon_{ijklm}$. A gene-specific ANOVA model was used to partition the remaining variation into gene-specific effects using the model $r_{ijklmn} = \mu + A_i + D_j + T_k + A_i (S_l) + \varepsilon_{ijklm}$. Least-squares mean estimates of gene-specific treatment effects were examined using hierarchical clustering in JMP (SAS Institute) and histograms in Excel (Microsoft) to visualize expression patterns for specific contiguous genomic locations. A subset of samples included in this analysis represented biological repeats of conditions examined previously with an array including a targeted subset of *T. maritima* genes (11). The correlations between the two sets of least squares mean estimates of gene-specific treatment effects for genes in common between both arrays (n=262) were as follows: galactomannan, barley glucan, and glucose, $r \geq 0.78$, starch and mannose, $r \geq 0.62$. An examination of fold changes for genes most highly differentially expressed between selected pairs of treatments (e.g., barley and starch) revealed good agreement between gene lists, although the full genome array used here resulted in more conservative estimates of fold changes than the targeted array used previously (11). Unless otherwise noted, original gene annotations have been checked against the COG database at NCBI (74) and the Conserved Domain Database at NCBI (40). For information on the magnitude and statistical significance of fold changes for all of the genes included on the array, follow the microarray link at http://www.che.ncsu.edu/extremophiles/page5.html (data to be posted upon publication of the manuscript).

Prediction of transcriptional regulator binding sites and promoters. Consensus binding sequences for LacI and XylR family proteins were taken from the literature (32, 61). The web-based RSA Tools was used to extract the 300 bases upstream of every gene in the *T. maritima* genome, and the RSA Tools DNA Pattern search used to identify matches to regulator consensus

sequences with two or fewer mismatches (77). Pattern searches with more degenerate matches were identified using the program FuzzNuc from the EMBOSS software suite (60).

Construction of phylogenetic trees. Protein sequences were obtained from Genbank Batch Entrez and aligned with ClustalX (75). In an attempt to draw information from sequence homology between related *T. maritima* proteins, phylogenetic analysis was constructed separately for Opp/Dpp ABC transporter subunits (substrate-binding proteins, ATP-binding proteins, and permeases), other predicted sugar transporter subunits, LacI family regulators and XylR family regulators of *T. maritima* using MEGA2 (30). For the *T. maritima* Opp/Dpp family proteins, the topologies of individual phylogenetic trees for the substrate binding proteins, ATP-binding subunits, and permease proteins showed consistent relationships among operons for three methods (neighbor-joining, minimum evolution and maximum parsimony) (63, 65, 69). A consensus tree is shown in Figure 3.3.

RESULTS

Expression results for selected *T. maritima* ABC transporters. Figure 3.2 summarizes the genomic locations and microarray expression results for selected sets of genes in the *T. maritima* genome. These include ABC-type bacterial carbohydrate uptake transporters from the two main families, CUT1 and CUT2 (64), as well as members of the Opp/Dpp ABC transporter family. Expression data are also shown for associated hydrolases and putative transcriptional regulators. Based on these results, predictions of transporter specificities are shown in Table 3.2, along with a summary of specificities predicted by previous work.

CUT1 ABC transport systems. CUT1 transporters with substrate binding proteins related to maltose binding proteins include maltose transporter subunits (TM1836 and TM1839, TM1202-TM1204) recently shown to have different expression patterns and varied transport capabilities (47, 51). Expression patterns during growth in the presence of other carbohydrates which were examined did not reveal new information about the specificity of a third related set of transport proteins for which no substrate has yet been suggested (TM0418-TM0422) (Table 3.2).

The CUT1 permeases and substrate binding proteins encoded by TM0810-TM0813 are found with genes whose functions relate to breakdown of N-acetylglucosamine polysaccharides. However, growth of *T. maritima* in the presence of the β-1,4 N-acetylglucosamine polymer chitin was similar to control cultures, consistent with a lack of differential expression of this locus. Sequence similarity searches suggest that *T. maritima* lacks an identifiable chitinase, and might instead utilize chitin in the presence of neighboring species capable of chitin hydrolysis. Alternatively, transcription of these genes may be higher in the presence of N-acetylglucosamine or another N-acetylglucosamine-containing oligosaccharide found in the natural environment of *T. maritima*.

CUT2 transport systems. The two CUT2 transporters found in *T. maritima* are comprised of a substrate binding protein, a single permease subunit presumed to form a homodimer in the functional transporter and a fusion protein consisting of two nucleotide binding domains. Previously we observed the up-regulation of the LacI family gene TM0949 and the predicted ribokinase TM0960 during growth on xylose (11). Computational analysis of LacI regulators has determined that TM0949 is most similar to RbsR, a negative regulator of ribose uptake (17).

Here, several genes within the TM0949-TM0960 gene string were up-regulated during growth on xylose, ribose and arabinose (Figure 3.3A), including *rbs*ABCD homologs not examined previously by Chhabra et al. (11). From expression results alone, it is unclear whether this system can import multiple pentose sugars, or whether transcription of the genes is triggered by the interconversion of xylose or arabinose to ribose via the pentose phosphate pathway (Figure 3.3C). Two strong matches to a LacI family consensus binding site are arrayed consecutively upstream of the ribokinase TM0960 (Table 3.S2), and a predicted rho-independent terminator located downstream of TM0949 is the only identifiable terminator within the gene cluster (Figure 3.3A) (16). Similarly to observations of other transport systems of *T. maritima* discussed below, the putative binding protein of this transport set (TM0958) was more highly up-regulated than other transporter components.

Expression results shown here suggest several clarifications of *T. maritima* genome annotation and the *T. maritima* pentose phosphate pathway as predicted by sequence similarity in the KEGG database (25). The predicted KEGG pathway identifies an RpiB homolog responsible for the interconversion of ribulose-5-phosphate to ribose-5-phosphate as TM1080, which was detected at similar levels on all substrates (data not shown). However, expression results and sequence similarity suggest TM0951 as a possible a candidate for a second, inducible ribose-5-phosphate isomerase. Two nearby transketolase subunits (TM0953 and TM0954) previously annotated as frameshifts are detected at higher levels during growth on xylose, ribose, and arabinose, consistent with their proposed role in the *T. maritima* pentose phosphate pathway (Figure 3.3C). While TM0952 is annotated as a glycerol kinase, a second *T. maritima* glycerol kinase homolog (TM1430, GK2) shares greater sequence identity with *B. subtilis glpK* (66% id/479 aa vs. 45% id/487 aa) and co-localizes with other glycerol utilization genes. Both

TM0952 and TM1430 belong to the FGGY family of carbohydrate kinases, which also include xylulokinases, fucokinases, and gluconokinases (Pfam00370). TM0116, a predicted *T. maritima* xylulokinase, is found within a distant operon (see below) but was not observed to be differentially expressed on any sugar substrate examined here (data not shown). Given the lack of additional glycerol utilization genes nearby TM0952, a role for the encoded protein as an inducible xylulokinase should be considered.

Two hypothetical proteins of unknown function within the TM0949-TM0960 locus are also differentially expressed. The functions of these proteins remain unclear, yet their up-regulation during growth on multiple pentoses suggests a plausible role in pentose uptake or catabolism. TM0950, which is related to a hypothetical protein in *Lactobacillus johnsonii* (LJ1257), contains no known domains. However, LJ1257 is located in a gene cluster with similar composition to the *T. maritima* pentose-responsive locus, including a putative sugar isomerase (LJ1064), a LacI family regulator (LJ1265), N- and C- terminal transketolase subunits (LJ1266-1267), and an FGGY family sugar kinase.

Homologs to the RbsABC ABC transporter subunits bearing 40-52% identity to the *T. maritima* homologs are found together in the genome of the hyperthermophile *Thermoanaerobacter tengcongensis* (2) with a putative N-acetylglucosamine kinase (TTE0216) classified into COG2971 of the Clusters of Orthologous Groups of proteins database (74). A related *T. maritima* protein (TM1280) was expressed much more highly during growth of *T. maritima* on xylose than on any other sugar examined. The expression of TM1280 was >30 fold higher during growth on xylose than on the β(1,4)-linked xylose polymer xylan. The annotation for sugar specificity of this putative kinase is apparently drawn from a distantly related human

N-acetylglucosamine kinase (22), as the specificities of closely related microbial homologs have not yet been determined.

The second *T. maritima* CUT2 transport set is found with the XylR family regulator TM0110. Despite the lack of an *rbsD* cytoplasmic sugar binding homolog, genes homologous to *rbsABC* are all present. Transcripts of TM0110 were detected at higher levels during growth on xylose as compared to all other substrates tested here except laminarin, although other genes within this gene string were not significantly differentially expressed between xylose and any other sugar. Unlike the xylose catabolic genes of many model organisms, the characterized *T. maritima* xylose isomerase (TM1667) (1) is not found with the predicted xylulokinase (TM0116). This separation might reflect a broader physiological specificity of the TM1667 enzyme, which has also been used in the conversion of glucose to fructose (1), or may reflect differential regulation of the two activities in response to different xylose-containing substrates.

In addition to genes within the TM0949-TM0960 locus described above, other *T. maritima* genes also respond to the simple sugar L-arabinose, including a characterized L-arabinose isomerase (TM0276) (33), an α-L-arabinofuranosidase (TM0281), an uncharacterized conserved protein (TM0280), and a homolog to the protein araM from the *B. subtilis* arabinose utilization operon (66) (Figure 3.3B). Located upstream of these genes is a LacI family regulator, TM0275, which is most similar to AraR from a *Geobacillus stearothermophilus* arabinose cluster (Table 3.S2). Sugar ABC permease subunits TM0278 and TM0279 do not show strong differential regulation, and together with a frame-shifted substrate binding protein (TM0277) may suggest a nonfunctional transporter.

Sequence analysis of Opp/Dpp transporters subunits in *T. maritima*. Taken together, the well-documented ability of *T. maritima* to use complex carbohydrates and the lack of annotated polysaccharides transporters suggested novel oligosaccharide transporters yet to be identified in the *T. maritima* genome. The high degree of identity between the Dpp/Opp family cellobiose transporter of *P. furiosus* and a likely cellobiose transporter of *T. maritima* (29) has raised the possibility that additional related transporters of *T. maritima* might transport oligosaccharides. The phylogeny of Opp/Dpp transport subunits in the COG database (74) and BLAST homology searches (Table 3.S2) suggested three different lineage-specific gene expansions likely to have taken place after the divergence of *T. maritima* from the next closest sequenced organism. A consensus tree based on substrate-binding protein relationships with operon organizations superimposed is shown in Figure 3.4. Duplication or acquisition of fully intact Opp/Dpp ABC transport operons (one substrate-binding protein, two permeases, and two ATP-binding subunits) can be inferred, although three solitary substrate binding proteins are also apparent (Figure 3.4, Table 3.2). In two instances, these proteins display high levels of homology (>60% identity) to substrate-binding proteins of full transport systems, perhaps suggesting interaction with subunits of other transport systems. Duplication of Opp/Dpp substrate binding proteins in *T. maritima* might accomplish expansion of sugar binding capabilities for related substrates, as the peptide specificities of two Opp/Dpp family transporters of *Lactococcus lactis* IL1403 have been largely attributed to features of substrate binding proteins (14, 67).

Expression of Group 1 Opp/Dpp transporters is elevated during growth on β-linked gluco-oligosaccharide substrates. Three related substrate-binding proteins detected at higher levels during growth with β-linked sugars can be classified into group 1 of Opp/Dpp family

transport operons (Figure 3.4). These proteins share considerable similarity with a *P. furiosus* transporter implicated in the uptake of *β*-1,4 linked glucose oligomers including cellobiose, cellotriose, and laminaribose (29). Sequence similarity patterns suggest that this group likely arose from lateral gene transfer of one or two transport systems from archaea followed by duplication of the sugar binding protein and divergence of regulatory strategies and expression specificity (74). We have previously noted the location of a tightly conserved palindromic sequence motif similar to the LacI family consensus upstream of selected genes responsive to growth on carboxymethyl cellulose, barley glucan, glucomannan (11), and cellobiose (Montero and Kelly, unpublished observation). The putative regulator binding sites are situated between the -35 and -10 elements of consensus σ^A promoter binding sites (Table 3.S2). Here, a biological repetition of the barley and glucomannan growth experiments confirmed up-regulation of genes encoding cellobiose phosphorylase (TM1848) (58), two endoglucanases (TM1524, TM1525), the LacI family regulator TM1218, and the Opp/Dpp family ABC transporter subunits encoded by TM1219-TM1223 (36) (Figure 3.5A). In keeping with the designation of the related *P. furiosus* transporter, we suggest the designation CbtABCDF for the *T. maritima* transport set. Consistent with observations in *P. furiosus*, the substrate binding protein CbtA (TM1223) was more highly up-regulated in response to β-1,4 linked glucooligomers than the other transporter subunits (29). We suggest the designation CelR for the LacI family regulator TM1218. A search of the intervening sequence between TM1222 and TM1223 revealed a GC-rich inverted repeat with a spacing of two bases flanked by a σ^A-like promoter (Table 3.S2). Subsequent searches of other Opp/Dpp family transporter strings revealed 5 additional cases of GC-rich inverted repeats located between coding sequences of Opp/Dpp family binding proteins and other transporter subunits, with spacing between the inverted repeats varying from n=2 to n=5 (Table 3.S2, Figure

3.4). Transcript levels detected from substrate-binding proteins responded more strongly during growth on the predicted transporter substrate than did other transporter components, raising the intriguing possibility that these inverted repeats might play a role in modulating transcriptional levels of transporter components.

A second gene string, separated from TM1218-TM1223 by 132 bases, encodes a XylR family regulatory protein (TM1224), a putative glycosylase (TM1225), a second CbtA homolog (TM1226, 60% identity with TM1223), and ManB (TM1227), a characterized β-mannanase (55). TM1224-TM1227 are up-regulated on mannose-containing carbon sources (Figure 3.5A), consistent with the ability of the carbohydrate-binding domain of TM1227 to accommodate mannooligosaccharides, galactomannan and glucomannan degradation products (8). The specificity of the up-regulation of TM1224 (here designated ManR) and TM1226 (here designated MbtA) on mannose, glucomannan and galactomannan is especially striking. It appears likely that TM1226 might interact with the ATP-binding and permease subunits of the cellobiose transporter. In agreement with their close phylogenetic grouping, TM1223 and TM1226 are reciprocal best BLAST homologs; this suggests possible past duplication and specialization of the binding protein for the cellobiose transporter to accommodate mannan oligosaccharides. It also appears likely that expression of TM1226 could be under the transcriptional control of TM1224. Other candidates for regulation by TM1224 include genes previously observed to be up-regulated on glucomannan and galactomannan within the TM1745-TM1752 gene string and the β-mannosidase TM1624 (11) (Figure 3.5B). Although the OppA-family binding protein TM1746 is more highly up-regulated during growth on xylose, other components of the transporter (TM1747-TM1750), two endoglucanases (Cel5A TM1751, Cel5B

TM1752), and a β-mannosidase (TM1624) are highest during growth on mannans. We propose to designate the transporter components TM1746-TM1750 as MtpABCDF (Table 3.2).

The third substrate binding protein of group 1 (TM0031) is located within a gene string encoding the laminaribiase BglB/Cel3 (TM0025) (89) and laminarinase TM0024 (7, 88), as well as components of an ABC transport complex homologous to CbtABCDF not examined by Chhabra et al. (11). A XylR family regulator (TM0032) is located upstream of the ABC transporter components. Higher transcript levels for components of this transporter during growth on the β-1,3 linked glucose polymer laminarin, the mixed β-1,3-β-1,4-linkage glucose polysaccharide barley, and the β-1,6 linked glucose polymer pustulan may suggest a general role in the uptake of β-linked sugars (Figure 3.5C). We suggest the designations BgtpABCDF and BglcR for the transporter and regulator, respectively (Table 3.2). Similarly to CbtA of the cellobiose transporter, TM0031 (BgtpA) was detected at higher levels in the presence of β-glucans than other transporter components, and a GC-rich inverted repeat was found in the intervening sequence between BgtpA and TM0030 (BgtpB) (Table 3.S2). A proposed pathway for the uptake and utilization of β-glucan and β-mannan oligosaccharides is shown in Figure 3.5D.

Higher transcript levels of Group 2 and 3 Opp/Dpp transporters during growth on xylose and xylose-containing oligosaccharides. Components of two distinct Opp/Dpp family transporters were detected at higher levels in the presence of the simple sugar xylose and the polysaccharide xylan (Table 3.1). The two sets of transport proteins are located nearby one another, separated by a set of genes predicted to encode enzymes for the catabolism of uronic acids. We have previously noted the similarities in functional composition of this gene cluster

(11) to the xylan utilization cluster of *Geobacillus stearothermophilus* T-6 (68). Both sets of *T. maritima* transporters are divergently transcribed from family 10 xylanases (*xylA/xyl10A*, TM0061; *xylB/xyl10B*, TM0070) (Figure 3.6), both reported previously to be active on xylan polysaccharides (13, 79, 87). The similarities in expression profiles and gene content of the two gene sets do not appear to be the result of a recent duplication, as reflected in the consensus phylogenetic tree of Opp/Dpp family transport components (Figure 3.4). Comparison with sequences from other sequenced organisms reveals that the TM0071-TM0075 gene set clusters with two other *T. maritima* ABC transporter sets in a grouping which apparently arose from a lateral gene transfer event with archaea (74), likely followed by duplication and divergence within an ancestral lineage (Group 2, Figure 3.4). In contrast, TM0056-TM0060 cluster within a group of bacterial transporter proteins (74). The genomic arrangement of the two xylose and xylan-responsive transporter gene sets also differs (Figure 3.6).

We propose to designate TM0071-TM0075 as XtpABCDF in keeping with the names assigned to orthologous proteins found in an unpublished cluster of xylan utilization genes from *T. neapolitana* (GI:23270356). Transcript levels of TM0071 were slightly higher on xylose than xylan, while other members of the transport operon showed varying degrees of preference for xylose over xylan (Figure 3.6B). In contrast, the substrate binding protein TM0056 was detected at much higher levels during growth on xylan than xylose (Figure 3.6A). Transcripts from the remaining transporter subunits (TM0057-TM0060) and the α-glucuronidase AguA (TM0055) (62) were detected at higher levels during growth on xylose and xylan when compared to non-xylose sugars. We propose to designate the transport proteins encoded by TM0056-TM0060 as XtpGHJLM (Table 3.2). The DppA substrate binding protein encoded by TM0309 (proposed designation XtpN) is closely related to TM0056 (Table 3.S2) and is found nearby a predicted α-

xylosidase (TM0308). The slight up-regulation of both TM0309 (proposed designation XtpN) and TM0310 on xylan may suggest a role for these proteins (and possibly TM0308) in uptake and hydrolysis of an undetermined xylose-containing polysaccharide (Figure 3.2).

The variation in expression patterns for the xylanase-associated ABC transporters may relate to differences in the carbohydrate binding specificity for *T. maritima* Xyl10A and Xyl10B. XylA contains four carbohydrate binding domains absent in XylB: the A1 and A2 domains of XylA have been shown to bind xylan while the C1 and C2 domains bind cellulose and a number of other monosaccharides and polysaccharides (7). The hydrolase content of the two gene strings also differs, suggesting likely specialization of the transporters for differently-substituted xylan degradation products. A β-xylosidase (86) and acetyl xylan esterase co-localize with TM0071-TM0075, and an α-glucuronidase co-localizes with TM0056-TM0060. While no regulatory proteins are located within either xylanase-transporter gene string, similar inverted repeat sequences found upstream of Xyl10A (TM0061), Xyl10B (TM0070), and a putative α-xylosidase of glycosyl hydrolase family 31 (TM0308) share similarity with the consensus for a XylR family regulator (Table 3.S2). The XylR family regulator (TM0110) is expressed more highly on xylose and laminarin than any other substrate. The observation of similar expression profiles on xylose and laminarin may relate to the co-occurrence of carbohydrate binding domains for binding xylan and mixed linkage glucan carbohydrates by distinct domains in Xyl10A (8), or reflect sequence similarity between the XylR family regulators TM0110 and TM0032 (BglcR).

Higher transcript levels of a Group 3 Opp/Dpp transporter during growth on the simple sugar rhamnose. Growth of *T. maritima* on L-rhamnose (a methyl pentose also known as

deoxy-L-mannose) had not been previously demonstrated. Within group 3 of the Opp/Dpp family transporters of *T. maritima* is a set of ABC transporter components which co-localize with predicted rhamnose catabolic genes (31, 48) (Figure 3.7A). Here, the majority of genes which showed higher transcript levels during growth on rhamnose are found in this locus (Figure 3.7A). Transcripts of nearly all genes encoding subunits of the transporter (TM1063-TM1067) were observed at higher levels during growth on rhamnose as compared to all other sugars examined here. We suggest the designation RtpABCDF for these transport components (Table 3.2). Similarly to related transport systems, a GC-rich inverted repeat was found in the intervening sequence between RtpA (TM1067) and RtpB (TM1066) (Table 3.S2). The presence of an α-glucuronidase (TM1068, Agu4C) and β-glucuronidase (TM1062) within this locus suggest that the ABC transporter encoded by TM1063-TM1067 might also be involved in the uptake of rhamnose-containing disaccharides or oligosaccharides which include glucuronic acid residues. A second candidate rhamnose transporter is encoded by TM1060, which shares sequence similarity with major facilitator superfamily sugar-proton symporters. While the likely L-rhamnulose aldolase RhaD (TM1072) (53) and predicted rhamnulokinase RhaB (TM1073) are homologous to *E. coli* K12 rhamnose catabolic genes, an RhaA rhamnose isomerase homolog is missing. A likely substitute is TM1071, annotated as a putative sugar isomerase, which is homologous to rhamnose isomerase RhaI of *Rhizobium leguminosarum* bv. trifolii (52) and *Bacteriodes thetaiomicron* VPI-5482 (85). Several hypothetical proteins within the rhamnose locus present interesting targets for further work (Table 3.S2).

Based on this analysis, a predicted pathway for rhamnose utilization in *T. maritima* is shown in Figure 3.7B. Expression data from this locus also suggest a potential mechanism for transcriptional regulation. A DeoR/GlpR family transcriptional regulator (TM1069, COG1349)

found within the rhamnose transport and catabolism cluster shares sequence identity with proteins found within rhamnose catabolic clusters of *Bacillus halodurans* (72) and *Oceanobacillus iheyensis* (73). Therefore, we propose to designate TM1069 as RhaR.

Opp/Dpp transporters of unknown specificity. Expression data and genomic neighborhood analysis did not reveal specific substrate preferences for several sets of Opp/Dpp family transporter components (Table 3.2). Further work will be necessary to clarify whether these proteins are involved in uptake of untested sugars or alternative substrates transported by other members of the Opp/Dpp transporter family, such as metal ions (84) or peptides (19, 57).

DISCUSSION

The combination of microarray data with gene neighborhood and sequence analysis represents a powerful high-throughput approach for examining gene regulation and predicting functional roles of genes for microorganisms which lack genetic systems. Here, the genomic contexts and transcriptional responses of *T. maritima* genes to 14 monosaccharide and polysaccharide substrates were examined to improve upon previous annotations of ABC transporter proteins (48, 59). However, similar approaches could be used to perform substrate-by-substrate analysis of transcriptional responses to additional carbohydrates, peptides, metals, antibiotics, or other elements in microbial species for which microarrays are available. As demonstrated here, a loop experimental design allows efficient collection of large microarray datasets. Mixed model analysis of these datasets then enables comparisons between transcript levels for all pairs of substrates, rather than limiting comparisons to an arbitrary reference

condition. In this case, this approach allowed greater flexibility in comparing responses to both changes in carbohydrate composition and branch type.

Transcriptional data presented here support the hypothesis that many members of the Opp/Dpp ABC transporter family of *T. maritima* are involved in carbohydrate transport, and explain the observation that glycoside hydrolases often co-localize with these genes. Given the differential regulation of related Opp/Dpp transport systems in response to carbohydrates, this strategy has likely allowed the acquisition of new uptake capabilities, perhaps assisting in the adaptation of *T. maritima* species to specific environments. Transcriptional information was especially helpful in suggesting candidate substrates for several Opp/Dpp gene sets resulting from apparent lateral gene transfer followed by duplication and divergence (Figure 3.4). In two cases, the next closest related transporter gene sets are found in archaea. In total, carbohydrate specificities were proposed for six full or partial operons of Opp/Dpp transporter subunits, and expression results were confirmed for two operons previously examined (11). The results obtained will assist in streamlining biochemical characterizations of substrate binding protein specificities for *T. maritima* in progress in our laboratory and others. Although *T. maritima* does not grow on peptides as a sole carbon source, it is still unclear whether any of its Opp/Dpp transport systems are involved in peptide import. However, transcripts of an Opp/Dpp family transporter operon (TM0500-TM0503) which lacks a substrate binding protein are detected at higher levels in high density co-cultures of *T. maritima* and *Methanococcus jannaschii* (24). This transporter may be involved in export of a small peptide (TM0504) located downstream of the transporter which has been implicated in quorum sensing and biofilm formation (24).

Differential expression information for predicted carbohydrate-responsive transcriptional regulators of *T. maritima* has now assisted in the prediction of putative functions for previously

unannotated members of the LacI (3 proteins), XylR (3 proteins), and GlpR/DeoR (one protein) families. These include candidates for the control of the uptake of β-glucans, (TM1218, TM0032) β-mannans (TM1224), xylose/xylan (TM0110, TM0949), arabinose (TM0275) and rhamnose (TM1069). Sequences resembling binding sites can be detected upstream of selected carbohydrate-responsive genes (Table 3.S2), further supporting the hypothesis that some or all of these proteins are involved in the regulation of carbohydrate import. The specificities of most of these regulators would have been impossible to determine from sequence analysis alone, but transcriptional data now offer insights into plausible substrates for further characterization efforts.

For several Opp/Dpp transporter operons, substrate binding proteins showed greater transcriptional responses to changes in carbon source than did other transporter subunits. If transcript levels correlate well with protein levels, this might indicate that transporter subunits are present in the absence of substrate. Increased transcription of substrate binding proteins could allow maximal capture of available carbohydrates to be transported by existing permease and ATPase subunits. A partial explanation for the differential regulation of Opp/Dpp substrate binding proteins relative to other transporter subunits is suggested by the presence of GC-rich hairpin structures in the intervening sequence between the subunits. The possibility that these hairpin sequences act as partial transcriptional terminators should be explored further.

In contrast to the Opp/Dpp family transporters, most members of known carbohydrate transporter families were not differentially expressed here. A notable exception was the CUT2 transporter which showed transcriptional responses to ribose, arabinose, and xylose. It is possible that other predicted sugar transporters respond to substrates not tested here. For example, maltose and lactose were not examined but CUT1 transporters of *T. maritima* do respond to the

presence of these sugars (51) (Table 3.2). The possibility remains that some *T. maritima* sugar transport operons are transcribed constitutively, perhaps independent of the control of local transcriptional regulators.

The lack of PTS system components in the *T. maritima* genome argues against mechanisms of global catabolite repression identical to those operating in gram-negative and gram-positive model organisms. Future work using microarrays to examine data from growth experiments with combinations of substrates will be needed to explore alternative mechanisms of preferred substrate utilization *T. maritima*. Although an alternative mechanism can not be ruled out, the proximity of regulators and differentially expressed genes involved in sugar utilization provides evidence that local transcriptional regulators play important roles in regulating uptake of individual sugars through inducible ABC transport systems. Inducible and independent transcriptional control of transport systems of varied specificities may assist *T. maritima* in discriminating between and responding to complex polysaccharides found in its natural environment.

ACKNOWLEDGMENTS

This work was supported in part by grants from the National Science Foundation (Grant No. BES-0317886) and the Department of Energy, Energy Biosciences Program (Grant No. DE-FG02-96ER20219).

CHAPTER 3 REFERENCES

1. **Bandlish, R. K., J. Michael Hess, K. L. Epting, C. Vieille, and R. M. Kelly.** 2002. Glucose-to-fructose conversion at high temperatures with xylose (glucose) isomerases from *Streptomyces murinus* and two hyperthermophilic *T. maritima* species. Biotechnol Bioeng **80:**185-94.
2. **Bao, Q., Y. Tian, W. Li, Z. Xu, Z. Xuan, S. Hu, W. Dong, J. Yang, Y. Chen, Y. Xue, Y. Xu, X. Lai, L. Huang, X. Dong, Y. Ma, L. Ling, H. Tan, R. Chen, J. Wang, J. Yu, and H. Yang.** 2002. A complete sequence of the *T. tengcongensis* genome. Genome Res **12:**689-700.
3. **Barrangou, R., E. Altermann, R. Hutkins, R. Cano, and T. R. Klaenhammer.** 2003. Functional and comparative genomic analyses of an operon involved in fructooligosaccharide utilization by *Lactobacillus acidophilus*. Proc Natl Acad Sci U S A **100:**8957-62.
4. **Bernal, A., U. Ear, and N. Kyrpides.** 2001. Genomes OnLine Database (GOLD): a monitor of genome projects world-wide. Nucleic Acids Res **29:**126-7.
5. **Bertram, R., M. Schlicht, K. Mahr, H. Nothaft, M. H. Saier, Jr., and F. Titgemeyer.** 2004. In silico and transcriptional analysis of carbohydrate uptake systems of *Streptomyces coelicolor* A3(2). J Bacteriol **186:**1362-73.
6. **Bibel, M., C. Brettl, U. Gosslar, G. Kriegshauser, and W. Liebl.** 1998. Isolation and analysis of genes for amylolytic enzymes of the hyperthermophilic bacterium *T. maritima maritima*. FEMS Microbiol Lett **158:**9-15.
7. **Boraston, A. B., A. L. Creagh, M. M. Alam, J. M. Kormos, P. Tomme, C. A. Haynes, R. A. Warren, and D. G. Kilburn.** 2001. Binding specificity and thermodynamics of a

family 9 carbohydrate-binding module from *T. maritima maritima* xylanase 10A. Biochemistry **40:**6240-7.

8. **Boraston, A. B., T. J. Revett, C. M. Boraston, D. Nurizzo, and G. J. Davies.** 2003. Structural and thermodynamic dissection of specific mannan recognition by a carbohydrate binding module, TmCBM27. Structure (Camb) **11:**665-75.

9. **Bronnenmeier, K., A. Kern, W. Liebl, and W. L. Staudenbauer.** 1995. Purification of *T. maritima maritima* enzymes for the degradation of cellulosic materials. Appl Environ Microbiol **61:**1399-407.

10. **Chhabra, S. R., and R. M. Kelly.** 2002. Biochemical characterization of *T. maritima maritima* endoglucanase Cel74 with and without a carbohydrate binding module (CBM). FEBS Lett **531:**375-80.

11. **Chhabra, S. R., K. R. Shockley, S. B. Conners, K. L. Scott, R. D. Wolfinger, and R. M. Kelly.** 2003. Carbohydrate-induced differential gene expression patterns in the hyperthermophilic bacterium *T. maritima maritima*. J Biol Chem **278:**7540-52.

12. **Chhabra, S. R., K. R. Shockley, D. E. Ward, and R. M. Kelly.** 2002. Regulation of endo-acting glycosyl hydrolases in the hyperthermophilic bacterium *T. maritima maritima* grown on glucan- and mannan-based polysaccharides. Appl Environ Microbiol **68:**545-54.

13. **Dakhova, O. N., N. E. Kurepina, V. V. Zverlov, V. A. Svetlichnyi, and G. A. Velikodvorskaya.** 1993. Cloning and expression in *Escherichia coli* of *T. maritima neapolitana* genes coding for enzymes of carbohydrate substrate degradation. Biochem Biophys Res Commun **194:**1359-64.

14. **Doeven, M. K., R. Abele, R. Tampe, and B. Poolman.** 2004. The binding specificity of OppA determines the selectivity of the oligopeptide ATP-binding cassette transporter. J Biol Chem **279:**32301-7.

15. **Elferink, M. G., S. V. Albers, W. N. Konings, and A. J. Driessen.** 2001. Sugar transport in *Sulfolobus solfataricus* is mediated by two families of binding protein-dependent ABC transporters. Mol Microbiol **39:**1494-503.

16. **Ermolaeva, M. D., H. G. Khalak, O. White, H. O. Smith, and S. L. Salzberg.** 2000. Prediction of transcription terminators in bacterial genomes. J Mol Biol **301:**27-33.

17. **Fukami-Kobayashi, K., Y. Tateno, and K. Nishikawa.** 2003. Parallel evolution of ligand specificity between LacI/GalR family repressors and periplasmic sugar-binding proteins. Mol Biol Evol **20:**267-77.

18. **Gao, J., M. W. Bauer, K. R. Shockley, M. A. Pysz, and R. M. Kelly.** 2003. Growth of hyperthermophilic archaeon *Pyrococcus furiosus* on chitin involves two family 18 chitinases. Appl Environ Microbiol **69:**3119-28.

19. **Garault, P., D. Le Bars, C. Besset, and V. Monnet.** 2002. Three oligopeptide-binding proteins are involved in the oligopeptide transport of *Streptococcus thermophilus*. J Biol Chem **277:**32-9.

20. **Hasseman, J.** 2001. TIGR microarray protocols. http://www.tigr.org/tdb/microarray/protocolsTIGR.shtml

21. **Hegde, P., R. Qi, K. Abernathy, C. Gay, S. Dharap, R. Gaspard, J. E. Hughes, E. Snesrud, N. Lee, and J. Quackenbush.** 2000. A concise guide to cDNA microarray analysis. Biotechniques **29:**548-50, 552-4, 556 passim.

22. **Hinderlich, S., M. Berger, M. Schwarzkopf, K. Effertz, and W. Reutter.** 2000. Molecular cloning and characterization of murine and human N-acetylglucosamine kinase. Eur J Biochem **267:**3301-8.

23. **Ihsanawati, T. Kumasaka, T. Kaneko, C. Morokuma, S. Nakamura, and N. Tanaka.** 2003. Crystallization and preliminary X-ray studies of xylanase 10B from *T. maritima maritima*. Acta Crystallogr D Biol Crystallogr **59:**1659-61.

24. **Johnson, M. R., C. I. Montero, S. B. Conners, K. R. Shockley, S. L. Bridger, and R. M. Kelly.** 2005. Population density-dependent regulation of exopolysaccharide formation in the hyperthermophilic bacterium *T. maritima maritima*. Mol Microbiol **55:**664-74.

25. **Kanehisa, M., S. Goto, S. Kawashima, Y. Okuno, and M. Hattori.** 2004. The KEGG resource for deciphering the genome. Nucleic Acids Res **32 Database issue:**D277-80.

26. **Kerr, M. K., M. Martin, and G. A. Churchill.** 2000. Analysis of variance for gene expression microarray data. J Comput Biol **7:**819-37.

27. **Kluskens, L. D., G. J. van Alebeek, A. G. Voragen, W. M. de Vos, and J. van der Oost.** 2003. Molecular and biochemical characterization of the thermoactive family 1 pectate lyase from the hyperthermophilic bacterium *T. maritima maritima*. Biochem J **370:**651-9.

28. **Koning, S. M., S. V. Albers, W. N. Konings, and A. J. Driessen.** 2002. Sugar transport in (hyper)thermophilic archaea. Res Microbiol **153:**61-7.

29. **Koning, S. M., M. G. Elferink, W. N. Konings, and A. J. Driessen.** 2001. Cellobiose uptake in the hyperthermophilic archaeon *Pyrococcus furiosus* is mediated by an inducible, high-affinity ABC transporter. J Bacteriol **183:**4979-84.

30. **Kumar, S., K. Tamura, I. B. Jakobsen, and M. Nei.** 2001. MEGA2: molecular evolutionary genetics analysis software. Bioinformatics **17:**1244-5.

31. **Kyrpides, N. C., C. A. Ouzounis, I. Iliopoulos, V. Vonstein, and R. Overbeek.** 2000. Analysis of the *T. maritima maritima* genome combining a variety of sequence similarity and genome context tools. Nucleic Acids Res **28:**4573-6.

32. **Laikova, O. N., A. A. Mironov, and M. S. Gelfand.** 2001. Computational analysis of the transcriptional regulation of pentose utilization systems in the gamma subdivision of Proteobacteria. FEMS Microbiol Lett **205:**315-22.

33. **Lee, D. W., H. J. Jang, E. A. Choe, B. C. Kim, S. J. Lee, S. B. Kim, Y. H. Hong, and Y. R. Pyun.** 2004. Characterization of a thermostable L-arabinose (D-galactose) isomerase from the hyperthermophilic eubacterium *T. maritima maritima*. Appl Environ Microbiol **70:**1397-404.

34. **Lee, M. H., Y. W. Kim, T. J. Kim, C. S. Park, J. W. Kim, T. W. Moon, and K. H. Park.** 2002. A novel amylolytic enzyme from *T. maritima maritima*, resembling cyclodextrinase and alpha-glucosidase, that liberates glucose from the reducing end of the substrates. Biochem Biophys Res Commun **295:**818-25.

35. **Liebl, W., R. Feil, J. Gabelsberger, J. Kellermann, and K. H. Schleifer.** 1992. Purification and characterization of a novel thermostable 4-alpha-glucanotransferase of *T. maritima maritima* cloned in *Escherichia coli*. Eur J Biochem **207:**81-8.

36. **Liebl, W., P. Ruile, K. Bronnenmeier, K. Riedel, F. Lottspeich, and I. Greif.** 1996. Analysis of a *T. maritima maritima* DNA fragment encoding two similar thermostable cellulases, CelA and CelB, and characterization of the recombinant enzymes. Microbiology **142 (Pt 9):**2533-42.

37. **Liebl, W., I. Stemplinger, and P. Ruile.** 1997. Properties and gene structure of the *T. maritima maritima* alpha-amylase AmyA, a putative lipoprotein of a hyperthermophilic bacterium. J Bacteriol **179:**941-8.

38. **Liebl, W., B. Wagner, and J. Schellhase.** 1998. Properties of an alpha-galactosidase, and structure of its gene *galA*, within an alpha-and beta-galactoside utilization gene cluster of the hyperthermophilic bacterium *T. maritima maritima*. Syst Appl Microbiol **21:**1-11.

39. **Lim, W. J., S. R. Park, C. L. An, J. Y. Lee, S. Y. Hong, E. C. Shin, E. J. Kim, J. O. Kim, H. Kim, and H. D. Yun.** 2003. Cloning and characterization of a thermostable intracellular alpha-amylase gene from the hyperthermophilic bacterium *T. maritima maritima* MSB8. Res Microbiol **154:**681-7.

40. **Marchler-Bauer, A., J. B. Anderson, C. DeWeese-Scott, N. D. Fedorova, L. Y. Geer, S. He, D. I. Hurwitz, J. D. Jackson, A. R. Jacobs, C. J. Lanczycki, C. A. Liebert, C. Liu, T. Madej, G. H. Marchler, R. Mazumder, A. N. Nikolskaya, A. R. Panchenko, B. S. Rao, B. A. Shoemaker, V. Simonyan, J. S. Song, P. A. Thiessen, S. Vasudevan, Y. Wang, R. A. Yamashita, J. J. Yin, and S. H. Bryant.** 2003. CDD: a curated Entrez database of conserved domain alignments. Nucleic Acids Res **31:**383-7.

41. **Meibom, K. L., X. B. Li, A. T. Nielsen, C. Y. Wu, S. Roseman, and G. K. Schoolnik.** 2004. The *Vibrio cholerae* chitin utilization program. Proc Natl Acad Sci U S A **101:**2524-9.

42. **Meissner, H., and W. Liebl.** 1998. *T. maritima maritima* maltosyltransferase, a novel type of maltodextrin glycosyltransferase acting on starch and malto-oligosaccharides. Eur J Biochem **258:**1050-8.

43. **Meissner, K., D. Wassenberg, and W. Liebl.** 2000. The thermostabilizing domain of the modular xylanase XynA of *T. maritima maritima* represents a novel type of binding domain with affinity for soluble xylan and mixed-linkage beta-1,3/beta-1, 4-glucan. Mol Microbiol **36:**898-912.

44. **Mirny, L. A., and M. S. Gelfand.** 2002. Using orthologous and paralogous proteins to identify specificity-determining residues in bacterial transcription factors. J Mol Biol **321:**7-20.

45. **Moore, J. B., P. Markiewicz, and J. H. Miller.** 1994. Identification and sequencing of the *T. maritima maritima* lacZ gene, part of a divergently transcribed operon. Gene **147:**101-6.

46. **Nakajima, M., H. Imamura, H. Shoun, and T. Wakagi.** 2003. Unique metal dependency of cytosolic alpha-mannosidase from *T. maritima maritima*, a hyperthermophilic bacterium. Arch Biochem Biophys **415:**87-93.

47. **Nanavati, D. M., T. N. Nguyen, and K. M. Noll.** 2005. Substrate specificities and expression patterns reflect the evolutionary divergence of maltose ABC transporters in *T. maritima maritima*. J Bacteriol **187:**2002-9.

48. **Nelson, K. E., R. A. Clayton, S. R. Gill, M. L. Gwinn, R. J. Dodson, D. H. Haft, E. K. Hickey, J. D. Peterson, W. C. Nelson, K. A. Ketchum, L. McDonald, T. R. Utterback, J. A. Malek, K. D. Linher, M. M. Garrett, A. M. Stewart, M. D. Cotton, M. S. Pratt, C. A. Phillips, D. Richardson, J. Heidelberg, G. G. Sutton, R. D. Fleischmann, J. A. Eisen, C. M. Fraser, and et al.** 1999. Evidence for lateral gene transfer between Archaea and bacteria from genome sequence of *T. maritima maritima*. Nature **399:**323-9.

49. **Nesbo, C. L., K. E. Nelson, and W. F. Doolittle.** 2002. Suppressive subtractive hybridization detects extensive genomic diversity in *T. maritima maritima*. J Bacteriol **184:**4475-88.

50. **Nguyen, T. N., K. M. Borges, A. H. Romano, and K. M. Noll.** 2001. Differential gene expression in *T. maritima neapolitana* in response to growth substrate. FEMS Microbiol Lett **195:**79-83.

51. **Nguyen, T. N., A. D. Ejaz, M. A. Brancieri, A. M. Mikula, K. E. Nelson, S. R. Gill, and K. M. Noll.** 2004. Whole-genome expression profiling of *T. maritima maritima* in response to growth on sugars in a chemostat. J Bacteriol **186:**4824-8.

52. **Oresnik, I. J., L. A. Pacarynuk, S. A. P. O'Brien, C. K. Yost, and M. F. Hynes.** 1998. Plasmid-Encoded Catabolic Genes in *Rhizobium leguminosarum* bv. trifolii: Evidence for a plant-inducible rhamnose locus involved in competition for nodulation. Mol Plant Microbe Interact **11:**1175-1185.

53. **Osipiuk, J., M. E. Cuff, O. Korolev, T. Skarina, A. Savchenko, A. Edwards, and A. Joachimiak.** Crystal Structure Of Sugar-Phosphate Aldolase From *T. maritima maritima*. Unpublished (GI:37927646).

54. **Parisot, J., A. Ghochikyan, V. Langlois, V. Sakanyan, and C. Rabiller.** 2002. Exopolygalacturonate lyase from *T. maritima maritima*: cloning, characterization and organic synthesis application. Carbohydr Res **337:**1427-33.

55. **Parker, K. N., S. R. Chhabra, D. Lam, W. Callen, G. D. Duffaud, M. A. Snead, J. M. Short, E. J. Mathur, and R. M. Kelly.** 2001. Galactomannanases Man2 and Man5 from *T. maritima* species: growth physiology on galactomannans, gene sequence analysis, and biochemical properties of recombinant enzymes. Biotechnol Bioeng **75:**322-33.

56. **Paulsen, I. T., L. Nguyen, M. K. Sliwinski, R. Rabus, and M. H. Saier, Jr.** 2000. Microbial genome analyses: comparative transport capabilities in eighteen prokaryotes. J Mol Biol **301:**75-100.

57. **Payne, J. W., and M. W. Smith.** 1994. Peptide transport by micro-organisms. Adv Microb Physiol **36:**1-80.

58. **Rajashekhara, E., M. Kitaoka, Y. K. Kim, and K. Hayashi.** 2002. Characterization of a cellobiose phosphorylase from a hyperthermophilic eubacterium, *T. maritima maritima* MSB8. Biosci Biotechnol Biochem **66:**2578-86.

59. **Ren, Q., K. H. Kang, and I. T. Paulsen.** 2004. TransportDB: a relational database of cellular membrane transport systems. Nucleic Acids Res **32 Database issue:**D284-8.

60. **Rice, P., I. Longden, and A. Bleasby.** 2000. EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet **16:**276-7.

61. **Rodionov, D. A., A. A. Mironov, and M. S. Gelfand.** 2001. Transcriptional regulation of pentose utilisation systems in the *Bacillus/Clostridium* group of bacteria. FEMS Microbiol Lett **205:**305-14.

62. **Ruile, P., C. Winterhalter, and W. Liebl.** 1997. Isolation and analysis of a gene encoding alpha-glucuronidase, an enzyme with a novel primary structure involved in the breakdown of xylan. Mol Microbiol **23:**267-79.

63. **Rzhetsky, A., and M. Nei.** 1993. Theoretical foundation of the minimum-evolution method of phylogenetic inference. Mol Biol Evol **10:**1073-95.

64. **Saier, M. H., Jr.** 2000. Families of transmembrane sugar transport proteins. Mol Microbiol **35:**699-710.

65. **Saitou, N., and M. Nei.** 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol **4:**406-25.

66. **Sa-Nogueira, I., and S. S. Ramos.** 1997. Cloning, functional analysis, and transcriptional regulation of the *Bacillus subtilis araE* gene involved in L-arabinose utilization. J Bacteriol **179:**7705-11.

67. **Sanz, Y., F. Toldra, P. Renault, and B. Poolman.** 2003. Specificity of the second binding protein of the peptide ABC-transporter (Dpp) of *Lactococcus lactis* IL1403. FEMS Microbiol Lett **227:**33-8.

68. **Shulami, S., O. Gat, A. L. Sonenshein, and Y. Shoham.** 1999. The glucuronic acid utilization gene cluster from *Bacillus stearothermophilus* T-6. J Bacteriol **181:**3695-704.

69. **Sourdis, J., and M. Nei.** 1988. Relative efficiencies of the maximum parsimony and distance-matrix methods in obtaining the correct phylogenetic tree. Mol Biol Evol **5:**298-311.

70. **Suresh, C., M. Kitaoka, and K. Hayashi.** 2003. A thermostable non-xylanolytic alpha-glucuronidase of *T. maritima maritima* MSB8. Biosci Biotechnol Biochem **67:**2359-64.

71. **Suresh, C., A. A. Rus'd, M. Kitaoka, and K. Hayashi.** 2002. Evidence that the putative alpha-glucosidase of *T. maritima maritima* MSB8 is a pNP alpha-D-glucuronopyranoside hydrolyzing alpha-glucuronidase. FEBS Lett **517:**159-62.

72. **Takami, H., K. Nakasone, Y. Takaki, G. Maeno, R. Sasaki, N. Masui, F. Fuji, C. Hirama, Y. Nakamura, N. Ogasawara, S. Kuhara, and K. Horikoshi.** 2000. Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and genomic sequence comparison with *Bacillus subtilis*. Nucleic Acids Res **28:**4317-31.

73. **Takami, H., Y. Takaki, and I. Uchiyama.** 2002. Genome sequence of *Oceanobacillus iheyensis* isolated from the Iheya Ridge and its unexpected adaptive capabilities to extreme environments. Nucleic Acids Res **30:**3927-35.

74. **Tatusov, R. L., N. D. Fedorova, J. D. Jackson, A. R. Jacobs, B. Kiryutin, E. V. Koonin, D. M. Krylov, R. Mazumder, S. L. Mekhedov, A. N. Nikolskaya, B. S. Rao, S. Smirnov, A. V. Sverdlov, S. Vasudevan, Y. I. Wolf, J. J. Yin, and D. A. Natale.** 2003. The COG database: an updated version includes eukaryotes. BMC Bioinformatics **4:**41.

75. **Thompson, J. D., D. G. Higgins, and T. J. Gibson.** 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res **22:**4673-80.

76. **Tsujibo, H., M. Kosaka, S. Ikenishi, T. Sato, K. Miyamoto, and Y. Inamori.** 2004. Molecular characterization of a high-affinity xylobiose transporter of *Streptomyces thermoviolaceus* OPC-520 and its transcriptional regulation. J Bacteriol **186:**1029-37.

77. **van Helden, J., B. Andre, and J. Collado-Vides.** 2000. A web site for the computational analysis of yeast regulatory sequences. Yeast **16:**177-87.

78. **Vargas, M., and K. M. Noll.** 1996. Catabolite repression in the hyperthermophilic bacterium *T. maritima neapolitana* is independent of cAMP. Microbiology **142 (Pt 1):**139-44.

79. **Velikodvorskaya, T. V., I. Volkov, V. T. Vasilevko, V. V. Zverlov, and E. S. Piruzian.** 1997. Purification and some properties of *T. maritima neapolitana* thermostable xylanase B expressed in *E. coli* cells. Biochemistry (Mosc) **62:**66-70.

80. **Wassenberg, D., W. Liebl, and R. Jaenicke.** 2000. Maltose-binding protein from the hyperthermophilic bacterium *T. maritima maritima*: stability and binding properties. J Mol Biol **295:**279-88.

81. **Wassenberg, D., H. Schurig, W. Liebl, and R. Jaenicke.** 1997. Xylanase XynA from the hyperthermophilic bacterium *T. maritima maritima*: structure and stability of the recombinant enzyme and its isolated cellulose-binding domain. Protein Sci **6:**1718-26.

82. **Winterhalter, C., P. Heinrich, A. Candussio, G. Wich, and W. Liebl.** 1995. Identification of a novel cellulose-binding domain within the multidomain 120 kDa xylanase XynA of the hyperthermophilic bacterium *T. maritima maritima*. Mol Microbiol **15:**431-44.

83. **Wolfinger, R. D., G. Gibson, E. D. Wolfinger, L. Bennett, H. Hamadeh, P. Bushel, C. Afshari, and R. S. Paules.** 2001. Assessing gene significance from cDNA microarray expression data via mixed models. J Comput Biol **8:**625-37.

84. **Wu, L. F., and M. A. Mandrand-Berthelot.** 1995. A family of homologous substrate-binding proteins with a broad range of substrate specificity and dissimilar biological functions. Biochimie **77:**744-50.

85. **Xu, J., M. K. Bjursell, J. Himrod, S. Deng, L. K. Carmichael, H. C. Chiang, L. V. Hooper, and J. I. Gordon.** 2003. A genomic view of the human-*Bacteroides thetaiotaomicron* symbiosis. Science **299:**2074-6.

86. **Xue, Y., and W. Shao.** 2004. Expression and characterization of a thermostable beta-xylosidase from the hyperthermophile, *T. maritima maritima*. Biotechnol Lett **26:**1511-5.

87. **Zverlov, V., K. Piotukh, O. Dakhova, G. Velikodvorskaya, and R. Borriss.** 1996. The multidomain xylanase A of the hyperthermophilic bacterium *T. maritima neapolitana* is extremely thermoresistant. Appl Microbiol Biotechnol **45:**245-7.

88. **Zverlov, V. V., I. Y. Volkov, T. V. Velikodvorskaya, and W. H. Schwarz.** 1997. Highly thermostable endo-1,3-beta-glucanase (laminarinase) LamA from *T. maritima neapolitana*: nucleotide sequence of the gene and characterization of the recombinant gene product. Microbiology **143 (Pt 5):**1701-8.

89. **Zverlov, V. V., I. Y. Volkov, T. V. Velikodvorskaya, and W. H. Schwarz.** 1997. *T. maritima neapolitana bglB* gene, upstream of *lamA*, encodes a highly thermostable beta-glucosidase that is a laminaribiase. Microbiology **143 (Pt 11):**3537-42.

Table 3.1. Carbon sources used in this study

Poly/monosaccharide	Source	Backbone structure	Side chain	Molecular mass (Da)
Arabinose	NA[b]	Ara[b]		150
Glucose	NA	Glc		180
Mannose	NA	Man		180
Rhamnose	NA	Rha		182
Ribose	NA	Rib		150
Xylose	NA	Xyl		150
Galactomannan	Carob	$(Man\beta1\rightarrow 4\ Man)_n$	$Gal(\alpha1\rightarrow 6)$	NA
Glucomannan	Konjac	$(Glc\beta1\rightarrow 4\ Man)_n$		~100,000
β-1,3/β-1,4 glucan	Barley	$(Glc\beta1\rightarrow 3,4\ Glc)_n$		~ 90,000
Laminarin	*L. digitata*	$(Glc\beta1\rightarrow 3\ Glc)_n$		~250,000
Pustulan	*U. papullosa*	$(Glc\beta1\rightarrow 6\ Glc)_n$		~ 5,000
Starch	Potato	$(Glc\alpha1\rightarrow 4\ Glc)_n$		NA
β-xylan	Birchwood	$(Xyl\beta1\rightarrow 4\ Xyl)_n$	O-methyl-D-Glr$(\alpha1\rightarrow 2)$[c]	NA
Chitin	Crab shells	$(Nag\beta1\rightarrow 4\ Nag)_n$[d]		NA

[a] NA, not available.
[b] Ara, arabinose.
[c] Glr, glucuronic acid.
[d] Nag, N-acetylglucosamine.

TABLE 3.2. List of predicted or confirmed sugar transport systems of *T. maritima.*

Family and predicted substrate	Genes	Genomic location	Source(s) or reference(s)
Opp/Dpp ABC family			
β-glucans	*bgtpABCDF* [a]	TM0027-TM0031	This work
Xylan, xylose	*xtpHJLMG* [a]	TM0056-TM0060	This work
Xylan, xylose	*xtpABCDF* [a]	TM0071-TM0075	This work, GI:23270356
	dppABCDF [b]	TM0300-TM0304	Unknown
Xylan	*xtpN* [a]	TM0309	This work
	dppA[b]	TM0460	Unknown
	dppFABC [b]	TM0530-TM0533	Unknown
Rhamnose	*rtpABCDF* [a]	TM1063-TM1067	This work
	dppCDFAB [b]	TM1149-TM1153	Unknown
Lactose	*ltpABCDF* [c]	TM1194, TM1196-9	(51)
Cellobiose, barley	*cbtABCDF* [a,d]	TM1219-TM1223	(11), this work
β-mannans, mannose	*mbtA* [a]	TM1226	This work
β-mannans	*mtpABCDF* [a,d]	TM1746-TM1750	(11), this work
Sugar ABC family			
CUT1 subfamily			
	malE3F3ugpE [e] *malK2*	TM0418-TM0421	Unknown
Uronic acid polysaccharides	*ugpBAE*[e]	TM0430-TM0432	This work
	ugpB, *ugpA*,,*ugpE* [e]	TM0595, TM0596, TM0598	Unknown
N-acetylglucosamine or Nag polysaccharides	*ugpBAE* [e]	TM0810-TM0812	This work
Maltose, maltotriose, β-(1→4) mannotetraose	*malE1F1G1*[f]	TM1202-TM1204	(47)
	ugpBAE [e]	TM1232-TM1235	Unknown
Maltose	*malK1*[a]	TM1276	This work
Maltose, maltotriose, trehalose	*malE2F2* [f]	TM1836, TM1839	(47, 80)
	ugpBAE [e]	TM1853-TM1855	Unknown
CUT2 subfamily			
Monosaccharides?	*rbsC2A2B2*[a]	TM0112, TM0114, TM0115	This work
Ribose, arabinose, xylose	*rbsDB1A1C1* [a]	TM0955-6, TM0958-9	This work
Unknown			
		TM0102-TM0105	Unknown

[a] Designation proposed in this work based on expression results and sequence analysis.
[b] General designation for oligopeptide/dipeptide family transport proteins with unknown specificity.
[c] Designation proposed in this work based on expression results from (51).
[d] Designation proposed in this work based on expression results from (11).
[e] General designation for sugar transport proteins of unknown specificity.
[f] Designation proposed by Nguyen et al. (51).

TABLE 3.S1. Glycoside hydrolases (confirmed and putative) encoded in the *Thermotoga maritima*

Locus	Known/Putative Activity	ENDO or EXO	Signal Peptide^{+}	Reference(s)
TM0024	Laminarinase (LamA/Lam16)	Endo	Y	(7, 9)
TM0025	β-glucosidase/laminaribiase (BglB)	Exo	N	(88)
TM0055	α-glucuronidase (AguA/Agu67)	Exo	N	(62)
TM0061	β-xylanase (Xyl10A)	Endo	Y	(43, 81, 82)
TM0070	β-xylanase (Xyl10B)	Endo	Y	(23, 79)
TM0076	β-xylosidase	Exo	N	(89)
TM0281	α-L-arabinofuranosidase, putative (GH51)	Exo?	N	
TM0305	Endo-1,4-glucanase (Cel74)	Endo	Y	(10)
TM0306	α-L-fucosidase	Exo	N	
TM0308	α-xylosidase, putative (GH31)		N	
TM0310	β-galactosidase, putative (GH42)		N	
TM0364	4-α-glucanotransferase (MgtA)		N	(35)
TM0434	α-glucuronidase (Agu4A)	Exo	N	(71)
TM0433	Pectate lyase (PelA)	Exo	Y	(27, 54)
TM0437	Exo-polygalacturonase	Exo	N	
TM0633	GH73		N	
TM0685	GH23		?	
TM0752	α-glucuronidase (Agu4B)	Exo	N	(70)
TM0767	Maltosyltransferase (MmtA)		N	(42)
TM0809	β-N-acetylglucosaminidase (CbsA)	Exo	N	Unpublished (GI:13242176)
TM0921	GH32		N	
TM1062	β-glucuronidase (putative)	Exo	N	
TM1068	α-glucuronidase (Agu4C)	Exo	N	(70)
TM1192	α-galactosidase (GalA/Gal36A)	Exo	N	(38)
TM1193	β-galactosidase Z (LacZ)	Exo	N	(45)
TM1195	β-galactosidase (BgalA)	Exo	N	(50)
TM1201	Arabinogalactan endo-1,4-β-galactosidase (GH53)		N	
TM1227	β-mannanase (ManB/Man5)	Endo	Y	(55)
TM1231	GH38		N	
TM1281	6-phospho-β-glucosidase (putative)	Exo	N	
TM1414	β-fructosidase (BfrA)	Exo	N	Unpublished (GI:2330880)
TM1524	Endoglucanase (Cel12A)	Endo	N	(36)
TM1525	Endoglucanase (Cel12B)	Endo	Y	(36)

(Table 3.S1, continued)

TM1624	β-mannosidase (ManA/Man2)	Exo	N	(55)
TM1650	α-amylase (AmyB/Amy13B)	Endo	N	(39)
TM1751	Endoglucanase (Cel5A)	Endo	N	(12)
TM1752	Endo-mannanase (Cel5B)	Endo	N	(12)
TM1834	α-glucosidase (AglA)	Exo	N	(6)
TM1835	Cyclomaltodextrinase		N	(34)
TM1840	α-amylase (AmyA/Amy13A)	Endo	Y	(37)
TM1845	Pullulanase (PulA/Pul13A)	Endo	Y	(6)
TM1848	Cellobiose phosphorylase (CepA)	Exo	N	(58)
TM1851	α-mannosidase (cytosolic)	Exo	N	(46)

[+]SignalP V1.1 (http://genome.cbs.dtu.dk/services/SignalP/)

TABLE 3.S2. *T. maritima* ABC transport systems examined in this work. In column "Putative regulator binding site(s)," predicted σ^A promoters are underlined or italicized where predicted sites overlap, predicted LacI or XylR regulator binding sites are in bold, and inverted repeat sequences are italicized.

Gene ID	Protein	Putative/Confirmed* Function	Relevant Homologs	Putative binding site(s)	Position
ORFs related to β-glucan utilization					
TM0024	LamA	laminarinase	91% id/646 aa (*T. neapolitana* LamA) 61%/249 aa (PF0076)		
TM0025	BglB/Cel3	β-glucosidase	85% id/721 aa (TN_BglB)		
TM0026		hypothetical protein			
TM0027	Bgtp	ABC ATP-binding protein	52% id/265 aa (TM1219 CbtF) 51% id/264 aa (PF1213 CbtF)		
TM0028	BgtpD	ABC ATP-binding protein	40% id/323 aa (TM1220 CbtD) 47% id/321 aa (PF1212 CbtD)		
TM0029	BgtpC	ABC permease	42% id/305 aa (PF1211 CbtC) 42% id/271 aa (TM1221 CbtC)		
TM0030	BgtpB	ABC permease	43% id/328 aa (PF1210 CbtB) 43% id/328 aa (TM1222 CbtB)	GTGATCTTTT*CCCGGCCCC*ACTTC*GGGGCCGGG*TCCACAAAAA Inverted repeat	*-54* *-33*
TM0031	BgtpA	ABC substrate-binding protein	34% id/567 aa (PF1209 CbtA) 33% id/554 aa (TM1223 CbtA) 31% id/572 aa (TM1226)		
TM0032	BglcR	Regulator of β-glucan uptake, putative	33% id/390 aa (TM0110)	TTGAGCTCCTTGGAAAGAGCATCGGGAATAAAAT TTGTACTTTTTGAGTCATATCGTTATAAT	-201 -168 -71 -43
TM0312		oxidoreductase, putative	67% id/325 aa (PF0554)	TTGAA**ATGAAAACATTTTCA**GACTATAAT	-54 -26 **-49** **-35**
TM1218	CelR	regulator of cellobiose uptake, putative	33% id/327 aa (TM1200)		
TM1219	CbtF	ABC ATP-binding protein	58% id/318 aa (PF1213)		
TM1220	CbtD	ABC ATP-binding protein	60% id/300 aa (PF1212)		
TM1221	CbtC	ABC permease	59% id/301 aa (PF1211)		
TM1222	CbtB	ABC permease	67% id/325 aa (PF1210)	ATGA*CCCCGCCCC*TC*GGGGCGGGG*TTTTAAAAA Inverted repeat	-51 -19 *-47* *-23*
TM1223	CbtA	cellobiose-binding protein	60% id/557 aa (TM1226) 56% id/564 aa (PF1209, CbtA)	TTGGA**ATGTAAACATTTTCA**CTGTACATTTACA	-75 -47 **-70** **-56**
TM1524	CelA	endoglucanase	73% id/256 aa (*T. neapolitana* CelA) 50% id/253 aa (TM1525)	TTGA**ATGTAAACATTTTCA**TAATAAGAT	-45 -18 **-41** **-27**
TM1525	CelB	endoglucanase	86% id/272 aa (*T. neapolitana* CelB)		

(Table 3.S2, continued)

TM1848	CepA	cellobiose phosphorylase	92% id/813 aa (*T. neapolitana* CbpA)	TTGA**ATGAAAACATTTTCA**GAATAAAAT	-63 **-59**	-36 **-45**

ORFs related to xylose, ribose and xylan utilization

TM0055	AguA/Agu67	α-glucuronidase	56% id/683 aa (GT_AguA))	TTGATGTGAATCTTGGTGAGGAAGGTTTCAT	-66	-36
TM0056	XtpG	ABC substrate binding protein	69% id/664 aa (TM0309) 41% id/687 aa (TM1067)			
TM0057	XtpM	ABC ATP-binding protein	70% id/338 aa (TM1063)			
TM0058	XtpL	ABC ATP-binding protein	61% id/331 aa (TM1064)			
TM0059	XtpJ	ABC permease protein	73% id/356 aa (TM1065)			
TM0060	XtpH	ABC permease protein	62% id/328 aa (TM1066)			
TM0061	XynA/Xyl10A	xylanase A	88% id/1057 aa (TN_XynA)	**TTTTCTTTACAAAAAATAA**CTTTAGGGTGATATAAT TTGAAAGTTCCTTTCACAATAATGGTATAAT	**-84** -79 -137	**-64** -49 -107
TM0070	XynB/Xyl10B	xylanase B	83% d/346 aa (TN_XynB)	TTTCTC**TTTTTTCCTTTCTGGTATTAT**ATTTTCGTTAAATCAT TTGTATGTTATA**TTTTTTCCTTTCGGGAAATAT**TTTAATGTAA	-57 **-57** -200 -193 **-188**	-37 **-34** -168 -162 **-168**
TM0071	XtpA	ABC substrate binding protein	76% id/626 aa (TN_XtpA)	TTGATTTTCTCGGAGATTTACATTAAAATATTTCCCGAAAGGAAAAAAT	-79 -63	-51 -31
TM0072	XtpB	ABC permease protein	84% id/332 aa (TN_XtpB)	AAACAGGGGG*CCGGGGG*GACAA*CCCCCGG*TCCTTTAAAGAGAA Inverted repeat	-48	-30
TM0073	XtpC	ABC permease protein	94% id/287 aa (TN_XtpC)			
TM0074	XtpD	ABC ATP binding protein	95% id/332 aa (TN_XtpD)			
TM0075	XtpF	ABC ATP binding protein	94% id/320 aa (TN_XtpF)			
TM0076	Xyl3	β-xylosidase	93% id/778 aa (*T. neapolitana* XloA)			
TM0077	AxeA	acetyl xylan esterase	95% id/395 aa (*T. neapolitana* AxeA			
TM0110		XylR-family regulator	33% id/390 aa (TM0031)	TCCGACCTTGATTTTAAATTATTTCCTGCATATAAT	-60	-40
TM0111		Alcohol dehydrogenase	72% id/390 aa (PF0075 AdhB)			
TM0112		ABC sugar permease protein	44% id/296 aa (TTE0764)			
TM0113		xylan deacetylase?	54% id/200 aa (Chte02001934)			
TM0114		ABC sugar-binding protein	36% id/291 aa (TTE0765)			
TM0115		ABC ATP-binding protein	48% id502 aa (TTE0763)			
TM0116	XylB	Xylulokinase	42% id/495 aa (TTE_XylB) 38% id/491 aa (BSU17610 XylB)			

(Table 3.S2, continued)

TM0308		α-xylosidase, putative (GH31)	54% id/761 aa (BH1905) 46% id/738 aa (*L. pentosus* XylQ)	TCGAACTCTGC**GTGAAAACAGTTTCA**TGATAAT **TGAAAAAGCCTTCA**GAAAAGAAACTCGACGTTAATCTTCTATAAT AATGTTTTTA**TTTTTTCCTCTAAAGAAATTT**CTTCGTTATTAT	-59 -27 **-48 -34** **-134 -120** -123 -88 **-290 -270** -288 -258
TM0309	XtpN	ABC substrate-binding protein	69% id/664 aa (TM0056)		
TM0949	RbsR	regulator of ribose operon, utative	37% id/334 aa (TTE0803)	TTGACGAGTGTTGCTGTTGACTCTTTATATT TTGAAGGAGTAGCAGCAATGGGAGGTTAAAAG TTAAAAGAATAAGGCTTCTCATCAGTGATAAAAT	-149 -119 -95 -64 -70 -37
TM0950		hypothetical protein	41% id/518 aa (LJ1257)		
TM0951	RpiB	Ribose-5-P isomerase, putative	51% id/466 aa (YPO3311)		
TM0952		Xylulokinase, putative	50% id/489 aa (TM1430 GK2)		
TM0953	Tkt_C	Transketolase, C-terminal subunit	53% id/302 aa (Chte02002569) 51% id/272 aa 51%/307 aa (Lmo1033)		
TM0954	Tkt_N	Transketolase, N-terminal subunit	58% id/268 aa (Lmo1032) 51% id/272 aa (Chte02003137)		
TM0955	RbsC	ABC permease protein	56% id/317 aa (RB3497)		
TM0956	RbsA	ABC ATP-binding protein	51% id/504 aa (RB3496)		
TM0957		hypothetical protein	22% id/218 aa (RB3495)		
TM0958	RbsB	ABC substrate-binding protein	58% id/320 aa (RB3493)		
TM0959	RbsD	cytoplasmic ribose-binding protein	47% id/133 aa (L85737) 43%id/134 aa (SAG0117)		
TM0960	RbsK	ribokinase	42% id/303 aa (TTE0202)	**TGAAAACGATTTCGAAAACGATTTCA**T	-43 -17 **-43 -31** **-30 -18**
TM1280		N-acetylglucosamine kinase, putative	28% id/301 aa (SSO3218) 28% id/264 aa (TTE0216)		
TM1667	TM_XylA	Xylose isomerase	95% id/444 aa (TN_XylA)	TTGAAATGATACCCCAAGATTTTATATAAT	-63 -34
TM1668		Hypothetical protein	45 id/124 aa (TTC1244)		

ORFs related to mannan utilization

TM1224	ManR	XylR family regulator	27% id/376 aa (BSU17590 BS_xylR)
TM1225		Uncharacterized conserved protein	58% id/318 aa (BT1033) 37% id/303 aa (Chte02001674)
TM1226	MbtA	ABC substrate-binding protein	60% id/557 aa (TM1223)

(Table 3.S2, continued)

TM1227	ManB/Man5	Endo-1,4-β-mannosidase	84% id/666 aa (TN_ManB)	TTGACTTTGCAGAAGTCCGAATGGTACATT TTCACAAATAAAACTCGAAAATCACGAAAAAT	-87 -58 -124 -93
TM1624	ManA/Man2	β-mannosidase	80% id/785 aa (TN_Man2)	TCGAAAAACAAAAGGAGGGTAAAAA	-106 -82
TM1746	MtpA	ABC substrate-binding protein	31% id/431 aa (BB0328)	TTAACCCTGTCGCTTCGAAATGACAATAAT TTGGAACTCAATATCACAGTCTGATAAAAT	-193 -164 -98 -69
TM1747	MtpB	ABC permease protein	42% id/316 aa (PM70)		
TM1748	MtpC	ABC permease protein	45% id/277 aa (TTE0613)		
TM1749	MtpD	ABC ATP-binding protein	56% id/311 aa (TM0501)		
TM1750	MtpF	ABC ATP-binding protein	56% id/322 aa (CAC3635)		
TM1751	Cel5A	endoglucanase	51% id/302 aa (RB5256)	TTGAAAAATTTTTTATTGTTTTTGCTTATGAT	-54 -23
TM1752	Cel5B	endoglucanase	28% id/312 aa (Chte02001321)		
ORFS related to arabinose utilization					
TM0275	AraR	LacI family regulator of arabinose utilization	33% id/341 aa (*Geobacillus stearothermophilis* AraR)	GTGATTTGCAATCCTTCCAAAATACAGAATAAT	-43 -11
TM0276	AraA	Arabinose isomerase	94% id/496 aa (*T. neapolitana* AI) 62% id/395 aa (BSU28800 AraA)	**TGGATAACGGTACGTACCGATGGTAGTATATCAT**	**-167 -134**
TM0277		Frameshift, ABC binding protein			
TM0278		Sugar ABC permease subunit	36% id/291 aa (PF1740 MalF)		
TM0279		Sugar ABC permease subunit	37% id/274 aa (DR1436)	TTGAGTGTTCTGCTATTGATTGTGGCGATAGT	-95 -64
TM0280		Uncharacterized conserved protein; also found in *B. halodurans* L-arabinose operon	40% id/630 aa (BT3674) 23% id/454 aa (BH1877)	TTGTCCATTTTTGTTAAATCAAGTATATT TTGA*TTT*AACAAAAATGGACAATGTAGAAT TGGACAATGTAGAATTTGCGTAAAGAATAAT	-91 -63 -84 -55 -80 -55 -69 -39
TM0281	AbfA	α-L-arabinofuranosidase	36% id /505 aa (BSU28720 AbfA) 34% id/457 aa (BL1156 AbfA2)		
TM0282		Aldose-1-epimerase	45% id/354 aa (PSPTO236 GalM)	TTGAGGTAGAATTGGAGTAAGAAAAA	-48 -23
TM0283	AraD	L-ribulose-5-P-4-epimerase	46%id/212 aa (PPA0882 AraD)	TTGAAGCACAGCACTTCCCCGATTCTCCAAAT	-66 -35
TM0284	AraB	L-ribulokinase, putative	45% id/506 aa (Bcepa02003975) 36% id/504 aa (Mther02000909)	CTGACGAGGTGGATCATCTCTACAAT	-67 -42
TM0285	AraM	Uncharacterized conserved protein also found in *B. subtilis* arabinose utilization operon	29% id/382 aa (WD0787 AraM)		
ORFS related to rhamnose utilization					
TM1060		Putative mfs permease	COG0477		

(Table 3.S2, continued)

TM1061		Uncharacterized conserved protein	43% id/386 aa (BT2916) 42% id/381 aa (Mdeg0812)	
TM1062		β-glucuronidase	40% id/569 aa (SSO3036, gusB)	
TM1063	RtpF	ABC ATP-binding protein	70%/338 aa (TM0057)	
TM1064	RtpD	ABC ATP-binding protein	61% id/331 aa (TM0058)	
TM1065	RtpC	ABC permease protein	73% id/356 aa (TM0059)	
TM1066	RtpB	ABC permease protein	62% id/328 aa (TM0060)	TTGAAAGTATCCGGAGAAGACGTGGATATT -178 -149 CGATAG*GGGAGGG*CCT*CCCTCCC*TTTATTTA -54 -38 Inverted repeat
TM1067	RtpA	ABC substrate-binding protein	43% id/680 aa (TM0309) 41% id/681 aa (TM0056)	TTGACCTTGTTGTATGTTTGTTGATAT*TATATT* -92 -63 -92 -60
TM1068	Agu4C	α-glucuronidase	99% id/464 aa (TM0434) 62% id/465 aa (TM0752)	TTGAAAGCAACGTACATTCTCGTCACAAAGT -175 -145 TTGAGGAGGCACTTGAAGGGGCAGATTTTGTCAT -96 -63
TM1069	RhaR	DeoR family regulator	31% id/248 aa (BH1553) 32% id/248 aa (OB0498) 23% id/256 aa (EC DeoR)	TTGAACTTCGTTGAAGATGGAGACAT*TATCTT* -46 -18 -46 -15 *TTA*ATCACAAAAGTCAACGATATTTGTGATAAAAT -61 -27 -58 -27
TM1070		Hypothetical protein conserved in bacteria (COG4288)	39% id/116 aa (Rxyl023019)	**ATGAA**T**TCGTACCCA**GAGGTGATGAATAAT **-28 -14** -27 +9 TTCATATCAACACC**TCCTTTCCTTTCGAAAAATTT**TATCAC -112 -80 **-98 -78**
TM1071	RhaI	Rhamnose isomerase	39% id/386 aa (*Rhizobium leguminosaurum* bv. *trifolii* RhaI); 23%/360 aa (BT3764)	
TM1072	RhaD	L-rhamnose aldolase	37% id/237 aa (lp_3592) 31%/252 aa (BT3766)	
TM1073	RhaB	Rhamnulokinase	40% id/465 aa (*Clostridium perfringens* rhamnulokinase)	
TM1074		Hypothetical protein	30% id/86 aa, 24% id/381 aa, 35% id/73 aa (SCO0370)	

ORFs related to use of unknown substrate

TM1850		Hypothetical protein	-
TM1851		α-mannosidase	37% id/1058 aa (Rxyl136101)
TM1852		Glycosylase, putative	51% id/298 aa (TTE0063)
TM1853		ABC sugar permease protein	52% id/274 aa (YPO2475)
TM1854		ABC sugar permease protein	54% id/288 aa (YPO2476)
TM1855		ABC sugar substrate binding protein	50% id/400 aa (YPO2477)
TM1856		LacI family regulator	36% id/343 aa (TTE0803)

TGTAAGCGTTTACATGAATTGCTGACTGGAATTAATCC **-125 -112**
-117 -89

TTGATTTAACACGGAAATTTTTTGTATAAACGATA**TGTAAGCGCTTACA** -74 -33
-27 -14

FIGURE LEGENDS

FIG. 3.1. Loop design used for the study of carbon source utilization of *T. maritima* in this study. The arrowheads correspond to the Cy5 channel and the dotted arrow ends correspond to the Cy3 channel. Abbreviations for sugar names used in subsequent expression histograms are shown in parentheses.

FIG. 3.2. Circular representation of the *T. maritima* genome showing locations of known carbohydrate transport proteins and Opp/Dpp family ABC transporter components. Least squares mean estimates (see Methods) of transcript levels corrected for systematic errors are shown for selected operons whose carbohydrate specificity is predicted in this work. In this context, red and green denote transcript levels above (red) and below (green) the mean expression across all genes, where 0 represents the mean rather than no expression. Oligopeptide transporter subunits are represented in black, CUT1 transporter subunits in gray and CUT2 transporter subunits in white.

FIG. 3.3. Pentose-responsive loci of *T. maritima*. Predicted σ^A promoters are represented by arrows. Substrate binding proteins are outlined in bold and boxed. Spacing between genes is less than 30 bases unless indicated otherwise. A.) A *T. maritima* locus which responds to the pentose sugars ribose, arabinose, and xylose contains genes for the utilization of the simple sugar D-ribose, a likely ribose transport system, and other genes likely to participate in the pentose phosphate pathway. B.) An arabinose utilization locus contains genes for the conversion of arabinose into D-xylulose-5-P. C.) Predicted pathway for the hydrolysis, transport, and utilization of xylan, xylose, ribose and arabinose by *T. maritima*. Extracellular enzymes

responsible for polysaccharide hydrolysis are shown, as well as periplasmic binding proteins, membrane-embedded permeases, associated ATP-binding subunits, and intracellular hydrolases. References for hydrolases shown in the pathway are listed in supplemental Table 3.S2.

FIG. 3.4. Representative phylogenetic tree of substrate binding proteins of peptide family transporters from *T. maritima*. All operons are shown on one strand to more clearly represent the relative positions of subunits. Black arrows represent substrate binding proteins of the DppA family (COG0747), white arrows with diagonal stripes represent substrate binding proteins of the OppA family (COG0747), dark gray arrows represent permease subunits of the DppB/OppB family (COG0601), light gray arrows represent permease subunits of the DppC/OppC family (COG1173), white arrows represent ATP-binding subunits of the DppD/OppD family (COG0444), and white arrows with spots represent ATP-binding subunits of the DppF/OppF family (COG4608). Other non-transport related genes located between transporter subunits are represented as dotted arrows. Asterisks (*) represent apparent lineage-specific gene expansions which have taken place since the divergence of *T. maritima* and the next closest fully sequenced organism. Black hairpins represent locations of GC-rich inverted repeats. The tree topology represented here is consistent with trees constructed using protein sequences for ATP-binding and permease subunits and was not altered by using pairwise or complete deletion of missing data, or by tree construction method (neighbor-joining, minimum evolution, maximum parsimony).

FIG. 3.5. Expression results for transcripts detected at higher levels on β-linked polysaccharides. Small hairpin symbols represent locations of GC-rich inverted repeats, while large hairpin

symbols represent locations of predicted rho-independent terminators (http://www.tigr.org/software/TransTermResults/btm.html). Predicted σ^A promoters are represented by arrows, and an asterisk denotes the position of a putative cellobiose regulator operator. Substrate binding proteins are outlined in bold and boxed. Spacing between genes is less than 30 bases unless indicated otherwise. A.) Genes within a putative cellobiose transport operon (proposed designation CbtABCDF) including a likely regulator of cellobiose uptake and utilization (proposed designation CelR) and a co-localized mannan-responsive locus (including TM1224, proposed designation ManR, and TM1226, proposed designation MbtA). B.) Genes within a glucomannan and galactomannan responsive locus include the Opp transporter components TM1746-TM1752 (proposed designation MtpABCDF), Cel5A, and Cel5B and the b-mannosidase TM1624. C.) Genes within a β-glucan responsive locus (proposed designation BgtpABCDF) including a putative regulator of β-glucan uptake (proposed designation BglcR). D.) Predicted pathway for the utilization of β-glucans and glucomannan by *T. maritima*. Extracellular enzymes responsible for polysaccharide hydrolysis are shown, as well as periplasmic binding proteins, membrane-embedded permeases, associated ATP-binding subunits, and intracellular hydrolases. References for hydrolases shown in the pathway are listed in supplemental Table 3.S1.

FIG. 3.6. β-Xylan and xylose-responsive operons from groups 2 and 3 of Opp/Dpp family transporters. Small hairpin symbols represent locations of GC-rich inverted repeats found between substrate-binding proteins and other transporter subunits, while large hairpin symbols represent locations of predicted rho-independent terminators (http://www.tigr.org/software/TransTermResults/btm.html). Predicted σ^A promoters are represented by arrows, and

an asterisk denotes the positions of putative xylan/xylose regulator operator. Substrate binding proteins are outlined in bold and boxed. Spacing between genes is less than 30 bases unless indicated otherwise. A.) XtpGHJLM, a predicted xylooligosaccharide transport system, is divergently transcribed from xylanase Xyl10A. B.) XtpABCDF, a predicted xylose/xyloside transport system, is divergently transcribed from xylanase Xyl10B.

FIG. 3.7. Rhamnose responsive locus containing Opp/Dpp family transporter from group 3 of Opp/Dpp family transporters. Small hairpin symbols represent locations of GC-rich inverted repeats found between substrate-binding proteins and other transporter subunits. Predicted σ^A promoters are represented by arrows. Substrate binding proteins are outlined in bold and boxed. Spacing between genes is less than 30 bases unless indicated otherwise. A.) A rhamnose-responsive locus contains candidate genes likely to encode enzymes responsible for the transport and hydrolysis of rhamnose-containing di- or oligosaccharides, and the complete catabolism of the simple sugar L-rhamnose. B.) A predicted pathway for the utilization of L-rhamnose by *T. maritima*. A periplasmic binding protein, membrane-embedded permeases, associated ATP-binding subunits, and rhamnose catabolic enzymes are shown.

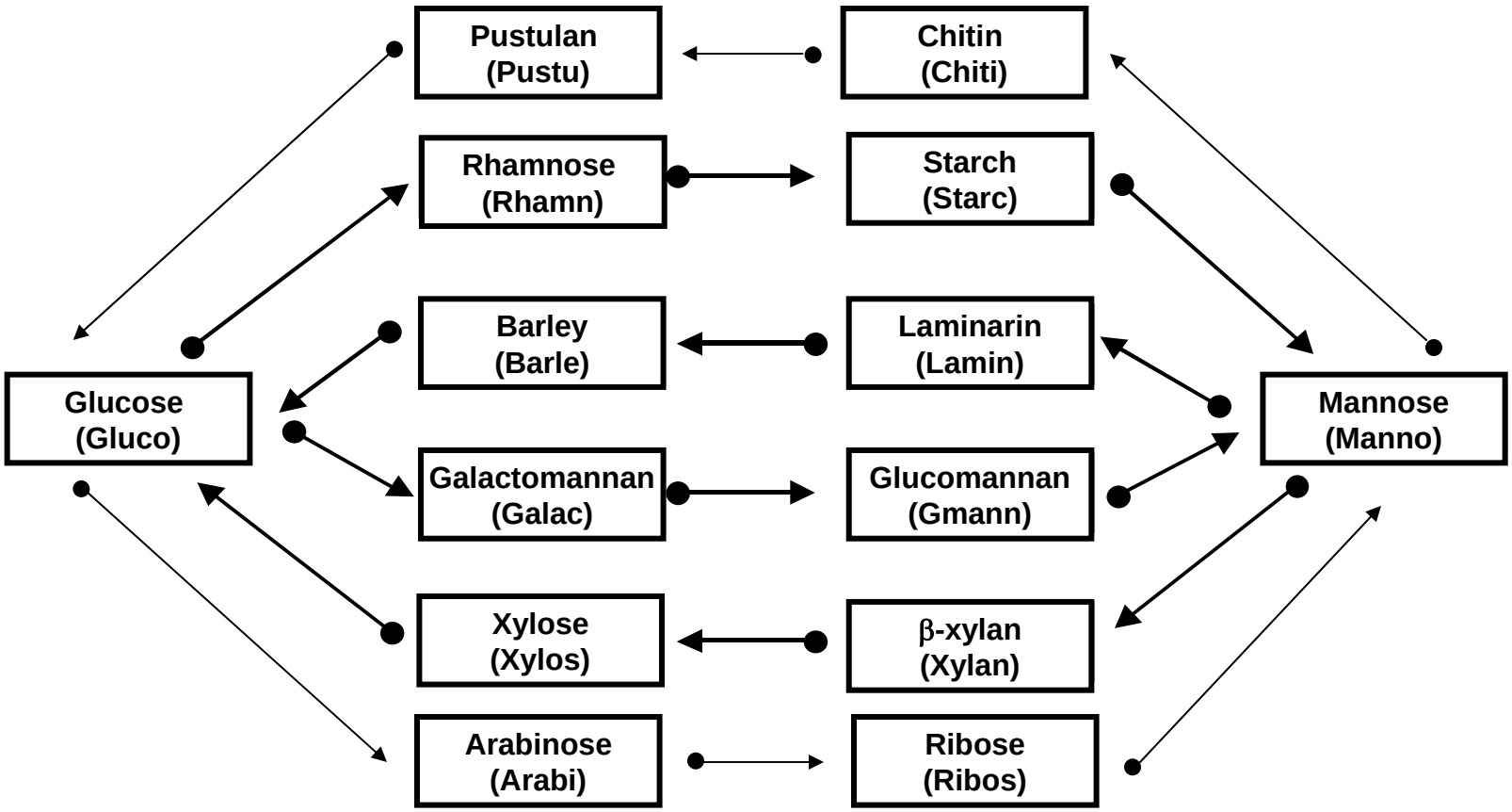

Pustulan
(Pustu)
Chitin
(Chiti)
Rhamnose
(Rhamn)
Starch
(Starc)
Barley
(Barle)
Laminarin
(Lamin)
Glucose
(Gluco)
Mannose
(Manno)
Galactomannan
(Galac)
Glucomannan
(Gmann)
Xylose
(Xylos)
β-xylan
(Xylan)
Arabinose
(Arabi)
Ribose
(Ribos)

Figure 3.1

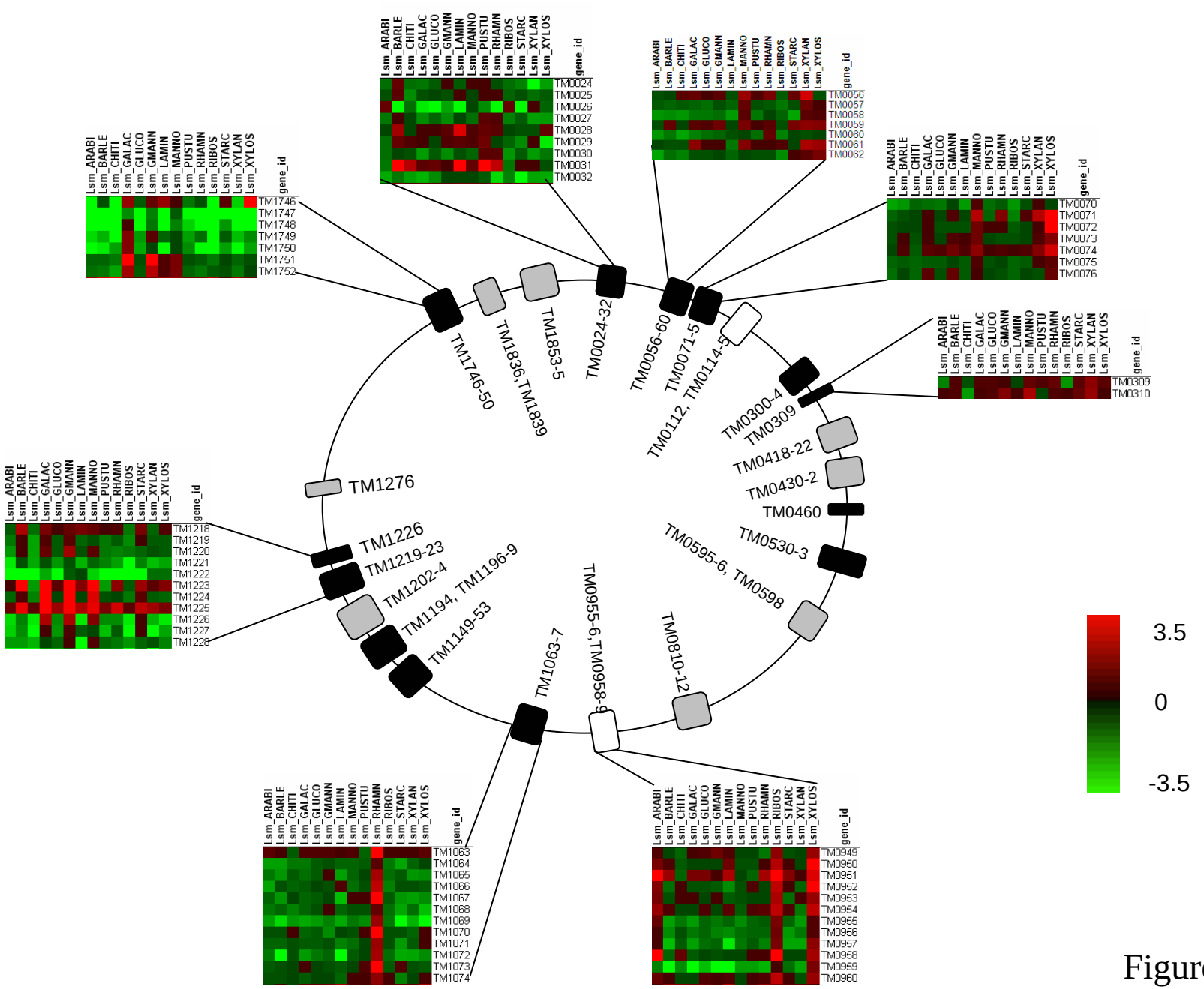

Lsm_ARABI Lsm_BARLE Lsm_CHITI Lsm_GALAC Lsm_GLUCO Lsm_GMANN Lsm_LAMIN Lsm_MANNO Lsm_PUSTU Lsm_RHAMN Lsm_RIBOS Lsm_STARC Lsm_XYLAN Lsm_XYLOS gene_id
TM0024 TM0025 TM0026 TM0027 TM0028 TM0029 TM0030 TM0031 TM0032
TM0056 TM0057 TM0058 TM0059 TM0060 TM0061 TM0062
TM0070 TM0071 TM0072 TM0073 TM0074 TM0075 TM0076
TM0309 TM0310
TM0949 TM0950 TM0951 TM0952 TM0953 TM0954 TM0955 TM0956 TM0957 TM0958 TM0959 TM0960
TM1063 TM1064 TM1065 TM1066 TM1067 TM1068 TM1069 TM1070 TM1071 TM1072 TM1073 TM1074
TM1218 TM1219 TM1220 TM1221 TM1222 TM1223 TM1224 TM1225 TM1226 TM1227 TM1228
TM1746 TM1747 TM1748 TM1749 TM1750 TM1751 TM1752
TM1746-50
TM1836,TM1839
TM1853-5
TM0024-32
TM0056-60
TM0071-5
TM0112, TM0114-5
TM0300-4
TM0309
TM0418-22
TM0430-2
TM0460
TM0530-3
TM0595-6, TM0598
TM0810-12
TM0955-6,TM0958-9
TM1063-7
TM1149-53
TM1194, TM1196-9
TM1202-4
TM1219-23
TM1226
TM1276
3.5
0
-3.5

Figure 3.2

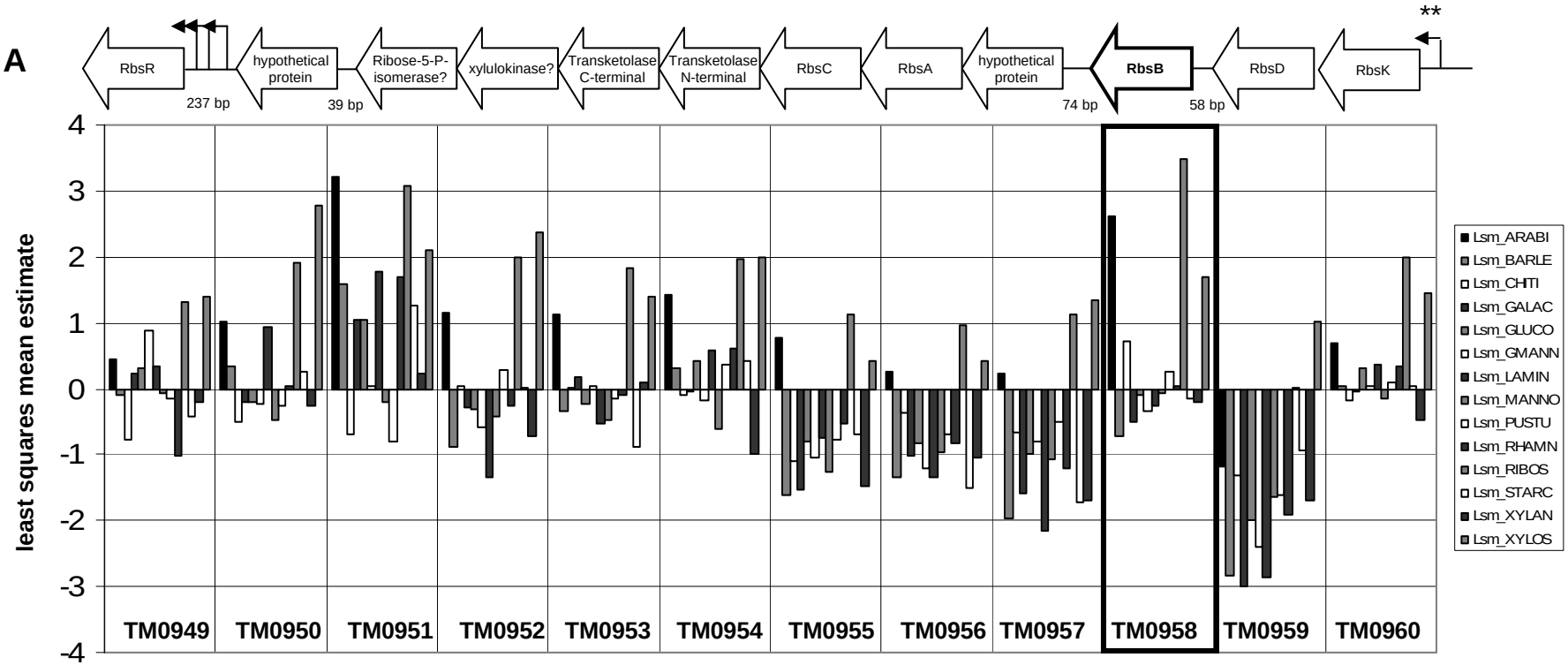

A
RbsR
237 bp
hypothetical protein
39 bp
Ribose-5-P-isomerase?
xylulokinase?
Transketolase C-terminal
Transketolase N-terminal
RbsC
RbsA
hypothetical protein
74 bp
RbsB
58 bp
RbsD
RbsK
**
least squares mean estimate
4
3
2
1
0
-1
-2
-3
-4
TM0949
TM0950
TM0951
TM0952
TM0953
TM0954
TM0955
TM0956
TM0957
TM0958
TM0959
TM0960
Lsm_ARABI
Lsm_BARLE
Lsm_CHITI
Lsm_GALAC
Lsm_GLUCO
Lsm_GMANN
Lsm_LAMIN
Lsm_MANNO
Lsm_PUSTU
Lsm_RHAMN
Lsm_RIBOS
Lsm_STARC
Lsm_XYLAN
Lsm_XYLOS

Figure 3.3A

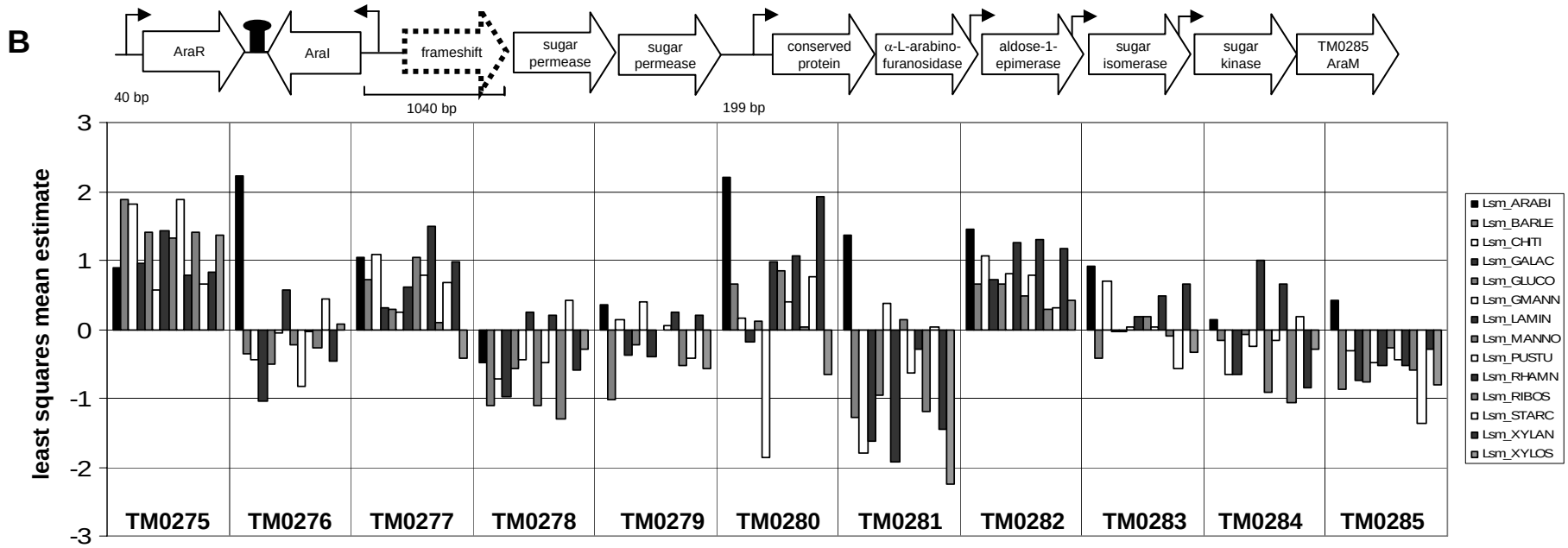

B
AraR
AraI
frameshift
sugar permease
sugar permease
conserved protein
α-L-arabino-furanosidase
aldose-1-epimerase
sugar isomerase
sugar kinase
TM0285 AraM
40 bp
1040 bp
199 bp
least squares mean estimate
3
2
1
0
-1
-2
-3
TM0275
TM0276
TM0277
TM0278
TM0279
TM0280
TM0281
TM0282
TM0283
TM0284
TM0285
Lsm_ARABI
Lsm_BARLE
Lsm_CHITI
Lsm_GALAC
Lsm_GLUCO
Lsm_GMANN
Lsm_LAMIN
Lsm_MANNO
Lsm_PUSTU
Lsm_RHAMN
Lsm_RIBOS
Lsm_STARC
Lsm_XYLAN
Lsm_XYLOS

Figure 3.3B

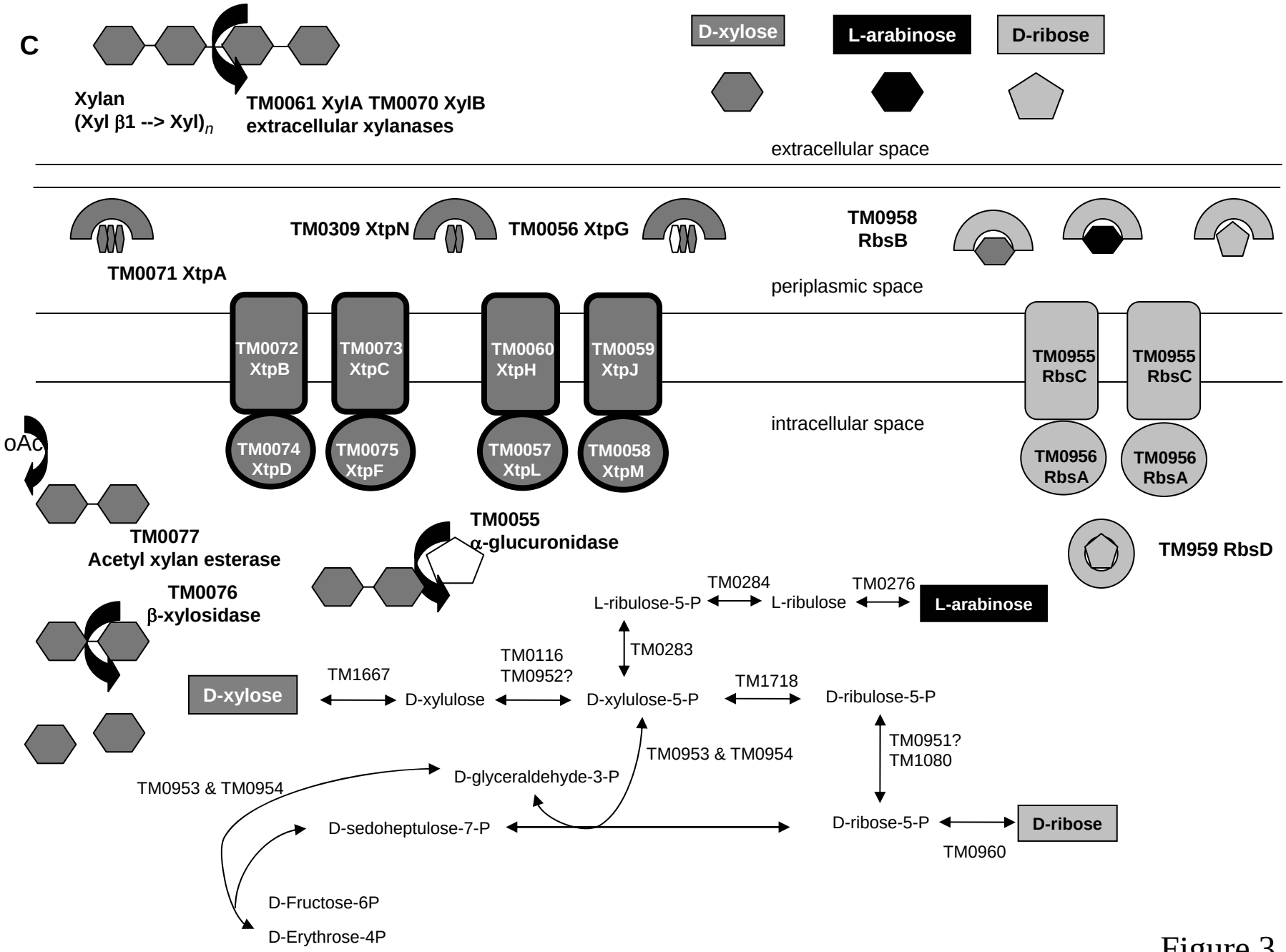

C
Xylan
(Xyl β1 --> Xyl)n
TM0061 XylA TM0070 XylB
extracellular xylanases
D-xylose
L-arabinose
D-ribose
extracellular space
TM0309 XtpN
TM0056 XtpG
TM0958
RbsB
TM0071 XtpA
periplasmic space
TM0072
XtpB
TM0073
XtpC
TM0060
XtpH
TM0059
XtpJ
TM0955
RbsC
TM0955
RbsC
intracellular space
TM0074
XtpD
TM0075
XtpF
TM0057
XtpL
TM0058
XtpM
TM0956
RbsA
TM0956
RbsA
oAc
TM0077
Acetyl xylan esterase
TM0055
α-glucuronidase
TM959 RbsD
TM0076
β-xylosidase
L-ribulose-5-P
TM0284
L-ribulose
TM0276
L-arabinose
TM0283
TM0116
TM0952?
TM1667
D-xylose
D-xylulose
D-xylulose-5-P
TM1718
D-ribulose-5-P
TM0951?
TM1080
TM0953 & TM0954
D-glyceraldehyde-3-P
TM0953 & TM0954
D-sedoheptulose-7-P
D-ribose-5-P
D-ribose
TM0960
D-Fructose-6P
D-Erythrose-4P

Figure 3.3C

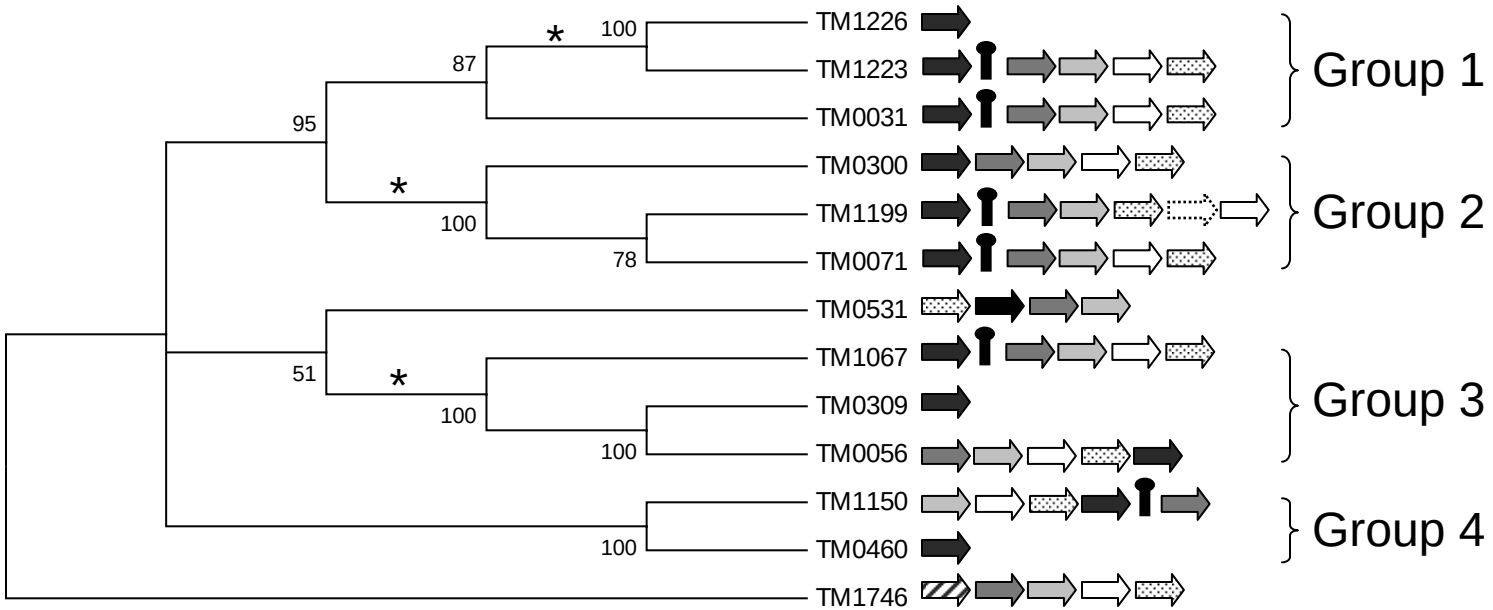

TM1226
TM1223
TM0031
TM0300
TM1199
TM0071
TM0531
TM1067
TM0309
TM0056
TM1150
TM0460
TM1746
100
87
95
100
78
51
100
100
100
*
*
*
Group 1
Group 2
Group 3
Group 4

Figure 3.4

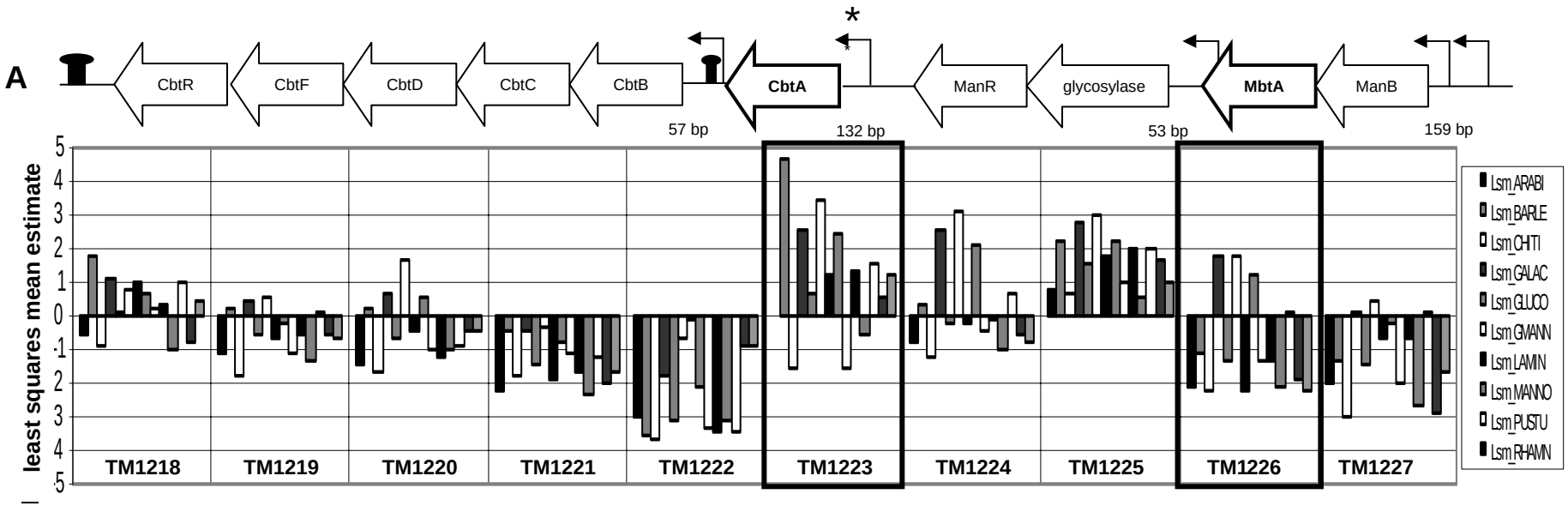

A
CbtR
CbtF
CbtD
CbtC
CbtB
CbtA
ManR
glycosylase
MbtA
ManB
57 bp
132 bp
53 bp
159 bp
least squares mean estimate
5
4
3
2
1
0
-1
-2
-3
-4
-5
TM1218
TM1219
TM1220
TM1221
TM1222
TM1223
TM1224
TM1225
TM1226
TM1227
Lsm_ARABI
Lsm_BARLE
Lsm_CHITI
Lsm_GALAC
Lsm_GLUCO
Lsm_GMANN
Lsm_LAMIN
Lsm_MANNO
Lsm_PUSTU
Lsm_RHAMN

Figure 3.5A

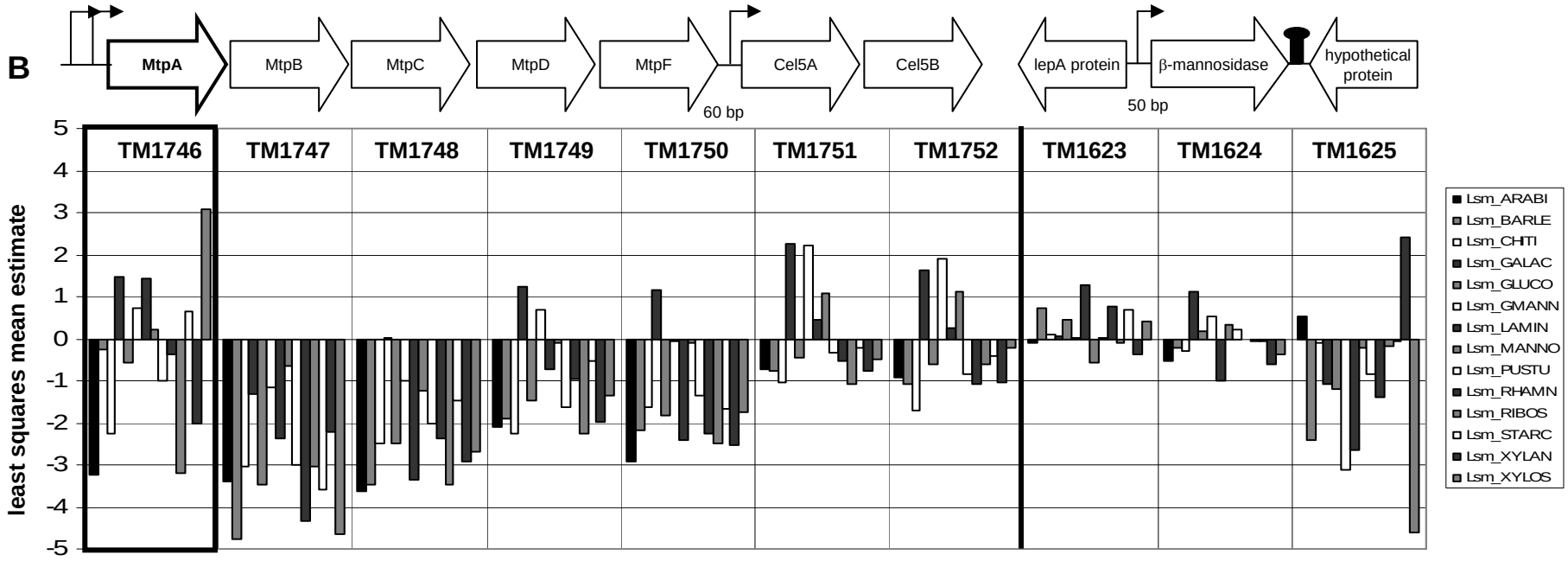

B
MtpA
MtpB
MtpC
MtpD
MtpF
Cel5A
Cel5B
lepA protein
β-mannosidase
hypothetical protein
60 bp
50 bp
TM1746
TM1747
TM1748
TM1749
TM1750
TM1751
TM1752
TM1623
TM1624
TM1625
least squares mean estimate
5
4
3
2
1
0
-1
-2
-3
-4
-5
Lsm_ARABI
Lsm_BARLE
Lsm_CHITI
Lsm_GALAC
Lsm_GLUCO
Lsm_GMANN
Lsm_LAMIN
Lsm_MANNO
Lsm_PUSTU
Lsm_RHAMN
Lsm_RIBOS
Lsm_STARC
Lsm_XYLAN
Lsm_XYLOS

Figure 3.5B

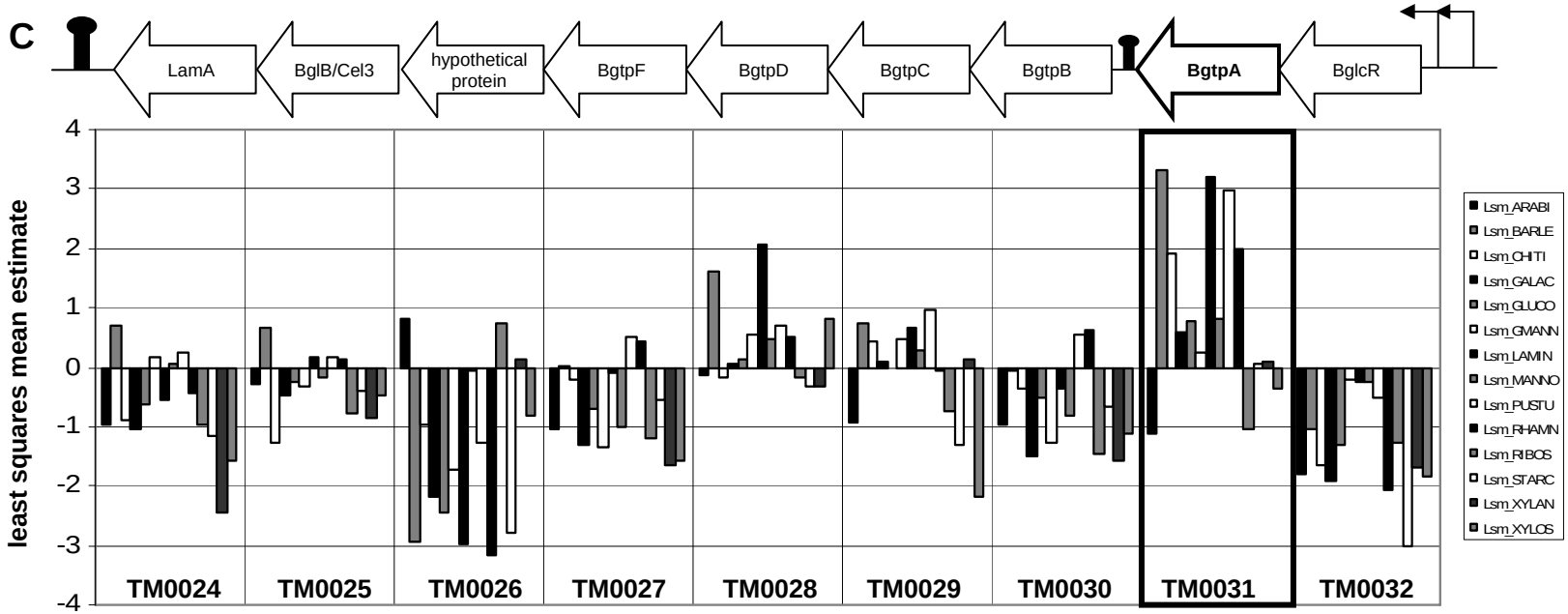

C
LamA
BglB/Cel3
hypothetical protein
BgtpF
BgtpD
BgtpC
BgtpB
BgtpA
BglcR
least squares mean estimate
4
3
2
1
0
-1
-2
-3
-4
TM0024
TM0025
TM0026
TM0027
TM0028
TM0029
TM0030
TM0031
TM0032
Lsm_ARABI
Lsm_BARLE
Lsm_CHITI
Lsm_GALAC
Lsm_GLUCO
Lsm_GMANN
Lsm_LAMIN
Lsm_MANNO
Lsm_PUSTU
Lsm_RHAMN
Lsm_RIBOS
Lsm_STARC
Lsm_XYLAN
Lsm_XYLOS

Figure 3.5C

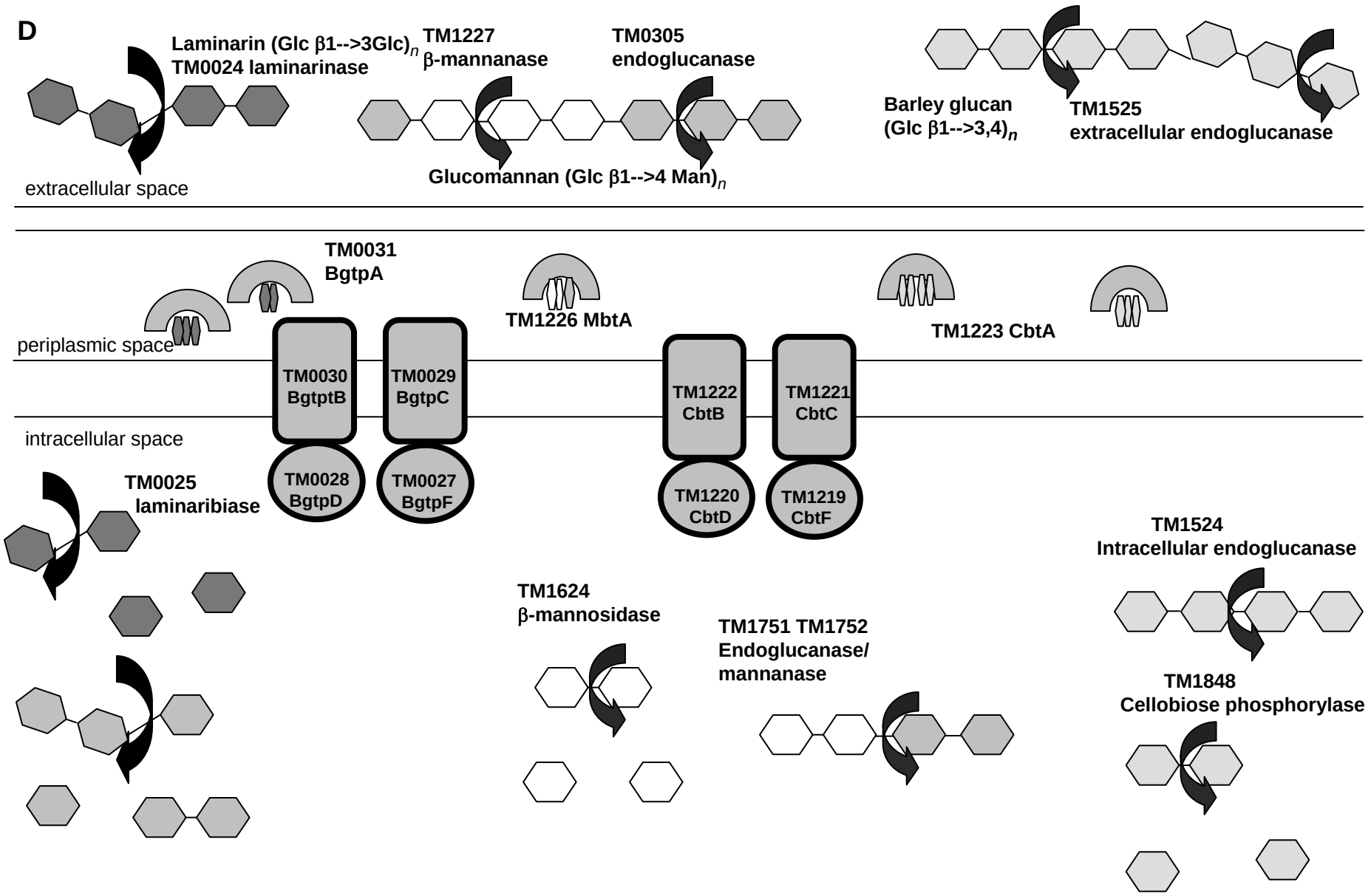

D
Laminarin (Glc β1-->3Glc)n
TM0024 laminarinase
TM1227
β-mannanase
TM0305
endoglucanase
Barley glucan
(Glc β1-->3,4)n
TM1525
extracellular endoglucanase
extracellular space
Glucomannan (Glc β1-->4 Man)n
TM0031
BgtpA
TM1226 MbtA
TM1223 CbtA
periplasmic space
TM0030
BgtptB
TM0029
BgtpC
TM1222
CbtB
TM1221
CbtC
intracellular space
TM0028
BgtpD
TM0027
BgtpF
TM1220
CbtD
TM1219
CbtF
TM0025
laminaribiase
TM1624
β-mannosidase
TM1751 TM1752
Endoglucanase/
mannanase
TM1524
Intracellular endoglucanase
TM1848
Cellobiose phosphorylase

Figure 3.5D

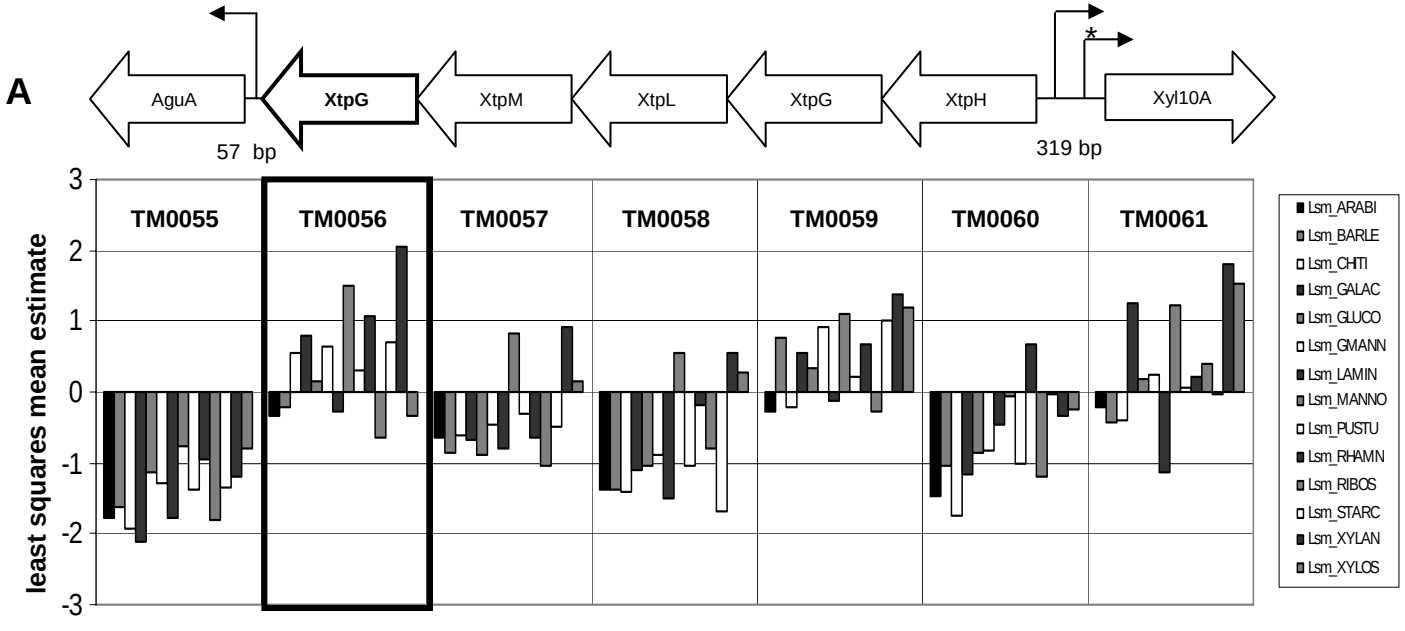

A
AguA
XtpG
XtpM
XtpL
XtpG
XtpH
Xyl10A
*
57 bp
319 bp
TM0055
TM0056
TM0057
TM0058
TM0059
TM0060
TM0061
least squares mean estimate
3
2
1
0
-1
-2
-3
Lsm_ARABI
Lsm_BARLE
Lsm_CHITI
Lsm_GALAC
Lsm_GLUCO
Lsm_GMANN
Lsm_LAMIN
Lsm_MANNO
Lsm_PUSTU
Lsm_RHAMN
Lsm_RIBOS
Lsm_STARC
Lsm_XYLAN
Lsm_XYLOS

Figure 3.6A

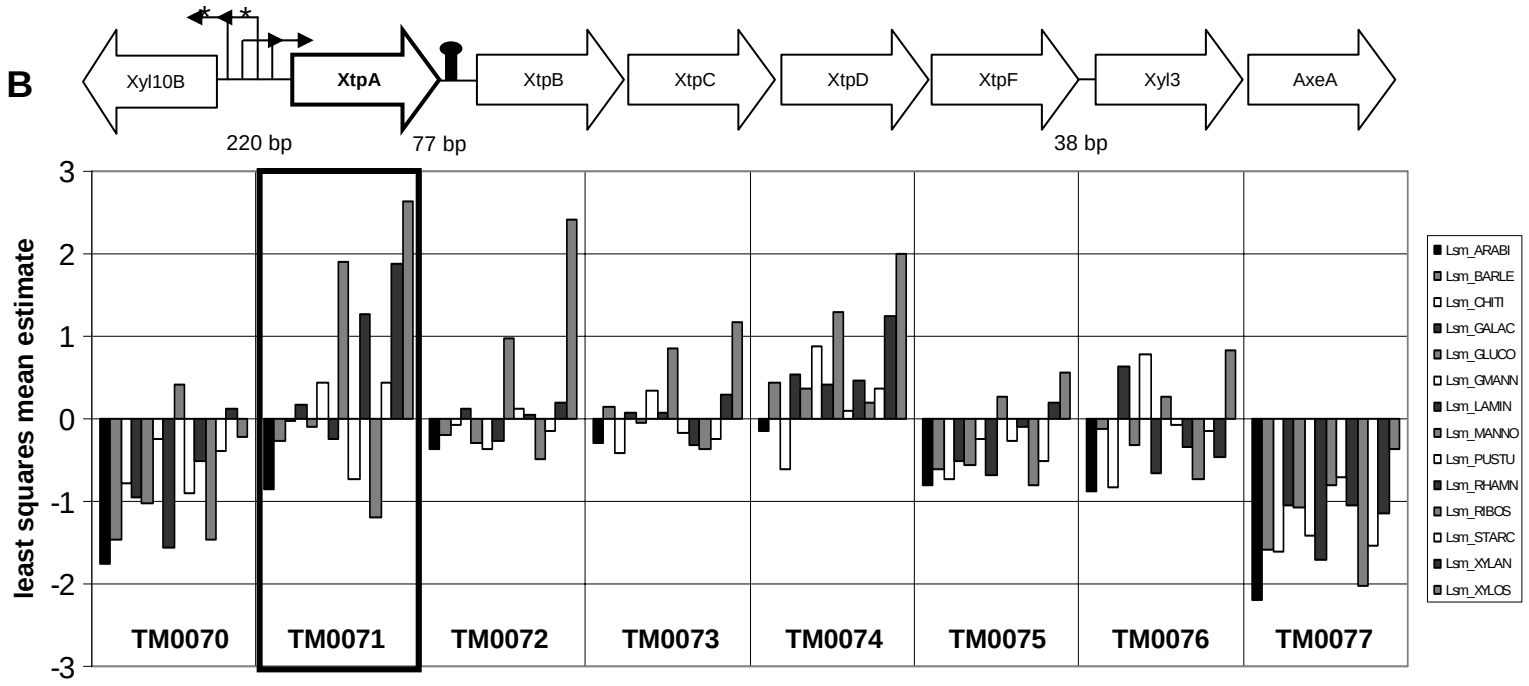

B
Xyl10B
XtpA
XtpB
XtpC
XtpD
XtpF
Xyl3
AxeA
220 bp
77 bp
38 bp
least squares mean estimate
3
2
1
0
-1
-2
-3
TM0070
TM0071
TM0072
TM0073
TM0074
TM0075
TM0076
TM0077
Lsm_ARABI
Lsm_BARLE
Lsm_CHITI
Lsm_GALAC
Lsm_GLUCO
Lsm_GMANN
Lsm_LAMIN
Lsm_MANNO
Lsm_PUSTU
Lsm_RHAMN
Lsm_RIBOS
Lsm_STARC
Lsm_XYLAN
Lsm_XYLOS

Figure 3.6B

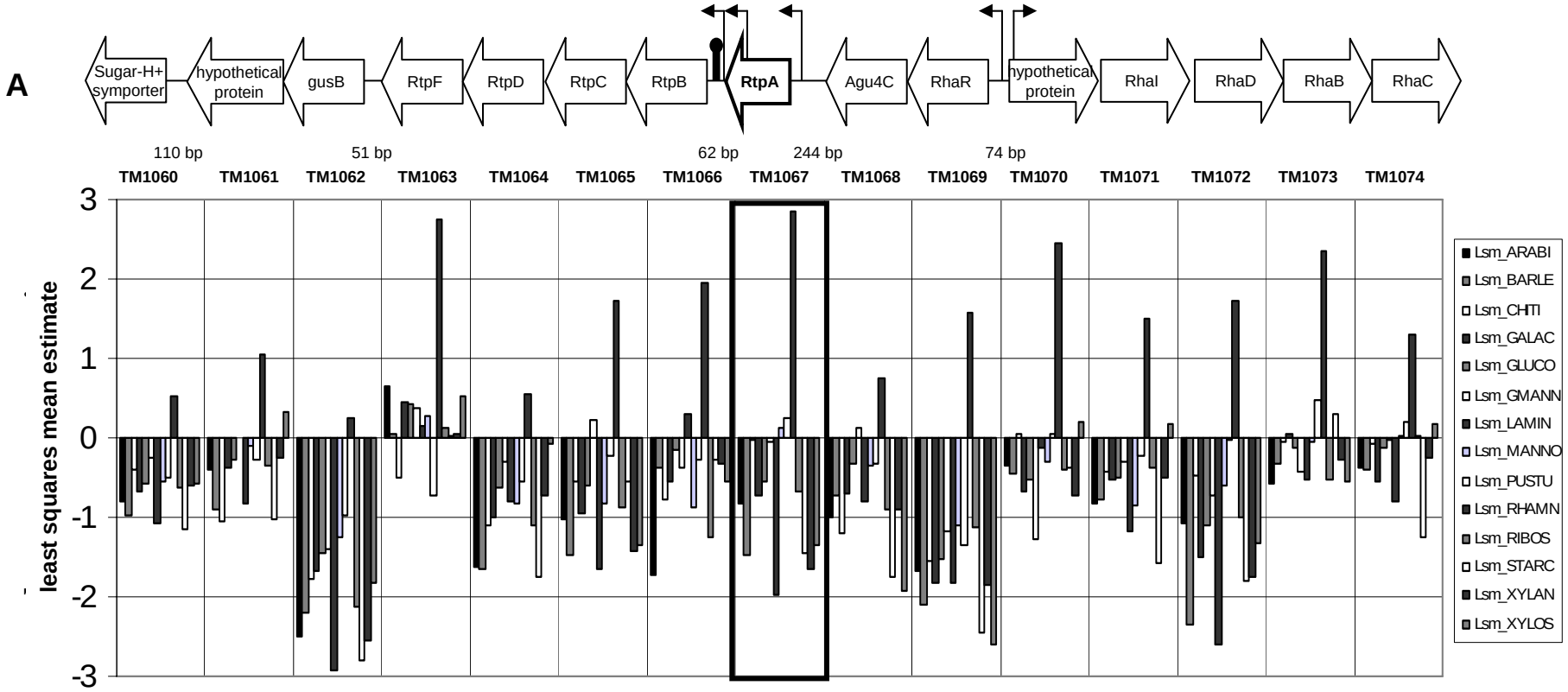

A
Sugar-H+ symporter
hypothetical protein
gusB
RtpF
RtpD
RtpC
RtpB
RtpA
Agu4C
RhaR
hypothetical protein
RhaI
RhaD
RhaB
RhaC
110 bp
51 bp
62 bp
244 bp
74 bp
TM1060
TM1061
TM1062
TM1063
TM1064
TM1065
TM1066
TM1067
TM1068
TM1069
TM1070
TM1071
TM1072
TM1073
TM1074
least squares mean estimate
3
2
1
0
-1
-2
-3
Lsm_ARABI
Lsm_BARLE
Lsm_CHITI
Lsm_GALAC
Lsm_GLUCO
Lsm_GMANN
Lsm_LAMIN
Lsm_MANNO
Lsm_PUSTU
Lsm_RHAMN
Lsm_RIBOS
Lsm_STARC
Lsm_XYLAN
Lsm_XYLOS

Figure 3.7A

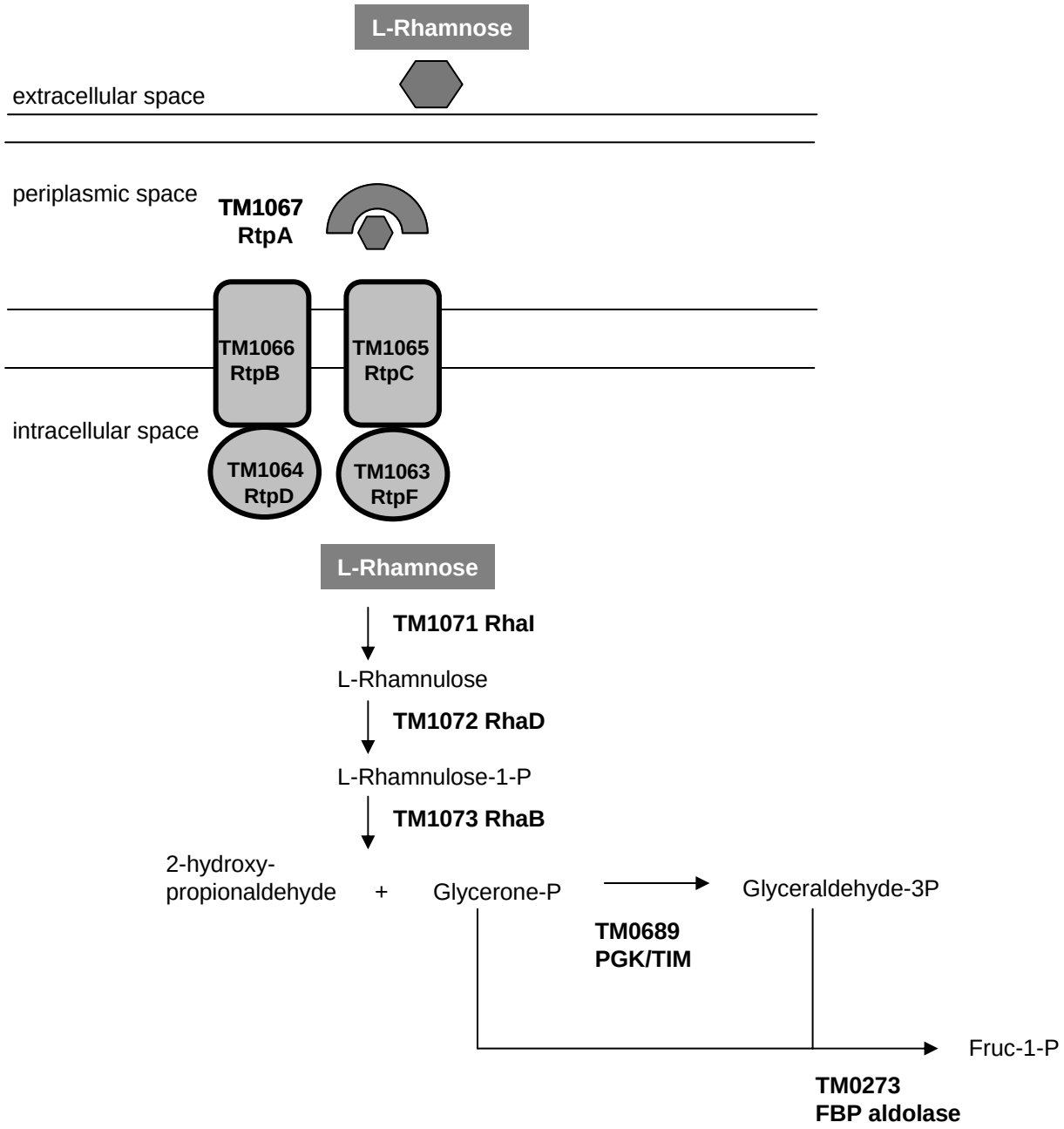

B
L-Rhamnose
extracellular space
periplasmic space
TM1067
RtpA
TM1066
RtpB
TM1065
RtpC
intracellular space
TM1064
RtpD
TM1063
RtpF
L-Rhamnose
TM1071 RhaI
L-Rhamnulose
TM1072 RhaD
L-Rhamnulose-1-P
TM1073 RhaB
2-hydroxy-
propionaldehyde
+
Glycerone-P
Glyceraldehyde-3P
TM0689
PGK/TIM
Fruc-1-P
TM0273
FBP aldolase

Figure 3.7B

Chapter 4:

Transcriptional profiles of cellobiose- and maltose-grown *T. maritima* cells suggest regulatory strategies optimized for ecologically relevant β-linked glucans

Shannon B. Conners, Steven R. Gray, Chung-Jung Chou, Matthew R. Johnson, Clemente I. Montero and Robert M. Kelly

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

ABSTRACT

Transcriptional responses of *Thermotoga maritima* cells to maltose and cellobiose were examined in batch and continuous culture conditions using a full genome microarray. Although minimal transcriptional changes were observed in response to sulfur addition, observed changes were highly consistent between maltose and cellobiose-grown cells. Transcription of genes previously implicated in import and utilization of maltose and cellobiose correlated reliably with growth conditions. Changes in transcripts of glycolytic genes under batch culture conditions may indicate response to differential hydrogen and metabolic by-product accumulation during maltose and cellobiose supplemented growth. Transcriptional data revealed up-regulation of KDG and glucuronate utilization genes during cellobiose growth, perhaps indicating a shift in the utilization patterns for 6-phosphogluconate. Increased transcription of a putative NADH:polysulfide oxidoreductase was observed during growth on cellobiose, correlating with reports of lower hydrogen evolution during cellobiose than maltose growth (55). β-linked exopolysaccharide (EPS) comprised mainly of glucose has been detected in *T. maritima* pure cultures and under co-culture conditions with *Methanococcus jannaschii* (21), and growth-phase dependent breakup of cellular communities held together by EPS is accompanied by expression of β-specific glycoside hydrolases (Johnson et. al, in press). Transcriptional regulation in response to cellobiose may indicate optimization of metabolic strategies for EPS formation and recycling in the natural environments of *T. maritima*.

INTRODUCTION

In the absence of a genetic system for the bacterium *Thermotoga maritima*, transcriptional data have been highly useful in revealing both physiological insights and potential mechanisms for ecological interaction (7, 9, 21, 34, 38, 39). In particular, transcriptional data have complemented predictions of existing genomic pathways for the utilization of simple sugars (e.g., xylose, ribose, rhamnose, arabinose) and complex carbohydrates (e.g., glucans and mannans of various linkage types, xylan) (7, 9). By coupling genomic and transcriptional data, candidate transporters and local regulators for β-1,4 mannan, cellobiose/β-1,4 glucan, β-1,3/1,6 glucan, rhamnose, arabinose, and ribose transport and utilization genes have been proposed (7, 9). These genes co-localize with hydrolytic and catabolic enzymes necessary to process the target sugars into forms which can be assimilated into central metabolism. Clearly, *T. maritima* has acquired or evolved pathways for the catabolism of a variety of carbohydrates, suggesting that varied carbohydrate linkages and compositions are present in its natural environments. One potential source of these varied carbohydrates is exopolysaccharide (EPS) from biofilms. Biofilm formation has been noted under continuous mono-culture and batch co-culture of *T. maritima* with *M. jannaschii* (21, 38). In high density maltose-grown mixed batch cultures, growth-inhibiting hydrogen produced by *T. maritima* is used by *M. jannaschii* to make methane (31), and hydrogen transfer is facilitated by cellular aggregations held together by exopolysaccharide (EPS) comprised of 91.2% glucose, 5.2% ribose, 1.7% mannose. Calcofluor staining suggests that β-linked glycans are present both in the presence and absence of *M. jannaschii*, and innoculation of a specific small peptide (TM0504) has been shown to induce EPS formation (21). Production of EPS appears to

be growth-phase dependent: aggregates form during mid-log phase and are degraded during early stationary phase. This may relate to the up-regulation of β-specific hydrolases during early stationary phase, as enzymatic degradation of biofilms for purposes of colonization has been observed in other species.

A mechanism of global catabolite repression has not been identified in *T. maritima*, which like other *Thermotoga* species examined thus far lacks the components of a phosphoenolpyruvate:sugar phospho-transferase system (PTS) system for the transport and phosphorylation of sugars and sugar-derived alcohols. Although the *T. maritima* genome encodes homologs to the *B. subtilis* CcpA and *E. coli* Crp proteins, a role in global catabolite repression has not been demonstrated for any of these proteins. A role has been proposed for a LacI/CcpA homolog (TM1218) as a regulator of cellobiose utilization genes. However, significant transcriptional differences exist between sugars which extend beyond processing and transport systems to genes of central metabolism, perhaps suggesting global transcriptional regulation under certain conditions.

The main products of *T. maritima* metabolism include acetate, lactate, alanine, CO_2 and H_2 (18). Differences in the ratio of H_2:CO_2 products have been observed when *T. maritima* cells are grown on various carbon sources (55). During growth on glucose, this ratio approaches 2:1; however, *T. maritima* grown on the β-1,4 linked glucan disaccharide cellobiose (1:2) and the α-1,4 glucan linked disaccharide maltose (4:3) release proportionally less hydrogen and more carbon dioxide (55). Here, transcriptional profiles of cells grown on maltose and cellobiose were compared during batch and chemostat growth. We have observed substantially different transcriptional responses of the archaeon *Pyrococcus furiosus* to cellobiose and maltose (Chou and Kelly, manuscript

in preparation). Since a large number of genes have been observed to be affected when the hyperthermophilic archaeon *P. furiosus* is grown on media supplemented with sulfur (44), the effect of sulfur on cellobiose and maltose batch cultures was also examined.

MATERIALS AND METHODS

Two methods were utilized to simulate possible growth conditions which might be observed in the natural environments of *T. maritima*. Continuous culture conditions include constant replenishment of nutrients through fresh media and the removal of potentially inhibitory metabolic by-products, while these metabolic products accumulate during batch growth. Cultures of *T. maritima* MSB8 were grown in artificial sea water using optical density measurements and epifluorescence microscopic cell density enumeration, as described previously (7, 9). Growth curves of *T. maritima* grown in batch culture in the presence of the disaccharides maltose and cellobiose, which differ only by linkage type, were similar and reproducible. Addition of sulfur did not affect growth rate or final cell density for either carbohydrate. All samples were taken during mid-exponential phase at a similar cell density. A loop design was constructed to compare the four batch conditions (Figure 4.1a). Continuous culture of *T. maritima* was performed using methods described in (38). Sampling was performed as follows: during growth on each sugar, an initial sample was taken (e.g., M1 or C1). Media was changed to a mixture of the two carbohydrates and a sample was taken once a steady state had been achieved (e.g., MC or CM with the first letter denoting which sugar was in the media at the start of the experiment). Finally, a third sample was taken upon re-establishment of the steady growth state on media containing the original carbohydrate

(e.g., M2 or C2). A loop design was constructed to compare all continuous culture samples (Figure 4.1b). Modified TIGR microarray protocols were used as described previously. A two-stage mixed model analysis was performed using normalization and gene models including both random effects (array, spot), fixed effects (dye, treatment), and random error (9, 54) using SAS (SAS Institute). For the batch carbohydrate-sulfur experiment, models including interactions between carbohydrate and sulfur was considered, but effects of sulfur and carbohydrate were additive in all cases. Therefore, a simpler model which treated each culture (e.g., cellobiose-sulfur, cellobiose-without sulfur, maltose-sulfur, maltose-without sulfur) independently was used. Least-squares mean estimates of treatment effects were obtained, and the statistical significance of pair-wise differences (analogous to $\log_2$-transformed fold changes) determined using proc mixed in SAS. Results were visualized in JMP (SAS Institute) and Excel using the program AFM 4.0 (4).

RESULTS AND DISCUSSION

Transcriptional effect of sulfur. Compared to the effect of carbohydrate linkage type, sulfur presence or absence had relatively little effect on transcriptional profiles for *T. maritima* grown on either maltose or cellobiose. This finding contrasts with results obtained by Schut and co-workers with the hyperthermophilic archaeon *P. furiosus*, in which substantial numbers of transcripts responded strongly to the presence or absence of sulfur (44). Despite the limited transcriptional response we observed, there was excellent agreement between the lists and magnitudes of genes which were differentially expressed in the presence or absence of sulfur during growth on maltose and cellobiose (Table 4.1).

Interestingly, in almost all cases, transcripts of genes which were repressed on sulfur were detected at higher levels during growth on maltose than cellobiose. Also, genes up-regulated during growth on sulfur tended to be up-regulated on cellobiose. Significant sulfur and carbohydrate effects appeared to be additive in all cases.

Among transcripts consistently detected at higher levels in the absence of sulfur were TM0980-2; transcripts of these genes and an associated gene (TM0979) were also detected at higher levels during growth on cellobiose. TM0979, TM0980, and TM0981 are related at the sequence level and homologs (e.g., *dsrEFH*) often co-localize in genomes. The crystal structure of TM0979 was recently determined by two groups and found to form a YchN-like fold (16, 36). DsrEFH are involved in the oxidation of intracellular sulfur in the phototrophic sulfur bacterium *Chromatium vinosum* D (now renamed *Allochromatium vinosum*), and transcription of these genes is up-regulated by the presence of reduced sulfur compounds (37). It has been proposed that these proteins may play a role in assembly or stabilization of sirohaem proteins in some species (37); however, they are also found in genomes of species which lack these proteins. The conserved cysteine and glycines of the YeeE/YedE membrane protein family are found within TM0982, a protein with no known function. TM0983 is a small protein (79 aa) which contains a SirA domain found in regulators of disulfide bond formation. TM0978, located downstream of the other orfs, has a HTH motif in addition to a SirA domain. The exact gene content and order of the TM0979-TM0983 locus of *T. maritima* is conserved in *Thermoanaerobacter tengcongensis* (1). Although the cellular roles of TM0978-TM0983 in *T. maritima* remain to be determined, sulfur-dependent transcriptional regulation of some of these genes may imply involvement in a sulfur-related cellular

process; however, TM0979 does not bind SO_4^{2-}, SO_3^{2-} , Na+, Ca^{2+}, Mg^{2+}, glucose, or ribose, among other potential ligands tested (16).

Other transcripts which declined in the presence of sulfur included a putative transcriptional regulator of the *tetR* family (TM0823) and a predicted iron-sulfur protein of the SAM superfamily (TM0824), which are separated by less than 20 nucleotides. A histone-like DNA-binding protein (TM0266) and an outer membrane protein (TM0477) are encoded by two other sulfur-repressed orfs. Transcripts of spermidine synthase (TM0654) and a single ribosomal protein (TM1485) were also affected by sulfur absence.

Transcript levels of a select few genes were consistently higher in the presence of sulfur, including a gene (TM1292) related to *sipB*, a sulfur induced putative polyferredoxin from *P. furiosus* (44) In *P. furiosus*, *sipA* and *sipB* are divergently transcribed, although their homologs in two other sequenced *Pyrococcus* species (*abyssi* and *horokoshii*) are distantly separated. The related *T. maritima* genes appear to comprise an operon-like structure (TM1290-TM1291-TM1292), although only TM1292 responds significantly to the presence of sulfur. Transcripts of a *ham1* gene (TM0159) homologous to XTP pyrophosphatases involved in purine salvage was also detected at higher levels during growth on sulfur, and TM0375 is a putative propanediol utilization gene whose function remains unclear. Dihydrofolate reductase (TM1641), which uses NADPH to reduce 7,8-dihydrofolate, had higher transcript levels in the presence of both sulfur and maltose.

The limited transcriptional response of *T. maritima* to sulfur presence or absence was somewhat surprising, given the sulfur-dependent expression changes of a large number of genes in *P. furiosus*. In particular, strong down-regulation of a large number of

hydrogenase subunits was observed during sulfur-dependent growth for *P. furiosus* (44). The *T. maritima* genome (33) contains a number of operons with genes similar to hydrogenase components, including genes encoding a characterized cytoplasmic αβγ NADH-dependent iron-only hydrogenase (TM1424-6) (22, 51, 52), an uncharacterized putative NADP-reducing hydrogenase (TM0010-2), and uncharacterized homologs to archaeal MbxA-MbxN hydrogenase subunits (TM1205-16) with an associated GltD-related gene proposed to be a bifunctional glutamate synthase and hydrogenase (6). The Mbx proteins constitute an uncharacterized putative Ni-Fe hydrogenase in *P. furiosis*, although the related *mbh* operon has been characterized and found to encode subunits of a membrane-bound complex with similarity to respiratory chains (43, 48). Although membrane-bound hydrogenase activity distinct from the cytoplasmic iron-only hydrogenase has been reported for *T. neapolitana*, the identity of the protein(s) involved has not been reported (23). Of the known and putative hydrogenases of *T. maritima*, only the NADH dependent iron-only hydrogenase (TM1424-6) and associated genes (TM1420-3, TM1427-8) displayed a consistent transcriptional response, but to cellobiose rather than sulfur.

Lack of a transcriptional response of hydrogenase genes to sulfur presence or absence may relate to findings that the activity of the characterized *T. maritima* iron-only hydrogenase is similar in the presence and absence of sulfur and sulfide (22). It has been proposed that the iron-only hydrogenase found in *Thermotoga* sp. plays a main role in hydrogen evolution rather than sulfur reduction (8). Rather, a NADH:polysulfide oxidoreductase (TM0379) may play a primary role in sulfur reduction, as the activity of a close *T. neapolitana* homolog increases in the presence of sulfur and has much higher

sulfur reduction activity than the hydrogenase (8). If activity of this enzyme is directly regulated by sulfur, this may explain the lack of transcriptional response of the corresponding gene to its presence or absence. In contrast, TM0379 is up-regulated nearly 8-fold during growth on cellobiose as compared to maltose, perhaps indicating a metabolic adjustment triggered by sugar type.

Transcriptional effect of carbohydrate linkage type. The sugar linkage type of the glucan disaccharide was clearly important in distinguishing the transcriptional response to each carbohydrate. During both batch and continuous culture conditions, specific differential transcripts were consistently observed for maltose (α-1,4-linked) and cellobiose (β-1,4 linked) which related to sugar transport and processing. However, a number of very consistent differences observed between cellobiose/maltose and cellobiose-sulfur/maltose-sulfur cells in batch culture were not observed during continuous culture. This effect likely results from the removal of metabolic byproducts during growth in continuous culture growth, and highlights the advantage of using both types of growth experiments.

Cellobiose response regulon. *T. maritima* cells showed characteristic transcriptional responses to cellobiose during both batch growth and continuous culture, consistent with previous observations during growth on barley glucan and carboxymethyl cellulose (batch) (7, 9), and cellobiose (continuous culture, Montero and Kelly, unpublished observations). A set of likely cellobiose ABC transporter genes (CbtABCDF) responded strongly under cellobiose-only and cellobiose/maltose mixture conditions. The gene

encoding CbtA (TM1223), a putative substrate binding protein for cellobiose and β-1,4-glucans which is closely related to *P. furiosus* CbtA (26), was most highly up-regulated (22-55 fold for different comparisons with maltose-only conditions). Transcripts of neighboring genes encoding ABC permease and ATP-binding subunits also increased when cellobiose was present (Table 4.2, Figure 4.2). Transcripts of the gene encoding the characterized *T. maritima* cellobiose phosphorylase (41) and two endoglucanases (CelA and CelB, TM1524 and TM1525) (27) showed similar preferences for cellobiose over maltose. While cellobiose phosphorylase is likely the primary cellobiose-processing enzyme in *T. maritima*, increased transcription of *celA* and *celB* may result from a general transcriptional response to β-1,4-glucans. We have proposed that this response is controlled by the putative LacI family repressor encoded by orf TM1218, located downstream of the cellobiose transporter. Transcripts of TM1218 increased approximately 9 fold when cellobiose was added to a maltose-only culture, but did not change significantly when maltose was added to a cellobiose-only culture. We have previously reported a motif upstream of the cellobiose response regulon which is highly conserved and matches the LacI family consensus (7, 9). It is of note that the changes in cellobiose utilization genes described here indicate a specific transcriptional response to β-1,4 glucans. Genes within loci previously noted to respond to β-glucans of different linkage types (β-1,3 laminarin and β-1,6 pustulan, TM0024-32), xylose and the β-1,4-linked xylose polymer xylan (TM0056-61, TM0070-5), the monosaccharides rhamnose (TM1061-74), and ribose, arabinose, and xylose (TM0949-60), did not show significant differences between the two disaccharides examined here.

Maltose response regulon. Two maltose transporters have been characterized in *T. maritima* (32, 34, 53). The binding proteins of these transporters have been observed to have different expression patterns but overlapping binding specificities, as both bind maltose and maltotriose, but TM1839 binds trehalose and is upregulated during growth on maltose, while TM1204 binds β-mannotetraose and responds transcriptionally to lactose (32, 34). Consistent with previous work, transcripts of TM1204 were detected at similar levels here during growth on maltose and cellobiose. Transcripts from TM1839 are usually detected at high levels under many conditions, suggesting either constitutive expression of these genes or very high transcript levels, which may impair accurate detection of true changes in transcript levels by microarray. However, transcripts of TM1839 clearly responded to the presence or absence of maltose under continuous culture conditions. The largest increase (2.8 fold) was observed upon addition of maltose to a cellobiose-only culture. Other transcripts relating to the hydrolysis and transport of alpha glucans (e.g. TM0766, TM1650, TM1834, and TM1841) displayed smaller changes suggestive of increased transcription. The only confirmed maltose hydrolase identified in the *T. maritima* MSB8 genome, the NAD+-dependent alpha glucosidase AglA (TM1834) (2), displayed a maximal fold change of 1.9 among all comparisons of maltose and cellobiose conditions. Despite the lack of apparent transcriptional regulators in the main maltose utilization operons, transcripts from a putative sugar fermentation stimulation protein (*sfsA*, TM0565) were detected at higher levels during growth on maltose. A SfsA homolog in *E. coli* has been implicated in the co-regulation of maltose utilization genes in cooperation with CRP (24, 49). Clearly, further characterization of *T.*

maritima SfsA will be necessary to determine its involvement in maltose regulatory processes.

Differences in transcriptional profiles extend beyond transporters and hydrolases. Differences in transcriptional patterns of transporters and disaccharide utilization genes between maltose- and cellobiose-grown cells can be easily related to sugar type, but differential transcription patterns extended beyond these genes. On the whole, multiple factors are likely to be responsible for the additional differences observed. Both cellobiose and maltose are taken into *T. maritima* cells via ATP-dependent ABC transporters, but the catabolic steps necessary for each disaccharide have differing ATP requirements. The α-glucosidase (TM1834) (28) produces two molecules of α-D-glucose per maltose unit, each of which must be phosphorylated by the ATP-dependent glucokinase (TM1469) (17). In contrast, cellobiose phosphorylase produces one molecule of α-D-glucose and one molecule of α-D-glucose-1-phosphate during the phosphorylysis of cellobiose, using free phosphate and conserving ATP in the process (57). In addition, cellular redox levels and NAD/NADH and NADP/NADPH ratios may affect transcriptional patterns and cellular activity of enzymes. The *T. maritima* alpha glucosidase (TM1834) (2) is NAD+-dependent (40), while two key glucose catabolic enzymes (glucose-6-phosphate dehydrogenase, 6-phosphogluconate dehydrogenase) are NADP+-dependent. Similar observations of changes in redox proteins have been made during growth of *T. maritima* on maltose and lactose as compared to glucose (34), further indicating this may be part of a global response.

During growth on glucose, *T. maritima* has been shown to metabolize approximately 85% of glucose by a conventional Embden-Meyerhof-Parnas pathway, and approximately 15% of glucose via a classical phophorylated Entner-Doudroff pathway (47). This results in the production of proportionally more hydrogen than carbon dioxide in an approximate ratio of 2:1, a fact which has been used to construct a cell-free hydrogen production system using glucose and *T. maritima* enzymes with *P. furiosus* NADP+-dependent hydrogenase (55). Although the relative proportions of glucose metabolized by the EMP and ED pathways have not been determined for *T. maritima* during growth on cellobiose and maltose, proportionally more H_2 than CO_2 is produced during growth on maltose (4:3) than cellobiose (1:2) (55). The characterized *T. maritima* hydrogenase is believed to use NADH as an electron donor for hydrogen evolution, since reduced ferredoxin generated by POR was not used as electron donor for the hydrogenase; however, the fact that the *T. maritima* hydrogenase also catalyzed reduction of a quinone perhaps hints at a more complex electron transfer strategy POR and the hydrogenase (51).

One possible explanation for differences in hydrogen production ratios between maltose and cellobiose is suggested by the strong up-regulation of a putative operon (4-16x) during cellobiose growth conditions (TM0379-81). A homolog of the characterized *T. neapolitana* NADH:polysulfide oxidoreductase (TM0379, 86% identical/443 residues) encoded within this operon may compete with the hydrogenase for NADH during growth on cellobiose (8), leading to polysulfide or sulfur reduction as an alternative to hydrogen production. Although biochemical verification of increased sulfur reduction during cellobiose growth will be necessary to confirm this hypothesis, indirect supporting

evidence is available in the form of increased transcription of the *T. maritima* homologs of *dsrEF* (TM0980-1), genes possibly involved in the oxidation of intracellular sulfur, during cellobiose- and sulfur-dependent growth (37). The physiological functions of two genes co-localized with TM0379 have not yet been determined. While TM0380 contains a genuine frameshift in the MSB8 genome sequence, truncating the size of the orf at 69 residues, the remaining protein includes a well-conserved domain with two conserved cysteines. Similarly strong hybridization signals are detected from the truncated orf and the third gene, suggesting that the truncated protein may have some functional relevance to *T. maritima* cells. The third orf, TM0381, is annotated as dihydrolipoamide dehydrogenase and shows similarity to NAD- and FAD-dependent dehydrogenases. Although dihydrolipoamide dehydrogenase plays the role of the E3 component of the pyruvate dehydrogenase complex in other organisms, the *T. maritima* genome lacks other identifiable subunits of the complex (e.g., E1, pyruvate dehydrogenase, decarboxylase component; and E2, dihydrolipoamide acyltransferase). This is consistent with observations that, like other hyperthermophiles, *T. maritima* uses a POR to accomplish pyruvate oxidation rather than pyruvate dehydrogenase (3, 25).

It is interesting that most genes in the operon encoding *T. maritima* iron-only hydrogenase subunits (TM1420-7), while not regulated by sulfur, are apparently differentially regulated between maltose and cellobiose in batch culture. This may relate to changing NAD+/NADH ratios triggered by transcriptional changes in glycolytic enzymes (Figure 4.3), if these correspond to changes in protein levels and activity in *T. maritima* cells. In particular, transcripts of the NAD+-dependent glyceraldehyde-3-phosphate dehydrogenase (TM0688), which utilizes free phosphate and generates

NADH, are higher during cellobiose growth. Electrons generated by this reaction have been found to preferentially reduce sulfur to sulfide in *Thermotoga* sp. FjSS3.B1 rather than contributing to hydrogen production (19). Despite the observed transcriptional changes of the hydrogenase subunits, the much greater enzyme activity of polysulfide oxidoreductase may reduce NADH utilization for hydrogen production (8). Clues to the mechanism by which the hydrogenase genes respond to redox cofactor ratios are provided by additional proteins encoded in the hydrogenase operon, including a putative transcriptional regulator (TM1427) related to Rex, a redox sensitive transcriptional regulator of *Streptomyces coelicolor*, which dissociates from its operator upon binding NADH (5). Although TM1427 lacks the conserved glycine residues implicated in NADH binding in related gram-positive proteins, its localization immediately downstream of the genes encoding subunits of the *T. maritima* NADH-dependent hydrogenase and similar expression patterns suggest a possible conserved role in sensing of NAD+/NADH ratios (33) (Figure 4.4). Further supporting this idea, we have detected a match to the *Streptomyces coelicolor* Rex repressor operator sequence upstream of TM1420 (35), the first gene in the hydrogenase operon. This sequence, TGTTACCTACTTCACA, is highly similar to the putative Rex-binding sequence TGTGACCTGCTTCACA found upstream of the *nuoA-N* operon encoding subunits of an NADH dehydrogenase in *Streptomyces coelicolor* A2(3) (5).

Beyond glycolysis, genes relating to pyruvate utilization were also up-regulated during cellobiose batch growth, including the ATP-dependent pyruvate kinase (TM0208) and subunits of the pyruvate ferredoxin oxidoreductase (POR, TM0015-18). If these transcriptional changes correlate to increased activity of these enzymes, increased carbon

dioxide released during pyruvate oxidation by POR may contribute to relatively greater amounts of carbon dioxide produced during growth on cellobiose. It is of note that differences in transcript levels for the POR subunits have also been observed during growth on arabinose, ribose, and mannose as compared to glucose in batch culture (Conners and Kelly, unpublished observation), consistent with lower H_2 and higher CO_2 production during growth on these carbohydrates. Differential regulation of acetate kinase and lactate dehydrogenase was also apparent, with transcripts of the acetate kinase (TM0274) higher and transcripts of the lactate dehydrogenase lower during growth on cellobiose. This may relate to the relative amounts of hydrogen produced during growth on maltose and cellobiose, as increased lactate production has been noted under high hydrogen partial pressures for *Thermotoga* sp. FjSS3.B1 (19). It must be noted that many of the changes observed in genes of central metabolism and glycolysis were apparent in batch grown cells but not cells grown in a chemostat, most likely as the result of differences in metabolic product accumulation between batch growth and continuous culture growth.

In addition to genes encoding proteins responsible for formation of metabolic end-products, other cellular pathway members were significantly up-regulated during cellobiose batch growth. These included changes ranging from 2-3.5 fold in a number of genes encoding enzymes of the glucuronate and pentose interconversion pathway, e.g, *uxaC* (TM0064, uronate isomerase), *kdgA* (TM0066, 2-dehydro-3-deoxyphospho-gluconate aldolase), *kdgK* (TM0067, 2-keto-3-deoxygluconate kinase), *uxuB* (TM0068, D-mannonate oxidoreductase), *uxuA* (TM0069, D-mannonate hydrolase), and a putative regulatory protein likely to play the role of *kdgR* (TM0065) (58). A possible reason could

be increased activity of the ED pathway enzyme 6-phosphogluconate dehydratase (TM0551, 31%/520 aa with the *E. coli* homolog), which has homology to enzymes which produce KDPG from 6-phosphogluconate. KDPG is an intermediate in the pentose and glucuronate interconversion pathway, and is dephosphoryated by KDG kinase to yield KDG, a likely inducer for the regulon. The 6-phosphogluconate dehydratase might be preferentially used during growth on cellobiose as an alternative to the NADP+ dependent 6-phosphogluconate dehydrogenase (TM0438), which decarboxylates 6-phosphogluconate to release ribulose-5-phosphate. It is of note that transcripts of the *T. maritima* ribokinase are detected at higher levels on maltose. Similarly to TM1155, neither TM0551 nor TM0438 appear to be regulated strongly at the transcriptional level, although their activity could be regulated by cellular redox cofactor ratios.

The EMP pathway glycolytic genes regulated between cellobiose and maltose are the same genes which showed significant regulation between *T. maritima* pure culture and high density co-culture (TM0208-9, TM0273, TM0688-9, TM0877, TM1469), although the latter comparison resulted in larger fold changes (21). However, the regulatory mechanism for these glycolytic genes remains unknown. *T. maritima* lacks an ortholog to *B. subtilis* CggR, a repressor of glycolysis genes (29). Although a glucokinase has been implicated in glucose-mediated catabolite repression in some *Streptomyces* and *Staphylococcus* species, the *T. maritima* glucokinase lacks a DNA-binding HTH motif. Work in *B. subtilis* has demonstrated a role for CcpA in glucose-mediated up-regulation of glycolysis (29). Matches to the CcpA consensus are apparent upstream of a number of genes in the *T. maritima* genome. Some matches can be predicted to be involved in local regulation of specific operons based on genomic context, but others are found far from

LacI homologs and carbohydrate transport or hydrolysis genes, perhaps suggesting a broader role for a LacI family protein or a hierarchy of LacI family proteins in global transcriptional regulation in response to sugars.

Common transcriptional features of biofilm cells and cellobiose-grown cells. Gene expression differences between maltose and cellobiose may indicate adaptations to community growth of *T. martima* in ecological settings where β-linked EPS is present. EPS formation has been observed in pure and co-cultures of *T. maritima* (21), and up-regulation of β-specific glycoside hydrolases occurs during growth-phase dependent degradation of EPS (Johnson et al, in press). A number of genes and operons in Table 4.2 are marked with a subscript B to indicate commonalities between transcriptional profiles of *T. maritima* biofilm cells and cellobiose-grown cells. These include in common several upregulated gene sets: iron transporter components FeoAB (TM0049-50), heat shock chaperones (TM0373-4), archaeal type multi-subunit glutamate synthase genes (TM0394-8), the putative NADH:polysulfide oxidoreductase (TM0379) and cellobiose phosphorylase (TM1848), which was found to be strongly differentially regulated between biofilm and planktonic cells, although growth was on maltose in both cases (38). Two genes were down-regulated both in the biofilm cells and cellobiose-grown cells (TM0560, putative bacterioferritin, and TM0659, rubredoxin). Perhaps these common features relate to the upregulation of a *sigA* sigma factor homolog (TM1451) during both biofilm growth and cellobiose growth, or relate to similar redox environments in the cells. It may be advantageous for pure *T. maritima* biofilm cells to prevent the buildup of growth inhibitory hydrogen by preferentially reducing sulfur, perhaps precipitating a shift in intracellular sulfur stores and/or redox cofactors, and causing the up-regulation of

components of a chaperone system for iron-sulfur proteins and transporters for iron and sulfur compounds (38). The details of biofilm formation in *T. maritima* are likely complex and may differ between pure and mixed cultures. Efforts to better understand the signaling and regulatory processes which control EPS formation and degradation are ongoing in our laboratory.

ACKNOWLEDGMENTS

This work was supported in part by grants from the National Science Foundation (Grant No. BES-0317886) and the Department of Energy, Energy Biosciences Program (Grant No. DE-FG02-96ER20219).

CHAPTER 4 REFERENCES

1. **Bao, Q., Y. Tian, W. Li, Z. Xu, Z. Xuan, S. Hu, W. Dong, J. Yang, Y. Chen, Y. Xue, Y. Xu, X. Lai, L. Huang, X. Dong, Y. Ma, L. Ling, H. Tan, R. Chen, J. Wang, J. Yu, and H. Yang.** 2002. A complete sequence of the *T. tengcongensis* genome. Genome Res **12:**689-700.
2. **Bibel, M., C. Brettl, U. Gosslar, G. Kriegshauser, and W. Liebl.** 1998. Isolation and analysis of genes for amylolytic enzymes of the hyperthermophilic bacterium *Thermotoga maritima*. FEMS Microbiol Lett **158:**9-15.
3. **Blamey, J. M., and M. W. Adams.** 1994. Characterization of an ancestral type of pyruvate ferredoxin oxidoreductase from the hyperthermophilic bacterium, *Thermotoga maritima*. Biochemistry **33:**1000-7.
4. **Breitkreutz, B. J., P. Jorgensen, A. Breitkreutz, and M. Tyers.** 2001. AFM 4.0: a toolbox for DNA microarray analysis. Genome Biol **2:**SOFTWARE0001.
5. **Brekasis, D., and M. S. Paget.** 2003. A novel sensor of NADH/NAD+ redox poise in *Streptomyces coelicolor* A3(2). Embo J **22:**4856-65.
6. **Calteau, A., M. Gouy, and G. Perriere.** 2005. Horizontal transfer of two operons coding for hydrogenases between bacteria and archaea. J Mol Evol **60:**557-65.
7. **Chhabra, S. R., K. R. Shockley, S. B. Conners, K. L. Scott, R. D. Wolfinger, and R. M. Kelly.** 2003. Carbohydrate-induced differential gene expression patterns in the hyperthermophilic bacterium *Thermotoga maritima*. J Biol Chem **278:**7540-52.

8. **Childers, S. E., and K. M. Noll.** 1994. Characterization and regulation of sulfur reductase activity in *Thermotoga neapolitana*. Appl Environ Microbiol **60:**2622-2626.

9. **Conners, S. B., C. I. Montero, D. A. Comfort, K. R. Shockley, M. R. Johnson, S. R. Chhabra, and R. M. Kelly.** 2005. An expression-driven approach to the prediction of carbohydrate transport and utilization regulons in the hyperthermophilic bacterium *Thermotoga maritima*. J Bacteriol **187:**7267-82.

10. **Dams, T., G. Auerbach, G. Bader, U. Jacob, T. Ploom, R. Huber, and R. Jaenicke.** 2000. The crystal structure of dihydrofolate reductase from *Thermotoga maritima*: molecular features of thermostability. J Mol Biol **297:**659-72.

11. **Dams, T., G. Bohm, G. Auerbach, G. Bader, H. Schurig, and R. Jaenicke.** 1998. Homo-dimeric recombinant dihydrofolate reductase from *Thermotoga maritima* shows extreme intrinsic stability. Biol Chem **379:**367-71.

12. **Dams, T., and R. Jaenicke.** 2001. Dihydrofolate reductase from *Thermotoga maritima*. Methods Enzymol **331:**305-17.

13. **Dams, T., and R. Jaenicke.** 1999. Stability and folding of dihydrofolate reductase from the hyperthermophilic bacterium *Thermotoga maritima*. Biochemistry **38:**9169-78.

14. **Ding, Y. R., R. S. Ronimus, and H. W. Morgan.** 2001. *Thermotoga maritima* phosphofructokinases: expression and characterization of two unique enzymes. J Bacteriol **183:**791-4.

15. **Engel, A. M., Z. Cejka, A. Lupas, F. Lottspeich, and W. Baumeister.** 1992. Isolation and cloning of Omp alpha, a coiled-coil protein spanning the periplasmic space of the ancestral eubacterium *Thermotoga maritima*. Embo J **11:**4369-78.

16. **Gaspar, J. A., C. Liu, K. A. Vassall, G. Meglei, R. Stephen, P. B. Stathopulos, A. Pineda-Lucena, B. Wu, A. Yee, C. H. Arrowsmith, and E. M. Meiering.** 2005. A novel member of the YchN-like fold: Solution structure of the hypothetical protein Tm0979 from *Thermotoga maritima*. Protein Sci **14:**216-23.

17. **Hansen, T., and P. Schonheit.** 2003. ATP-dependent glucokinase from the hyperthermophilic bacterium *Thermotoga maritima* represents an extremely thermophilic ROK glucokinase with high substrate specificity. FEMS Microbiol Lett **226:**405-11.

18. **Huber, R., T. A. Langworthy, H. Konig, M. Thomm, C. R. Woese, U. B. Sleytr, and K. O. Stetter.** 1986. *Thermotoga maritima* sp. nov. represents a new genus of unique extremely thermophilic eubacteria growing up to 90 DegreesC. Archives of Microbiology **144:**324-333.

19. **Janssen, P. H., and H. W. Morgan.** 1992. Heterotrophic sulfur reduction by *Thermotoga* sp. strain FjSS3.B1. FEMS Microbiol Lett **75:**213-7.

20. **Johnsen, U., T. Hansen, and P. Schonheit.** 2003. Comparative analysis of pyruvate kinases from the hyperthermophilic archaea *Archaeoglobus fulgidus*, *Aeropyrum pernix*, and *Pyrobaculum aerophilum* and the hyperthermophilic bacterium *Thermotoga maritima*: unusual regulatory properties in hyperthermophilic archaea. J Biol Chem **278:**25417-27.

21. **Johnson, M. R., C. I. Montero, S. B. Conners, K. R. Shockley, S. L. Bridger, and R. M. Kelly.** 2005. Population density-dependent regulation of exopolysaccharide formation in the hyperthermophilic bacterium *Thermotoga maritima*. Mol Microbiol **55:**664-74.

22. **Juszczak, A., S. Aono, and M. W. Adams.** 1991. The extremely thermophilic eubacterium, *Thermotoga maritima*, contains a novel iron-hydrogenase whose cellular activity is dependent upon tungsten. J Biol Chem **266:**13834-41.

23. **Kaslin, S. A., S. E. Childers, and K. M. Noll.** 1998. Membrane-associated redox activities in *Thermotoga neapolitana*. Arch Microbiol **170:**297-303.

24. **Kawamukai, M., R. Utsumi, K. Takeda, A. Higashi, H. Matsuda, Y. L. Choi, and T. Komano.** 1991. Nucleotide sequence and characterization of the *sfs1* gene: *sfs1* is involved in CRP*-dependent *mal* gene expression in *Escherichia coli*. J Bacteriol **173:**2644-8.

25. **Kletzin, A., and M. W. Adams.** 1996. Molecular and phylogenetic characterization of pyruvate and 2-ketoisovalerate ferredoxin oxidoreductases from *Pyrococcus furiosus* and pyruvate ferredoxin oxidoreductase from *Thermotoga maritima*. J Bacteriol **178:**248-57.

26. **Koning, S. M., M. G. Elferink, W. N. Konings, and A. J. Driessen.** 2001. Cellobiose uptake in the hyperthermophilic archaeon *Pyrococcus furiosus* is mediated by an inducible, high-affinity ABC transporter. J Bacteriol **183:**4979-84.

27. **Liebl, W., P. Ruile, K. Bronnenmeier, K. Riedel, F. Lottspeich, and I. Greif.** 1996. Analysis of a *Thermotoga maritima* DNA fragment encoding two similar

thermostable cellulases, CelA and CelB, and characterization of the recombinant enzymes. Microbiology **142 (Pt 9):**2533-42.

28. **Liebl, W., B. Wagner, and J. Schellhase.** 1998. Properties of an alpha-galactosidase, and structure of its gene *galA*, within an alpha-and beta-galactoside utilization gene cluster of the hyperthermophilic bacterium *Thermotoga maritima*. Syst Appl Microbiol **21:**1-11.

29. **Ludwig, H., G. Homuth, M. Schmalisch, F. M. Dyka, M. Hecker, and J. Stulke.** 2001. Transcription of glycolytic genes and operons in *Bacillus subtilis*: evidence for the presence of multiple levels of control of the *gapA* operon. Mol Microbiol **41:**409-22.

30. **Lupas, A., S. Muller, K. Goldie, A. M. Engel, A. Engel, and W. Baumeister.** 1995. Model structure of the Omp alpha rod, a parallel four-stranded coiled coil from the hyperthermophilic eubacterium *Thermotoga maritima*. J Mol Biol **248:**180-9.

31. **Muralidharan, V., K. D. Rinker, I. S. Hirsh, E. J. Bouwer, and R. M. Kelly.** 1997. Hydrogen transfer between methanogens and fermentative heterotrophs in hyperthermophilic cocultures. Biotechnology and Bioengineering **56:**268-278.

32. **Nanavati, D. M., T. N. Nguyen, and K. M. Noll.** 2005. Substrate specificities and expression patterns reflect the evolutionary divergence of maltose ABC transporters in *Thermotoga maritima*. J Bacteriol **187:**2002-9.

33. **Nelson, K. E., R. A. Clayton, S. R. Gill, M. L. Gwinn, R. J. Dodson, D. H. Haft, E. K. Hickey, J. D. Peterson, W. C. Nelson, K. A. Ketchum, L. McDonald, T. R. Utterback, J. A. Malek, K. D. Linher, M. M. Garrett, A. M.**

Stewart, M. D. Cotton, M. S. Pratt, C. A. Phillips, D. Richardson, J. Heidelberg, G. G. Sutton, R. D. Fleischmann, J. A. Eisen, C. M. Fraser, and et al. 1999. Evidence for lateral gene transfer between Archaea and bacteria from genome sequence of *Thermotoga maritima*. Nature **399:**323-9.

34. **Nguyen, T. N., A. D. Ejaz, M. A. Brancieri, A. M. Mikula, K. E. Nelson, S. R. Gill, and K. M. Noll.** 2004. Whole-genome expression profiling of *Thermotoga maritima* in response to growth on sugars in a chemostat. J Bacteriol **186:**4824-8.

35. **Pan, G., A. L. Menon, and M. W. Adams.** 2003. Characterization of a [2Fe-2S] protein encoded in the iron-hydrogenase operon of *Thermotoga maritima*. J Biol Inorg Chem **8:**469-74.

36. **Peti, W., T. Etezady-Esfarjani, T. Herrmann, H. E. Klock, S. A. Lesley, and K. Wuthrich.** 2004. NMR for structural proteomics of *Thermotoga maritima*: screening and structure determination. J Struct Funct Genomics **5:**205-15.

37. **Pott, A. S., and C. Dahl.** 1998. Sirohaem sulfite reductase and other proteins encoded by genes at the *dsr* locus of *Chromatium vinosum* are involved in the oxidation of intracellular sulfur. Microbiology **144 (Pt 7):**1881-94.

38. **Pysz, M. A., S. B. Conners, C. I. Montero, K. R. Shockley, M. R. Johnson, D. E. Ward, and R. M. Kelly.** 2004. Transcriptional analysis of biofilm formation processes in the anaerobic, hyperthermophilic bacterium *Thermotoga maritima*. Appl Environ Microbiol **70:**6098-112.

39. **Pysz, M. A., D. E. Ward, K. R. Shockley, C. I. Montero, S. B. Conners, M. R. Johnson, and R. M. Kelly.** 2004. Transcriptional analysis of dynamic heat-shock

response by the hyperthermophilic bacterium *Thermotoga maritima*. Extremophiles **8:**209-17.

40. **Raasch, C., W. Streit, J. Schanzer, M. Bibel, U. Gosslar, and W. Liebl.** 2000. *Thermotoga maritima* AglA, an extremely thermostable NAD+-, Mn2+-, and thiol-dependent alpha-glucosidase. Extremophiles **4:**189-200.

41. **Rajashekhara, E., M. Kitaoka, Y. K. Kim, and K. Hayashi.** 2002. Characterization of a cellobiose phosphorylase from a hyperthermophilic eubacterium, *Thermotoga maritima* MSB8. Biosci Biotechnol Biochem **66:**2578-86.

42. **Sakane, I., M. Ikeda, C. Matsumoto, T. Higurashi, K. Inoue, K. Hongo, T. Mizobata, and Y. Kawata.** 2004. Structural stability of oligomeric chaperonin 10: the role of two beta-strands at the N and C termini in structural stabilization. J Mol Biol **344:**1123-33.

43. **Sapra, R., M. F. Verhagen, and M. W. Adams.** 2000. Purification and characterization of a membrane-bound hydrogenase from the hyperthermophilic archaeon *Pyrococcus furiosus*. J Bacteriol **182:**3423-8.

44. **Schut, G. J., J. Zhou, and M. W. Adams.** 2001. DNA microarray analysis of the hyperthermophilic archaeon *Pyrococcus furiosus*: evidence for anNew type of sulfur-reducing enzyme complex. J Bacteriol **183:**7027-36.

45. **Schwarzenbacher, R., J. M. Canaves, L. S. Brinen, X. Dai, A. M. Deacon, M. A. Elsliger, S. Eshaghi, R. Floyd, A. Godzik, C. Grittini, S. K. Grzechnik, C. Guda, L. Jaroszewski, C. Karlak, H. E. Klock, E. Koesema, J. S. Kovarik, A. Kreusch, P. Kuhn, S. A. Lesley, D. McMullan, T. M. McPhillips, M. A.**

Miller, M. D. Miller, A. Morse, K. Moy, J. Ouyang, A. Robb, K. Rodrigues, T. L. Selby, G. Spraggon, R. C. Stevens, H. van den Bedem, J. Velasquez, J. Vincent, X. Wang, B. West, G. Wolf, K. O. Hodgson, J. Wooley, and I. A. Wilson. 2003. Crystal structure of uronate isomerase (TM0064) from *Thermotoga maritima* at 2.85 A resolution. Proteins **53:**142-5.

46. **Schwarzenbacher, R., F. von Delft, P. Abdubek, E. Ambing, T. Biorac, L. S. Brinen, J. M. Canaves, J. Cambell, H. J. Chiu, X. Dai, A. M. Deacon, M. DiDonato, M. A. Elsliger, S. Eshagi, R. Floyd, A. Godzik, C. Grittini, S. K. Grzechnik, E. Hampton, L. Jaroszewski, C. Karlak, H. E. Klock, E. Koesema, J. S. Kovarik, A. Kreusch, P. Kuhn, S. A. Lesley, I. Levin, D. McMullan, T. M. McPhillips, M. D. Miller, A. Morse, K. Moy, J. Ouyang, R. Page, K. Quijano, A. Robb, G. Spraggon, R. C. Stevens, H. van den Bedem, J. Velasquez, J. Vincent, X. Wang, B. West, G. Wolf, Q. Xu, K. O. Hodgson, J. Wooley, and I. A. Wilson.** 2004. Crystal structure of a putative PII-like signaling protein (TM0021) from *Thermotoga maritima* at 2.5 A resolution. Proteins **54:**810-3.

47. **Selig, M., K. B. Xavier, H. Santos, and P. Schonheit.** 1997. Comparative analysis of Embden-Meyerhof and Entner-Doudoroff glycolytic pathways in hyperthermophilic archaea and the bacterium *Thermotoga*. Arch Microbiol **167:**217-32.

48. **Silva, P. J., E. C. van den Ban, H. Wassink, H. Haaker, B. de Castro, F. T. Robb, and W. R. Hagen.** 2000. Enzymes of hydrogen metabolism in *Pyrococcus furiosus*. Eur J Biochem **267:**6541-51.

49. **Takeda, K., C. Akimoto, and M. Kawamukai.** 2001. Effects of the *Escherichia coli sfsA* gene on *mal* genes expression and a DNA binding activity of SfsA. Biosci Biotechnol Biochem **65:**213-7.

50. **Van de Casteele, M., C. Legrain, V. Wilquet, and N. Glansdorff.** 1995. The dihydrofolate reductase-encoding gene *dyrA* of the hyperthermophilic bacterium *Thermotoga maritima*. Gene **158:**101-5.

51. **Verhagen, M. F., T. O'Rourke, and M. W. Adams.** 1999. The hyperthermophilic bacterium, *Thermotoga maritima*, contains an unusually complex iron-hydrogenase: amino acid sequence analyses versus biochemical characterization. Biochim Biophys Acta **1412:**212-29.

52. **Verhagen, M. F., T. W. O'Rourke, A. L. Menon, and M. W. Adams.** 2001. Heterologous expression and properties of the gamma-subunit of the Fe-only hydrogenase from *Thermotoga maritima*. Biochim Biophys Acta **1505:**209-19.

53. **Wassenberg, D., W. Liebl, and R. Jaenicke.** 2000. Maltose-binding protein from the hyperthermophilic bacterium *Thermotoga maritima*: stability and binding properties. J Mol Biol **295:**279-88.

54. **Wolfinger, R. D., G. Gibson, E. D. Wolfinger, L. Bennett, H. Hamadeh, P. Bushel, C. Afshari, and R. S. Paules.** 2001. Assessing gene significance from cDNA microarray expression data via mixed models. Journal of Computational Biology **8:**625-637.

55. **Woodward, J., N. I. Heyer, J. P. Getty, H. M. O'Neill, E. Pinkhassik, and B. R. Evans.** 2002. Presented at the Proceedings of the 2002 U.S. DOE Hydrogen Program Review.

56. **Wu, S. P., S. S. Mansy, and J. A. Cowan.** 2005. Iron-sulfur cluster biosynthesis. Molecular chaperone DnaK promotes IscU-bound [2Fe-2S] cluster stability and inhibits cluster transfer activity. Biochemistry **44:**4284-93.

57. **Yernool, D. A., J. K. McCarthy, D. E. Eveleigh, and J. D. Bok.** 2000. Cloning and characterization of the glucooligosaccharide catabolic pathway beta-glucan glucohydrolase and cellobiose phosphorylase in the marine hyperthermophile *Thermotoga neapolitana*. Journal of Bacteriology **182:**5172-5179.

58. **Zhang, R. G., Y. Kim, T. Skarina, S. Beasley, R. Laskowski, C. Arrowsmith, A. Edwards, A. Joachimiak, and A. Savchenko.** 2002. Crystal structure of *Thermotoga maritima* 0065, a member of the IclR transcriptional factor family. J Biol Chem **277:**19183-90.

Table 4.1. Sulfur-dependent changes in *T. maritima* transcripts during batch growth

Gene ID	Gene name	Putative function	LF_CE_CS	LF_MA_MS	LF_CE_MA	References
Genes with lower transcript levels in the presence of sulfur						
TM0266	*dbh*	DNA-binding protein HU	1.0	1.0	-1.6	(15, 30)
TM0477	*ompA*	Outer membrane protein alpha	1.0	1.1	1.0	
TM0654	*speE*	Spermidine synthase (polyamine synthesis)	0.9	1.0	-0.6	
TM0823	*tetR*	TetR family transcriptional regulator	1.0	1.1	1.7	
TM0824		Predicted Fe-S protein, SAM superfamily	0.9	0.9	0.5	
TM0980	*dsrF*	DsrF-like protein	1.2	1.2	0.5	
TM0981	*dsrE*	DsrE-like protein	1.3	1.3	1.6	
TM0982		Predicted membrane protein, possible transporter	1.2	1.2	1.4	
TM1485		Ribosomal protein L6	1.1	1.1	2.1	
Genes with higher transcript levels in the presence of sulfur						
TM0159	*ham1*	Xanthosine triphosphate pyrophosphatase, putative	-0.9	-0.9	-0.8	(10, 11)
TM0375	*pduL*	Propanediol utilization protein, putative	-0.9	-0.9	-1.0	
TM1292	*sipB*	Iron-sulfur cluster binding protein, putative	-1.1	-1.2	-1.2	
TM1641	*dyrA*	Dihydrofolate reductase	-1.3	-1.3	-1.5	(50)
TM1681		71 aa hypothetical protein	-1.0	-1.0	-1.3	

Table 4.2. Carbohydrate-dependent changes in *T. maritima* transcripts during batch growth

Gene ID	Gene name	Putative function	Log_2 Fold (CE_MA)	Log_2 Fold (CS_MS	Structure	Reference
Transcripts detected at higher levels during batch growth on cellobiose						
TM0013		Putative metallo beta-lactamase	1.1	1.1		
TM0014		Methyl-accepting chemotaxis protein	0.7	0.6		
TM0015	*porC*	Pyruvate ferredoxin oxidoreductase, gamma subunit	1.2	1.2		(25)
TM0016	*porD*	Pyruvate ferredoxin oxidoreductase, delta subunit	1.5	1.6		
TM0017	*porA*	Pyruvate ferredoxin oxidoreductase, alpha subunit	1.7	1.7		
TM0018	*porB*	Pyruvate ferredoxin oxidoreductase, beta subunit	1.0	1.0		
TM0019		Putative oxidoreductase/dehydrogenase	0.7	0.6		
TM0020		Protein crcB homolog.	0.7	0.7		
TM0021	*glnB*	PII like protein	0.5	0.5	1O51	(46)
TM0050[B]	*feoA*	Iron(II) transport protein A	1.1	1.2		
TM0051[B]	*feoB*	Iron(II) transport protein B	1.4	1.5		
TM0064	*uxaC/ uxuC*	Uronate isomerase, putative	0.8	0.8		(45)
TM0065	*kdgR*	Transcriptional regulator, IclR family	1.0	1.0		(58)
TM0066	*kdgA*	2-dehydro-3-deoxyphosphogluconate (KDG) aldolase	1.4	1.5		
TM0067	*kdgK*	2-keto-3-deoxygluconate kinase	1.0	1.0		
TM0068	*uxuB*	D-mannonate oxidoreductase	1.8	1.8		
TM0069	*uxuA*	D-mannonate hydrolase	1.2	1.3		
TM0169	*rex1*	Redox-sensing transcriptional repressor, putative	1.2	1.2		
TM0208	*pykA*	Pyruvate kinase	1.6	1.6		(20)
TM0209	*6-pfk*	6-phosphofructokinase, PPi-dependent	1.0	1.0		(14)
TM0272	*ppdK*	Pyruvate, orthophosphate dikinase	1.7	1.7		
TM0273	*fba*	Fructose-bisphosphate aldolase	1.9	1.9		
TM0274	*ackA*	Acetate kinase	1.7	1.7		
TM0275	*araR*	Transcriptional regulator, putative arabinose repressor	1.5	1.6		

(Table 4.2, continued)

TM0291	*leuC*	3-isopropylmalate dehydratase, large subunit [4Fe-4S] cluster	1.3	1.4	
TM0292	*leuD*	3-isopropylmalate dehydratase, small subunit	1.0	1.0	
TM0373[B]	*dnaK*	Molecular chaperone DnaK; binds and inhibits Fe-S cluster transfer activity of IscU, a scaffold protein for Fe-S cluster assembly	1.0	1.0	(56)
TM0374[B]	*sHsp*	Small heat shock protein	2.0	2.1	
TM0379[B]	*npo*	Homolog of *T. neapolitana* NADH:polysulfide oxidoreductase (GI:21702687)	2.9	2.9	(8)
TM0380		Frameshift	1.7	1.7	
TM0381		Dihydrolipoamide dehydrogenase homolog	4.1	4.1	
TM0393		XylR family regulator	1.3	1.3	
TM0394[B]	*gltB_3*	Archaeal-type glutamate synthase domain 3	1.9	1.9	
TM0395[B]		NADH oxidase, putative	2.5	2.5	
TM0396[B]		Iron-sulfur cluster-binding protein.	2.2	2.3	
TM0397[B]	*gltB_2*	Archaeal-type glutamate synthase domain 2	1.0	0.9	
TM0398[B]	*gltB_1*	Archaeal-type glutamate synthase domain 1	0.9	0.9	
TM0505	*groES*	10 kDa chaperonin (CPN10, GroES)	2.7	2.8	(42)
TM0506	*groEL*	60 kDa chaperonin (CPN60, GroEL)	3.3	3.4	
TM0687		Hypothetical UPF0133 protein	1.3	1.3	
TM0688	*gapDH*	Glyceraldehyde 3-phosphate dehydrogenase	1.0	1.0	
TM0689	*pgk/tim*	Bifunctional phosphoglycerate kinase/triose-phosphate isomerase	1.7	1.8	
TM0690		Hypothetical protein TM0690	1.6	1.6	
TM0691		Hypothetical protein TM0691	1.4	1.3	
TM0692		Holo-[acyl-carrier protein] synthase] (Holo-ACP synthase) (4'-phosphopantetheinyl transferase acpS).	1.0	1.0	
TM0693		Hypothetical protein TM0693	1.9	1.9	
TM0694		Trigger factor (TF)	2.2	2.3	
TM0695		ATP-dependent Clp protease proteolytic subunit (Endopeptidase Clp)	1.5	1.5	

(Table 4.2, continued)

TM0697	*fliQ*	Flagellar biosynthesis protein FliQ	0.9	0.9	
TM0701		Purine-binding chemotaxis protein	0.9	0.9	
TM0702	*cheA*	Chemotaxis sensor histidine kinase	2.1	2.2	
TM0902	*fliA*	Flagellar sigma factor σ^{28}	0.9	0.9	
TM0903		Chemotaxis methylation protein	1.0	1.0	
TM0904	*cheC*	Chemotaxis protein CheC	1.3	1.4	
TM0905		Uncharacterized conserved protein with similarity to glycosyltransferases (COG5581)	1.1	1.2	
TM0979	*dsrH*	DsrH-like protein	1.3	1.3	
TM0980	*dsrF*	DsrF-like protein	1.7	1.7	
TM0981	*dsrE*	DsrE-like protein	1.4	1.4	
TM0982		Predicted membrane protein, possible transporter	1.4	1.4	
TM1041	*hisD*	Histidinol dehydrogenase L-histidinol + 2 NAD(+) = L-histidine + 2 NADH	1.0	1.0	
TM1218	*celR*	Putative regulator of β-glucan response genes	1.4	1.6	
TM1219	*cbtF*	Cellobiose ABC transporter, ATP-binding subunit	2.4	2.4	
TM1220	*cbtD*	Cellobiose ABC transporter, ATP-binding subunit	3.0	3.0	
TM1221	*cbtC*	Cellobiose ABC transporter, permease subunit	2.4	2.4	
TM1222	*cbtB*	Cellobiose ABC transporter, permease subunit	1.3	1.4	
TM1223	*cbtA*	Cellobiose ABC transporter, substrate-binding protein	4.5	4.5	
TM1420		[2Fe-2S] protein, possibly involved in hydrogenase metal cluster assembly	0.8	0.8	(35)
TM1421		Putative hydrogenase component or ferredoxin	0.7	0.7	
TM1422		Iron-sulfur protein	1.0	1.1	
TM1423			0.6	0.6	
TM1424	*hydC*	Fe-hydrogenase, gamma subunit 1 [2Fe-2S]	0.0	0.0	(51, 52)
TM1425	*hydB*	Fe-hydrogenase, beta subunit 1 [2Fe-2S], 3 [4Fe-4S]	1.3	1.4	(51)
TM1426	*hydA*	Fe-hydrogenase, alpha subunit 2[2Fe-2S], 3 [4Fe-4S], H-cluster (catalytic site)	1.6	1.6	(51)

(Table 4.2, continued)

TM1427	*rex2*	Redox-sensing transcriptional repressor, putative	1.1	1.1		
TM1428		Methyl accepting chemotaxis protein	1.2	1.2		
TM1451[B]	*sigA*	General sigma factor σ^A	1.1	1.1		
TM1468		Uncharacterized conserved protein	1.0	1.0	1VPV	
TM1469		ATP-dependent glucokinase	0.8	0.8		(17)
TM1470		Rho transcription termination factor	1.7	1.7		
TM1471		Ribosomal protein L17	1.3	1.3		
TM1472		RNA polymerase alpha subunit	1.6	1.6		
TM1473		Ribosomal protein S4	2.2	2.3		
TM1474		Ribosomal protein S11	1.7	1.7		
TM1475		Ribosomal protein S13	1.8	1.9		
TM1476		Ribosomal protein L36	1.1	1.2		
TM1477			0.8	0.8		
TM1478		Methionine aminopeptidase	1.2	1.2		
TM1479		Adenylate kinase	2.1	2.1		
TM1480		Preprotein translocase SecY subunit	1.7	1.8		
TM1481		Ribosomal protein L15	1.1	1.1		
TM1482		Ribosomal protein L30	1.7	1.7		
TM1483		Ribosomal protein S5	2.0	2.0		
TM1484		Ribosomal protein L18	1.9	2.0		
TM1485		Ribosomal protein L6	2.1	2.1		
TM1486		Ribosomal protein S8	2.0	2.0		
TM1487		Ribosomal protein S14	1.1	1.1		
TM1488		Ribosomal protein L5	2.3	2.3		
TM1489		Ribosomal protein L24	2.2	2.2		
TM1490		Ribosomal protein L14	1.1	1.1		
TM1491		Ribosomal protein S17	2.2	2.2		
TM1492		Ribosomal protein L29	1.7	1.8		
TM1493		Ribosomal protein L16	1.4	1.4		
TM1494		Ribosomal protein S3	2.1	2.2		
TM1495		Ribosomal protein L22	2.0	2.0		
TM1496		Ribosomal protein S19	-0.6	-0.6		
TM1497		Ribosomal protein L2	1.8	1.8		
TM1498		Ribosomal protein L23	2.2	2.3		
TM1499		Ribosomal protein L4	2.5	2.5		
TM1500		Ribosomal protein L3	1.8	1.9		
TM1501		Ribosomal protein S10	1.5	1.4		
TM1502		Translation elongation factor Tu	1.7	1.8		
TM1503		Translation elongation factor G	1.3	1.3		
TM1504		Ribosomal protein S7	1.7	1.7		
TM1505		Ribosomal protein S12	0.7	0.7		
TM1524[B]		Engoglucanase	1.8	1.8		
TM1525[B]		Endoglucanase	2.2	2.2		

(Table 4.2, continued)

TM1565		Signal recognition particle GTPase	1.2	1.2	
TM1566		Ribosomal protein S16	1.2	1.2	
TM1567		Predicted KH domain RNA-binding protein	0.6	0.6	
TM1568	*rimM*	16S rRNA processing protein	1.3	1.3	
TM1569		tRNA guanine-N1 methyltransferase	1.7	1.7	
TM1570		Uncharacterized conserved protein	1.1	1.1	
TM1571		ribosomal protein L19	1.2	1.2	
TM1572		Signal peptidase I	1.5	1.5	
TM1578		Preprotein SecA subunit	1.1	1.2	
TM1590		Translation initiation factor IF-3	1.0	1.0	
TM1591		Ribosomal protein L35	1.0	1.1	
TM1592		Ribosomal protein L20	0.5	0.5	
TM1593		ATPase involved in chromosome partitioning	1.0	1.1	
TM1605		Translation elongation factor Ts	2.1	2.1	
TM1606		Putative ribonuclease	1.3	1.3	
TM1607		SSU ribosomal protein S30P	1.2	1.2	
TM1608		Uncharacterized DegV family protein (COG1307)	1.7	1.7	
TM1609		ATP synthase F1, subunit epsilon	0.4	0.3	
TM1610		ATP synthase F1, subunit beta	0.5	0.5	
TM1611		ATP synthase F1, subunit gamma	1.3	1.4	
TM1612		ATP synthase F1, subunit alpha	0.8	0.7	
TM1613		ATP synthase F1, subunit delta	1.5	1.5	
TM1627	*ctc*	50S ribosomal protein L25, general stress protein	1.2	1.3	
TM1848[B]	*cepA*	Cellobiose phosphorylase	2.3	2.2	(41)
Genes detected at higher levels during maltose batch growth					
TM0178		Primosomal protein N	-1.0	-1.1	
TM0179		Uncharacterized small protein conserved in *Thermotoga* species	-1.0	-0.8	
TM0180		Uncharacterized protein conserved in *Thermotoga* species	-2.1	-2.1	
TM0181		Hypothetical protein related to DNA polymerase III delta subunit	-1.3	-1.2	

(Table 4.2, continued)

TM0182		Putative iron-sulfur oxidoreductase also matching to pfam04055, Radical_SAM superfamily	-1.2	-1.2
TM0186		Response regulator	-0.9	-0.9
TM0187		Sensor histidine kinase	-0.9	-1.0
TM0188		Uncharacterized conserved protein	-0.5	-0.3
TM0189		Iron(III) ABC transporter, substrate-binding protein	-0.9	-0.9
TM0190		Iron (III) ABC transporter, permease protein	-1.0	-1.2
TM0191		Iron (III) ABC transporter, ATP-binding protein	-0.4	-0.3
TM0192		SpoVS-related protein (sporulation in *B. subtilis*, function unclear in *Thermotoga*)	-1.3	-1.3
TM0196		Uncharacterized conserved protein	-1.4	-1.5
TM0197		Predicted ATPase involved in cell cycle control	-1.1	-1.1
TM0211		Aminomethyltransferase-glycine cleavage system T protein	-1.6	-1.6
TM0228		NADP-reducing hydrogenase, subunit C	-1.1	-1.0
TM0231	*murC*	UDP-N-acetylmuramate--L-alanine ligase, peptidoglycan synthesis	-0.9	-0.9
TM0233		SpoVE-related protein, cell cycle protein	-1.2	-1.2
TM0236	*murF*	UDP-MURNAC-pentapeptide sythetase	-1.0	-1.0
TM0266		DNA-binding protein, HU	-1.6	-1.5
TM0267		Protein related to thiophene oxidation proteins	-0.3	-0.2
TM0268		5-methyltetrahydrofolate S-homocysteine methyltransferase	-0.9	-1.0
TM0269		Methionine synthase I, activation domain	-1.5	-1.5
TM0345		3-phosphoshikimate-1-carboxyvinyltransferase	-1.1	-1.1
TM0346		Shikimate 5-dehydrogenase	0.1	0.0
TM0347		chorismate synthase	-1.0	-1.0

(Table 4.2, continued)

TM0348		shikimate kinase/3-dehydroquinate synthas	-0.8	-0.8		
TM0356		Threonine dehydratase	-1.1	-1.0		
TM0357		Small orf (68 aa)	0.2	0.2		
TM0358		Enzyme related to eukaryotic diacylglycerol kinase	-1.2	-1.3		
TM0412		Zn-dependent hydrogenase, similarity to threonine dehydrogenase	-1.2	-1.1		
TM0413		Putative amidase	-1.0	-1.0		
TM0414		Putative myo-inositol 2-dehydrogenase	-1.3	-1.2		
TM0415		Putative carbohydrate kinase	-0.9	-1.0	1VK4	
TM0450		Small (100 aa) hypothetical protein	-1.5	-1.5		
TM0451		Ribosomal protein L33	-1.8	-1.8		
TM0480	*uvrA*	UvrA ATPase	-1.2	-1.2		
TM0481		Transmembrane protein	-1.1	-1.0		
TM0487		Conserved hypothetical protein	-0.9	-0.9	1WCJA	
TM0488		N5-Glutamine Methyltransferase	-1.2	-1.2		
TM0504		Signaling peptide	-1.3	-1.4		(21)
TM0515		Predicted iron-sulfur oxidoreductase	-1.2	-1.2		
TM0519		GAF domain-containing protein	-0.9	-0.9		
TM0560[B]		Bacterioferritin-like domain	-1.6	-1.7		
TM0561		Magnesium and Cobalt transporter CorA	-1.1	-1.0		
TM0565	*sfsA*	Sugar fermentation stimulation protein homolog; may be involved in maltose regulation and a homolog binds DNA	-1.3	-1.3		
TM0594		Putative permease	-1.0	-1.2		
TM0595		ABC sugar binding protein	-1.7	-1.7		
TM0619		Glycosyltransferase group 1	-1.0	-1.0		
TM0659[B]		Rubredoxin	-1.7	-1.7		
TM0662		Acyl carrier protein	-1.3	-1.2		

(Table 4.2, continued)

TM0669		Hypothetical protein	-1.2	-1.2		
TM0670		Hypothetical protein	-1.0	-1.0		
TM0728		Drug/metabolite superfamily (DMT) transporter	-1.0	-1.1		
TM0729		(p)ppGpp synthetase	-1.4	-1.4		
TM0766		GntR family transcriptional regulator	-1.3	-1.2		
TM0767		Maltosyltransferase	-0.9	-0.9		
TM0809		beta-N-acetylglucosaminidase CbsA	-0.9	-0.9		
TM0810		Sugar ABC transporter, substrate-binding protein	-0.9	-0.9		
TM0811		Sugar ABC transporter, permease subunit	-1.4	-1.4		
TM0812		Sugar ABC transporter, permease subunit	-1.5	-1.5		
TM0813		Glucosamine-6-phosphate deaminase	-1.4	-1.5	1J5X	
TM0927		Ferredoxin	-1.8	-1.8		
TM0935		CBS domain-containing protein	-1.1	-1.1	1O50	
TM0936		Putative metal-dependent hydrolase	-0.9	-0.9	1P1MA, 1J6PA	
TM0937		Predicted transcriptional regulator, PadR family	-0.9	-1.0		
TM0948		Predicted Fe-S-oxidoreductase	-1.2	-1.3		
TM1161	*mgtE*	Predicted Mg2+ transporter	-1.4	-1.4		
TM1432	*gspA*	Putative NAD+ or NADP+-dependent glycerol-3-phosphate dehydrogenase, *gspA* homolog (*A. fulgidus*) [2Fe-2S]	-1.2	-1.2	1Z82B	
TM1433		Pyridine nucleotide-disulphide oxidoreductase	-1.0	-0.9		
TM1434		Uncharacterized conserved protein, 4 conserved cysteines	0.0	0.0		
TM1435		Predicted CoA-binding protein (COG1832)	-1.0	-1.0		
TM1436		Putative glycerol-3-phosphate responsive antiterminator	-1.5	-1.4	1VKFD	
TM1439		Uncharacterized conserved protein (COG3330)	-1.4	-1.4		
TM1641	*dfrA*	NADP+-dependent dihydrofolate reductase	-1.5	-1.5		(10-13)

(Table 4.2, continued)

TM1650	*amy13B*	α-amylase	-1.1	-1.1
TM1681		71 aa hypothetical protein	-1.3	-1.3
TM1700		Haloacid dehalogenase-like hydrolase	-1.0	-1.0
TM1701		Conserved hypothetical with MatE (Multi Antimicrobial Extrusion) domain	-0.9	-1.0
TM1724		Short chain alcohol dehydrogenase	-1.0	-1.0
TM1834	*aglA*	α-glucosidase	-0.7	-0.6
TM1839	*tmmbp*	Maltose-binding protein	-0.9	-0.9
TM1840	*amy13A*	α-amylase	-0.6	-0.6
TM1841		Putative alpha-linked glucanase	-1.1	-1.1
TM1867	*ldh*	L-lactate dehydrogenase	-1.1	-1.1
TM1874	*cspA*	Cold shock protein	-1.9	-1.9

FIGURE LEGENDS

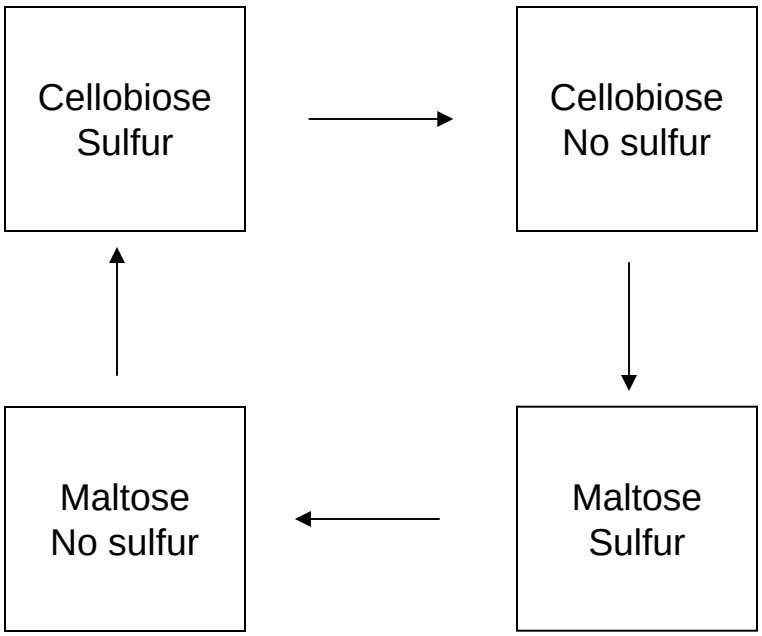
FIG. 4.1. Loop designs for batch growth and continuous culture microarray experiments. Arrows represent arrays, with the arrowhead representing the Cy5 channel and the arrow tail representing the Cy3 channel. a.) Four independent samples were taken from mid-log batch cultures containing maltose and cellobiose, with or without sulfur. Each sample was labeled once with Cy3 and once with Cy5. b.) Six samples were taken from a continuous culture run.

FIG. 4.2. Least squares mean estimates of treatment effects for cellobiose utilization genes (TM1218-1223) and genes within the locus containing a NADH: polysulfide oxidoreductase (TM0379).

FIG. 4.3. Maltose and cellobiose transport, hydrolysis and catabolic pathways include the glycolytic pathways of *T. maritima*, including enzymes from the EMP and ED pathways. Genes whose transcripts were detected at higher levels during cellobiose growth in batch cultures are highlighted in yellow, while genes with higher detected transcripts during growth on maltose are highlighted in blue. Abbreviations used include P (phosphate), G-6-P (glucose-6-phosphate), PFK (phosphofructokinase), FBP (fructose-bisphosphate), KDPG (2-keto-3-deoxy-6-phosphogluconate), KDG (2-keto-3-deoxy-6-gluconate), 6-PG-DH (6-phosphogluconate dehydrogenase), 6-PG (6-phosphogluconate), PEP (phosphoenolpyruvate), POR (pyruvate:ferredoxin oxidoreductase),

FIG. 4.4. Organization of the hydrogenase operon of *T. maritima*. Asterisk represents a potential binding site for a redox-sensitive transcriptional regulator, potentially TM1427, a member of the Rex family of NADH-sensitive transcriptional regulators. Grey shading indicates genes from this operon which have been characterized.

a.


Cellobiose
Sulfur
Cellobiose
No sulfur
Maltose
No sulfur
Maltose
Sulfur

b.

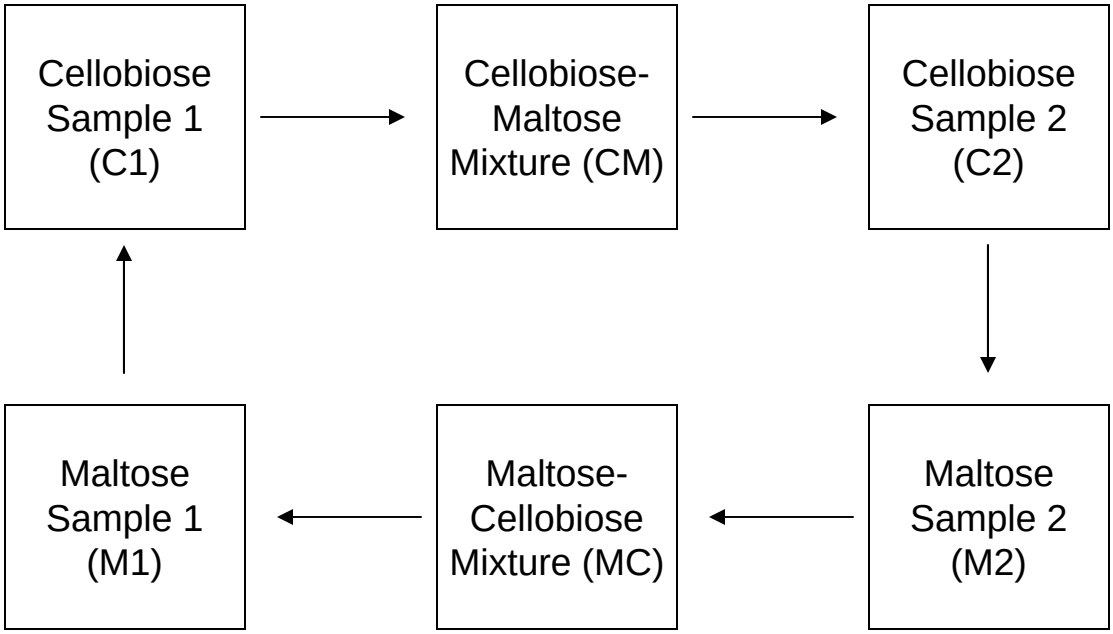
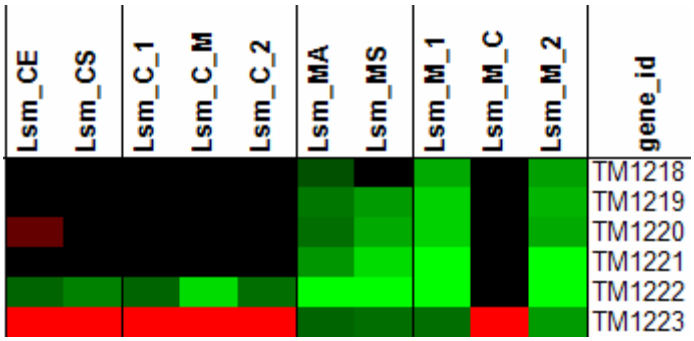

Cellobiose
Sample 1
(C1)
Cellobiose-
Maltose
Mixture (CM)
Cellobiose
Sample 2
(C2)
Maltose
Sample 1
(M1)
Maltose-
Cellobiose
Mixture (MC)
Maltose
Sample 2
(M2)

Figure 4.1


Lsm_CE
Lsm_CS
Lsm_C_1
Lsm_C_M
Lsm_C_2
Lsm_MA
Lsm_MS
Lsm_M_1
Lsm_M_C
Lsm_M_2
gene_id
TM1218
TM1219
TM1220
TM1221
TM1222
TM1223

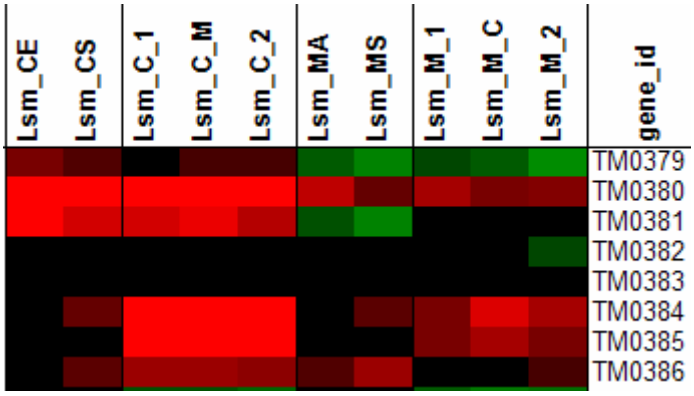

Lsm_CE
Lsm_CS
Lsm_C_1
Lsm_C_M
Lsm_C_2
Lsm_MA
Lsm_MS
Lsm_M_1
Lsm_M_C
Lsm_M_2
gene_id
TM0379
TM0380
TM0381
TM0382
TM0383
TM0384
TM0385
TM0386

Figure 4.2

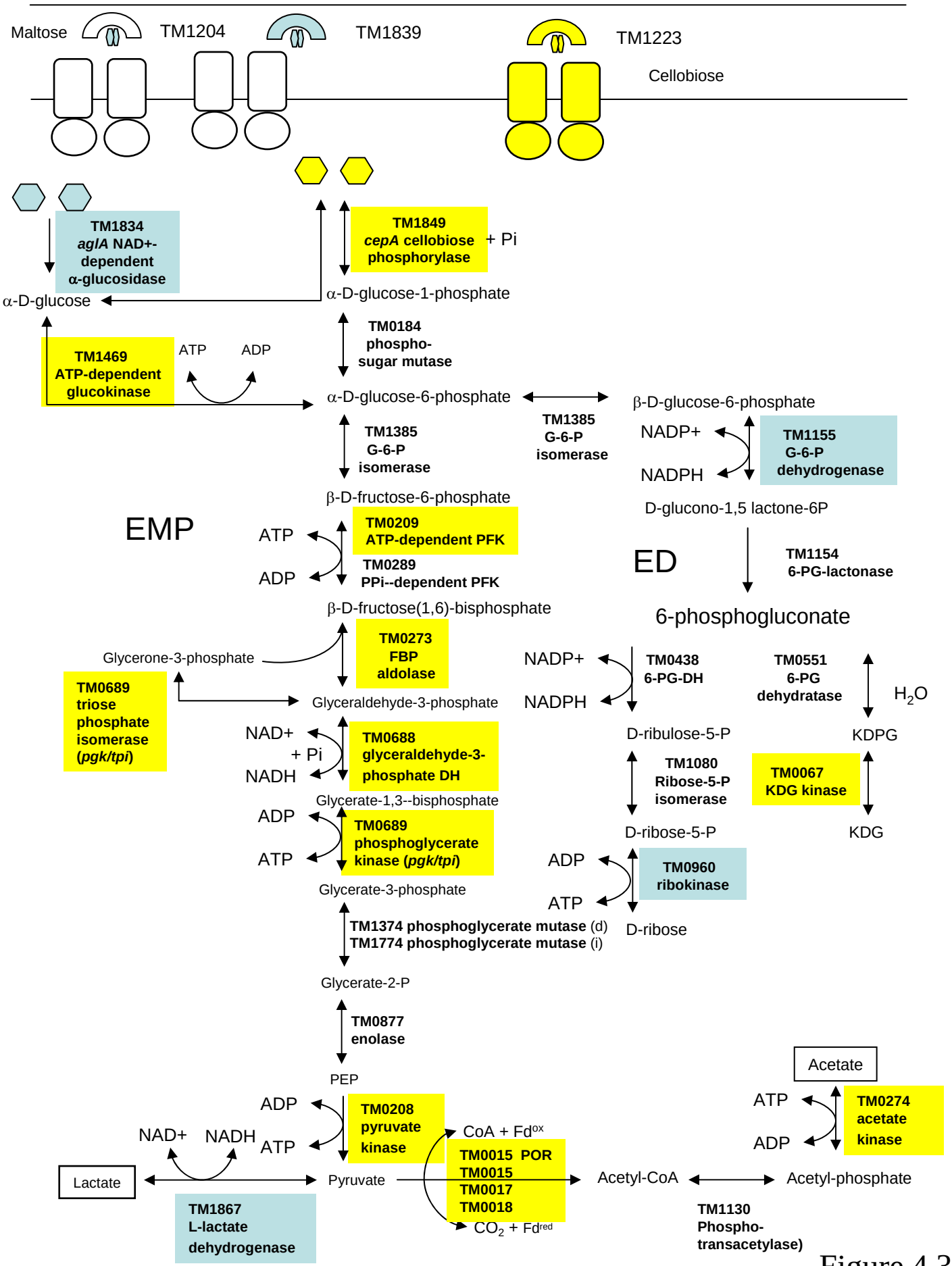
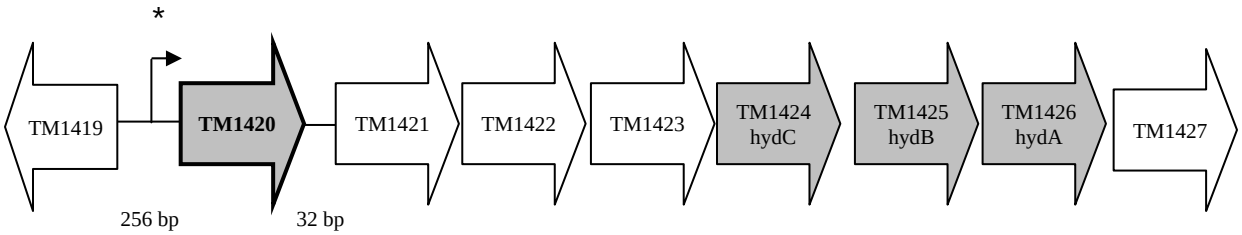

Maltose
TM1204
TM1839
TM1223
Cellobiose
TM1834 aglA NAD+-dependent α-glucosidase
TM1849 cepA cellobiose phosphorylase
+ Pi
α-D-glucose
α-D-glucose-1-phosphate
TM1469 ATP-dependent glucokinase
ATP
ADP
TM0184 phospho-sugar mutase
α-D-glucose-6-phosphate
TM1385 G-6-P isomerase
β-D-glucose-6-phosphate
NADP+
NADPH
TM1155 G-6-P dehydrogenase
TM1385 G-6-P isomerase
β-D-fructose-6-phosphate
D-glucono-1,5 lactone-6P
EMP
ATP
ADP
TM0209 ATP-dependent PFK
TM0289 PPi--dependent PFK
ED
TM1154 6-PG-lactonase
β-D-fructose(1,6)-bisphosphate
6-phosphogluconate
Glycerone-3-phosphate
TM0273 FBP aldolase
TM0689 triose phosphate isomerase (pgk/tpi)
Glyceraldehyde-3-phosphate
NADP+
NADPH
TM0438 6-PG-DH
TM0551 6-PG dehydratase
H2O
D-ribulose-5-P
KDPG
NAD+
+ Pi
NADH
TM0688 glyceraldehyde-3-phosphate DH
TM1080 Ribose-5-P isomerase
TM0067 KDG kinase
Glycerate-1,3--bisphosphate
D-ribose-5-P
KDG
ADP
ATP
TM0689 phosphoglycerate kinase (pgk/tpi)
ADP
ATP
TM0960 ribokinase
Glycerate-3-phosphate
D-ribose
TM1374 phosphoglycerate mutase (d)
TM1774 phosphoglycerate mutase (i)
Glycerate-2-P
TM0877 enolase
PEP
Acetate
ATP
ADP
TM0274 acetate kinase
ADP
ATP
TM0208 pyruvate kinase
NAD+
NADH
CoA + Fdox
TM0015 POR
TM0015
TM0017
TM0018
Lactate
Pyruvate
Acetyl-CoA
Acetyl-phosphate
TM1867 L-lactate dehydrogenase
CO2 + Fdred
TM1130 Phospho-transacetylase)

Figure 4.3




```
TGTTACCTACTTCACA  -94  -79 TM1420
TGTGACCTGCTTCACA -107  -92 SC_nuo
```

Figure 4.4

RESEARCH CONTRIBUTIONS

1. **Chhabra, S. R., K. R. Shockley, S. B. Conners, K. L. Scott, R. D. Wolfinger, and R. M. Kelly.** 2003. Carbohydrate-induced differential gene expression patterns in the hyperthermophilic bacterium *Thermotoga maritima*. J Biol Chem **278:**7540-52.
2. **Comfort, D. A., S. R. Chhabra, S. B. Conners, C. J. Chou, K. L. Epting, M. R. Johnson, K. L. Jones, A. C. Sehgal, and R. M. Kelly.** 2004. Strategic biocatalysis with hyperthermophilic enzymes. Green Chemistry **6:**459-465.
3. **Conners, S. B., C. I. Montero, D. A. Comfort, K. R. Shockley, M. R. Johnson, S. R. Chhabra, and R. M. Kelly.** 2005. An expression-driven approach to the prediction of carbohydrate transport and utilization regulons in the hyperthermophilic bacterium *Thermotoga maritima*. J Bacteriol **187:**7267-82.
4. **Johnson, M. R., C. I. Montero, S. B. Conners, K. R. Shockley, S. L. Bridger, and R. M. Kelly.** 2005. Population density-dependent regulation of exopolysaccharide formation in the hyperthermophilic bacterium *Thermotoga maritima*. Mol Microbiol **55:**664-74.
5. **Price, M. S., S. B. Conners, S. Tachdjian, R. M. Kelly, and G. A. Payne.** 2005. Aflatoxin conducive and non-conducive growth conditions reveal new gene associations with aflatoxin production. Fungal Genet Biol **42:**506-18.
6. **Pysz, M. A., S. B. Conners, C. I. Montero, K. R. Shockley, M. R. Johnson, D. E. Ward, and R. M. Kelly.** 2004. Transcriptional analysis of biofilm formation processes in the anaerobic, hyperthermophilic bacterium *Thermotoga maritima*. Appl Environ Microbiol **70:**6098-112.

7. **Pysz, M. A., D. E. Ward, K. R. Shockley, C. I. Montero, S. B. Conners, M. R. Johnson, and R. M. Kelly.** 2004. Transcriptional analysis of dynamic heat-shock response by the hyperthermophilic bacterium *Thermotoga maritima*. Extremophiles **8:**209-17.

8. **Shockley, K. R., K. L. Scott, M. A. Pysz, S. B. Conners, M. R. Johnson, C. I. Montero, R. D. Wolfinger, and R. M. Kelly.** 2005. Genome-wide transcriptional variation within and between steady states for continuous growth of the hyperthermophile *Thermotoga maritima*. Appl Environ Microbiol **71:**5572-6.

9. **Shockley, K. R., D. E. Ward, S. R. Chhabra, S. B. Conners, C. I. Montero, and R. M. Kelly.** 2003. Heat shock response by the hyperthermophilic archaeon *Pyrococcus furiosus*. Appl Environ Microbiol **69:**2365-71.

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