Title

Simultaneous Catabolism of Multiple Sugars and Development of Genetic Tools for Metabolic Engineering in Sulfolobus acidocaldarius

Author
Joshua, Chijioke Johnson

Publication Date
2012

Peer reviewed|Thesis/dissertation

Simultaneous Catabolism of Multiple Sugars and Development of Genetic Tools for Metabolic Engineering in *Sulfolobus acidocaldarius*

By

Chijioke Johnson Joshua

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Microbiology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Jay D. Keasling, Chair
Professor Bob B. Buchanan
Professor Krishna K. Niyogi
Professor Adam Arkin

Spring 2012

Simultaneous Catabolism of Multiple Sugars and Development of Genetic Tools for Metabolic Engineering in *Sulfolobus acidocaldarius*

By
Chijioke Johnson Joshua

ABSTRACT

Simultaneous Catabolism of Multiple Sugars and Development of Genetic Tools for Metabolic Engineering in *Sulfolobus acidocaldarius*

by

Doctor of Philosophy in Microbiology
University of California, Berkeley

Professor Jay D. Keasling, Chair

The demand for thermophilic microbial host systems for the production of bio-products has increased over the years due to the inadequacy of the currently available mesophilic systems. However, genetic tools for studying thermophilic organisms are sparingly available. This study evaluated the potential of *S. acidocaldarius*, a hyperthermophilic archaea, as a platform for metabolic engineering, especially cellulosic biofuels. Bioconversion of ligno-cellulosic biomass into biofuels involves the use of a wide array of thermophilic enzymes such as endo-glucanases and β-glycosidase, most of which are not properly expressed in *E. coli* and *S. cerevisiae*. *S. acidocaldarius* grows optimally at 75 – 80°C and pH 2-3. This organism utilizes most cellulosic sugars with the exception of cellobiose. The dissertation reports the absence of glucose-induced diauxie during consumption of multiple cellulosic sugars such as arabinose and xylose as co-carbon sources. The organism utilized combinations of 1 g/L each of glucose and xylose simultaneously with a specific growth rate of 0.079 h^{-1}. The organism did not show preference to glucose or any of the sugars tested. However, the organism grew faster on 2 g/L xylose (0.074 h^{-1}) than on equal amount of glucose (0.022 h^{-1}). The consumption of most cellulosic sugars by this organism makes it a potential candidate for cellulosic bio-products engineering. During growth on multiple sugars, the organism consumed each sugar at a rate that was roughly proportional to its concentration in the growth medium. The mechanism of this novel regulation is not fully understood and is currently being investigated. This study also focused on developing reliable genetic tools such as recipient strains with selectable traits and effective plasmid system to achieve the goal of making this organism a potent platform for metabolic engineering. A number of pRN1-based shuttle vectors were developed for heterologous expression in *S. acidocaldarius*. The role of a 241-bp region downstream of the *orf904* as the putative origin of replication of pRN1 was also investigated. The results indicated that *orf56, orf904* and the putative origin were the minimum replicon of pRN1. A 100-bp stem-loop structure within this putative origin was proposed to be the double-strand origin of replication of the plasmid. The impact of a functional β-glycosidase on cellobiose metabolism in *S. acidocaldarius* was evaluated by inserting *lacS* from *S. solfataricus* into the chromosome of the organism. The gene (*lacS*) established β-glycosidase activity *in vivo* but did not establish cellobiose metabolism, suggesting that *S. acidocaldarius* lacks a transporter for cellobiose. Overall, this study enhanced our knowledge of sugar metabolism in *S. acidocaldarius* and highlighted the development of genetic tools and strategy for utilizing the organism as a platform for metabolic engineering.

This is dissertation is dedicated

To the memory of my beloved brother, Charles Ndubuisi Joshua

And

To my beloved mother Nnenna Joy Igwe

For their unending love and support

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

This dissertation would not be possible without the mentorship of my research advisor Dr. Jay D. Keasling. I would love to thank him for guidance throughout the course of my study and for providing an enabling environment for me to succeed in my research endeavor. I have learned to be the best scientist that I can possibly be under his tutelage. Words can never describe my gratitude to Dr. Keasling for the opportunity to pursue and exercise my passion for knowledge, especially in the field of microbiology. I will forever remain grateful him for his advice, encouragements, constructive criticism and patience during my study.

I would love to extend my sincere appreciation and gratitude to the members of my thesis committee members: Dr. Bob Buchanan, Dr. Kris Niyogi and Dr. Adam Arkin, for their guidance and encouragements during my study. I would particularly thank my thesis committee members for taking time to review this dissertation. I am also grateful to the members of my qualifying examination committee: Dr. Steve Lindow, Dr. Bob Buchanan, Dr. Kris Niyogi and Dr. Mary Wildermuth for having faith in my ability as an aspiring young scientist.

I am indebted to my mother, Nnenna Igwe for love and support for my pursuit of education. I wish my brother Charles Joshua, was alive to celebrate this accomplishment with me, I will forever remain grateful to you! Adieu! Charles! My special regards go to my uncle Francis Igwe and his family for their encouragements. My heartfelt appreciations go to the following people and their families for their support and encouragements during the course of this study: Mr. Tunde Adedayo, Nwachukwu Nmaju, Ozioma, Madam Chika Emelike, Pastor and Mrs. Nosike.

My life at Berkeley would not have been balanced without the love and encouragements from members of the Berkeley United Methodist Church to whom I would remain loyal and appreciative. I am particularly indebted to Marsh and Irene Tekawa, and the Tekawa-Pons (Kevin, Cheryl, Justin and Ethan) for always being there for me.

I remain grateful to Dr. Edward Baidoo and Dr. Clem Fortman for being great friends and mentors, especially for always being there during difficult times. My appreciations go to Dr. Peter Benke and Dr. Rob Dahl for their immense contributions to my research. I am grateful to Dr. Blake Simmons and Dr. Leonard Katz for their encouragements and genuine interest in academic and professional well being. I have to specially thank Professor Sidney Kustu for being a great inspiration to me, words cannot express my gratitude for the knowledge that I acquired during my rotation in her lab.

I would seize this opportunity to express my appreciations to Juan, Malik, Luis and Kulika for their contributions to my research during the period they worked with me. I would also love to thank Dr. Brad Holmes, Dr. Danielle Augustine, Kevin Hockett (Now Dr. Hockett), Karen Wong, Beverly Bitagon, Dr. Jorge Alonso-Guiterrez, Dr. James Carothers, Mary Agnitsch and Eric Luning for their unique contributions to my research and general well-being.

I would love to thank the entire members of Keasling lab and JBEI (past and present) for their immeasurable contributions to my scientific knowledge. My appreciation goes to Dr. Dennis Grogan for his immense contributions to my knowledge of *Sulfolobus*.

I would love to thank everyone who has contributed in any capacity to my research and well-being during the course of my study. Please forgive me if I did not mention your name; just know that I sincerely appreciate your contributions.

My study would not be possible with funding from JBEI supported by the Department of Energy, Plant and Microbial Biology and University of California, Berkeley.

Chapter One: *Sulfolobus acidocaldarius*: A Metabolic Engineering Platform for Cellulosic Biofuel Production

Introduction

The inefficient bioconversion of ligno-cellulose (cellulosic) biomass into biofuels by mesophilic microbes and overall high cost of production has increased the demand for the development of thermophilic host systems (40, 52). Cellulosic biofuels are no doubt, attractive replacements for corn-based fuels, especially considering the adverse effect of the latter on the cost of food. However, the cost of producing these biofuels accrued from the use of expensive pre-treatment step that involves extensive use of cocktails of thermophilic enzymes has impeded progress (11, 17, 28, 33, 39). The recalcitrance of cellulosic biomass and associated metabolic bottle-necks that emanate from biomass deconstruction make bioconversion with mesophiles such as *Escherichia coli* and *Saccharomyces cerevisiae*, a daunting process (28). Most microbial strains used for cellulosic biofuels production tend to ferment hexoses such as glucose better than the pentoses; despite the fact that pentoses constitute over 20 % of cellulosic biomass (14, 28). Advancement of current state of cellulosic biofuel production to a sustainable level would require a cutting-edge technology that would consolidate the deconstruction and bioconversion processes. Since most pre-treatment steps require elevated temperature, it would be logical to consider developing a thermophilic microbial host system for cellulosic biofuel production.

In this study, we propose the development of *Sulfolobus acidocaldarius*, an aerobic hyperthermophilic crenarchaea as a prospective host system for cellulosic biofuels production. *S. acidocaldarius* grows optimally at 75 - 80°C and pH 2 – 3 (6, 9) and maintains a cytosolic pH of approximately 6.5. The difference in the pH of the organism's external and internal environments creates a trans-membrane pH gradient that is exploited for energy generation and uptake/export of materials (37, 48). Growth of *S. acidocaldarius* at low pH would be advantageous to the development of the organism as a host for biofuel production if dilute-acid treatment is the method of choice for biomass deconstruction (39, 56). In addition, *S. acidocaldarius* lacks active insertion (IS) elements that are characteristic of known *Sulfolobus* species, indicating that it has the most stable genome (9, 48). The stability of the genome of *S. acidocaldarius* and its adaptation to extreme temperature and pH environment makes the organism a plausible candidate for cellulosic biofuel production.

The organism metabolizes cellulosic sugars such as glucose and xylose as sole sources of carbon and energy (16, 19, 22). Recently, we showed that *S. acidocaldarius* was capable of catabolism of different combinations glucose, xylose, arabinose and galactose without an apparent glucose-induced catabolite repression (15, 22, 34). Presence of glucose-induced catabolite repression in model organisms such as *E. coli* and *S. cerevisiae* has impacted the use of these organisms in cellulosic biofuel production, due to inefficient bioconversion of the sugars (15, 34). Numerous metabolic engineering studies are currently being directed towards developing *E. coli* and yeast strains that are capable simultaneous metabolism of multiple sugars in the presence of glucose. Therefore absence of glucose-induced catabolite repression in *S. acidocaldarius* enhances the credentials of the organism as a prospective host for cellulosic biofuel production. Simultaneous metabolism of cellulosic sugars by *S. acidocaldarius* would reduce steps involved in the bioconversion process, therefore reducing the overall cost of production. Most *Sulfolobus* species are known to harbor cellulose metabolic genes such as those

encoding endo-β-glucanases, however there has been no report of cellulose metabolism by any *Sulfolobus* species (20, 32, 36, 48).

However, the development of *S. acidocaldarius* as host for cellulosic biofuel production is not without challenges, especially being a thermophilic organism. Among these challenges is the inadequate availability of genetic and molecular biology tools. In general, there are fewer genetic tools for studying thermophilic organism compared to their mesophilic counterpart. Although numerous genetic tools have been developed for the *Sulfolobus* species over the years, very few have been independently reproduced. Development of genetic tools for *Sulfolobus* species has been plagued by limited availability of non-viral free replicating plasmids and inadequate transformation strategy (5). Effective selectable markers are not readily available for the organism; like most archaea, *Sulfolobus* species are generally known to be resistant to a wide range of antibiotics (1, 5, 8). Overall improvement of the genetic toolbox for *S. acidocaldarius* will greatly enhance its suitability as a platform for cellulosic biofuels production.

S. acidocaldarius does not metabolize cellobiose; a critical metabolite in the degradation of cellulose (28). Cellobiose is metabolized by other *Sulfolobus* species suggesting that inability to catabolize this sugar is peculiar to *S. acidocaldarius* (16, 19, 48). Although absence of cellobiose metabolism undermines the utility of *S. acidocaldarius* for engineering cellulosic biofuels; it is very plausible to install cellobiose metabolic pathway from the other *Sulfolobus* species into the organism. Metabolism of cellobiose is initiated in *S. solfataricus* mediated by the action of β-glycosidase encoded by *lacS*; inactivation of this gene in the organism consequently terminated metabolism of cellobiose and lactose (18). A homolog of *lacS* that was annotated as beta-galactosidase (*bgaS*) is present in *S. acidocaldarius,* but the organism does cleave X-gal (5-bromo-4-chloro-indolyl-β-D-galactopyranoside), a characteristic of *Sulfolobus* species. Although a recent studies showed show that BgaS was able to cleave X-gal and cellobiose *in vitro*, it failed to explain the inability to cleave X-gal *in vivo* (42). It would be important to conduct exhaustive study on cellobiose metabolism in the organism to fully understand its inability to utilize cellobiose, as numerous factors could be responsible.

The growth of *S. acidocaldarius* at elevated temperature makes it a perfect system for production of bio-ethanol (boiling point approximately 78.5°C). Current strategy for ethanol production involves production and distillation steps; with *S. acidocaldarius* these two steps could be consolidated by growing the organism at 80°C, which is above the boiling of point of ethanol. *S. acidocaldarius* does not possess a native ethanol biosynthetic pathway; but this pathway can be engineered into the organism with the availability of an effective genetic toolbox. *S. acidocaldarius* could also be employed in the biosynthesis of next generation biofuels (24), especially fuels based on fatty acids (41, 51) and isoprenoids (31, 43). To achieve the goal of employing *S. acidocaldarius* for engineering cellulosic biofuels, we have to understand physiology and metabolism in the organism. There is also the need to develop effective genetic tools and establish cellobiose metabolism.

Sugar metabolism in *Sulfolobus acidocaldarius*

Sugar transport in archaea is generally achieved via an active transport system that is mediated by ATP Binding Cassette (ABC) transporters; complete phosphotransferase systems (PTS) are seldom found in these group of organism (23, 30). ABC transporters import sugars into the organism without modification. A number of genes encoding ABC transporters have been described in *S. acidocaldarius*, but only few of these genes have been properly annotated or verified (3, 9, 48). The prevalence of ABC transporters for sugar uptake in archaea might be influenced by the nutrient-deficient environments of these organisms (6, 58). *Sulfolobus* species are known for their broad sugar metabolism (16), and recently for the absence of catabolite repression that is common among bacteria and eukaryotes (15, 22, 34).

Glycolytic pathway for sugar catabolism in archaea is usually a variant of the classical pathway in bacteria and eukaryotes (53). Glycolysis in *S. acidocaldarius* is achieved via two modified Entner-Doudoroff (ED) pathways, termed non-phosphorylating and semi-phosphorylating ED pathways (2, 49). Unlike the classical ED pathway that yields one net ATP, the modified pathways that are present in *Sulfolobus* species, generate no net ATP from the breakdown of glucose (53). The organism relies on oxidative phosphorylation via the tri-carboxylic acid (TCA) cycle to generate the necessary ATP required for cellular processes. Glycolysis in *S. acidocaldarius* starts with the oxidation of glucose to gluconate by glucose dehydrogenase and followed by dehydration of gluconate to 2-keto-3-deoxygluconate (KDG) by gluconate dehydratase (2, 12). KDG is cleaved into pyruvate and glyceraldehyde by KDG aldolase in the non-phosphorylating ED pathway (2). The glyceraldehyde moiety is subsequently converted to pyruvate via a number of steps (2, 12, 49, 53). Semi-phosphorylating ED differs from the non-phosphorylating version by the way KDG is metabolized. In the semi-phosphorylating ED pathway, KDG is phosphorylated to 2-keto-3-deoxy-6-phosphogluconate (KDPG) by KDG kinase and subsequently dissimilated into pyruvate and glyceraldehyde-3-phosphate by the promiscuous KDG aldolase or KDPG aldolase (2, 49). Glyceraldehyde-3-phosphate generated from KDPG is not converted to 1, 3-bisphospho-glycerate; rather, it is converted directly to 3-phosphoglycerate resulting in net gain in ATP (2, 49). The pyruvate generated from these modified ED pathways is converted to acetyl-CoA by ferrodoxin-dependent oxoglurate synthase, which is also involved in the conversion of α-keto-glutarate to Succinyl-CoA (49).

Embden-Meyerhof-Parnas (EMP) pathway is present in *S. acidocaldarius*, but this pathway is not involved in glycolysis because two key genes, glucokinase and phosphofructokinase are absent (48). The EMP pathway in *Sulfolobus* species is believed to be essential for gluconeogenesis (2, 48-49). Similarly, only the non-oxidative phase of pentose phosphate pathway, which plays important role in anaplerotic reactions and pentose sugar metabolism, is present in the organism. Genes for the oxidative phase of the pathway are absent in the organism. Absence of complete pentose phosphate pathway does not impart generation of NADPH because the molecule is generated in the TCA cycle by Isocitrate dehydrogenase (49). NADPH is the preferred reducing equivalent in *S. acidocaldarius* and most archaea, and it is converted to ATP via oxidative phosphorylation (48-49).

Molecular biology in *Sulfolobus acidocaldarius*

Detailed understanding of *Sulfolobus* species and hyper-thermophilic archaea in general has been limited by the availability of reproducible genetic and molecular biology tools. These genetic tools have improved from generation systems that are based on archaea viruses such as *Sulfolobus Shibatae* virus 1 (SSV1) to current system that involves the use of free replicating shuttle vectors or homologous recombination with selectable markers (5, 46, 50, 57). These systems are still inadequate and need to be improved substantially. A number of plasmid systems have been developed for *Sulfolobus* species, especially those based on pRN1 or pRN2 (4, 13, 25-26); but only few of these plasmids have been independently reproduced. Besides the low availability of plasmid system, the key challenges facing advancement of genetic tools for *Sulfolobus* is the absence of effective selectable markers and effective transformation strategy (5). The most successful selection for *Sulfolobus* species are based on uracil or lactose auxotrophies (4, 21, 29, 44-45). Antibiotics-based selection has not been very successful in *Sulfolobus* species because the organism, like most hyper-thermophilic archaea is not very susceptibility to antibiotics (1, 8). Targeted gene disruption or replacement via homologous recombination has also been demonstrated in the organism (54-55). Homologous recombination is generally important when there is the need for gene deletion or disruption. But achieving homologous this species is very challenging, therefore making the use of plasmids a more practical option for routine heterologous gene expression. Some of the most reproducible *Sulfolobus* shuttle vectors are those based pRN1 such as pC, pJ-lacS and pCmalLacS, but these shuttle vectors are usually in the excess of 8 kb (3, 6). These recent advancements in *Sulfolobus* genetic tools have increased our knowledge of functions in the organism, but there is the need for further improvements.

Biofuel potential of *Sulfolobus acidocaldarius*

The ability of *S. acidocaldarius* to withstand harsh environmental conditions such as high temperature and low pH makes it a potential platform for the production of cellulosic biofuel like bio-ethanol. The major challenge is that *S. acidocaldarius* is incapable of anaerobic growth and ethanol production; but it tolerates greater than 3% ethanol (10, 14, 16). The organism lacks genes encoding pyruvate decarboxylase, but possesses all the genes necessary for the conversion of pyruvate to ethanol via acetyl-CoA except for acetylaldehyde dehydrogenase (9, 14, 47). Bio-ethanol production might be achieved under aerobic condition by the expression thermophilic acetyyaldehyde dehydrogenase; particularly the enzyme from *Thermoanaerobacter pseudoethanolicus 39E* (7). The organism could be used for the production of other classes of biofuels, especially isoprenoid or fatty acid based fuels. The lipid in the membrane of archaea is not made of fatty-acid-based hydrocarbons as in bacteria and eukaryotes, rather they are made up of isoprenoid-based hydrocarbon (27). The isoprenoid pathway is well established in archaea and could serve as a foundation for the development of next-generation biofuels (35, 43). Development of these potential fuels could be achieved with improved genetic tools and strategy.

Research rationale and thesis organization

The study was motivated by the need for a robust microbial system that can serve as a substitute for *E. coli* and yeast as platform for cellulosic biofuel production. Most studies in cellulosic biofuel engineering have been conducted in *E. coli* and *Saccharomyces cerevisiae*, but the use of these systems has some inherent challenges. Conversion of cellulosic biomass to individual sugar components requires the use of thermophilic enzymes that are not readily expressed in the mesophilic hosts (14, 28, 39). The current high cost of producing cellulosic biofuel could be attributed to the use of multiple and expensive pre-treatment steps to release fermentable sugars, and the inefficient conversion of these sugars into fuels by fermenting organisms (39, 56). Bioconversion of most cellulosic sugars is limited by the existence of carbon catabolite repression in *E. coli* and *S. cerevisiae* (34, 38); these species mostly do not consume other sugars such as xylose in the presence of glucose because of catabolite repression. Understanding sugar metabolism in *S. acidocaldarius* would enhance its potential as a platform cellulosic biofuel engineering and metabolic engineering at large. The growth of the organism at high temperature and low pH makes it suitable for the currently established pre-treatment technology. *S.acidocaldarius* could also serve as platform for the production of thermophilic enzyme employed in the pre-treatment of cellulosic biomass.

This dissertation describes the ability of *S. acidocaldarius* to utilize a mixture of sugars in the presence of glucose without an apparent catabolite repression in Chapter 2, in contrast to previous report that glucose represses metabolism of other sugars in *Sulfolobus* species. In the process we uncovered the existence of a novel but well coordinated mechanism of simultaneous sugar metabolism. Chapter 3 describes the development of pRN1-based shuttle vector system for heterologous genes expression in *Sulfolobus* species. The chapter focuses on understanding the origin of replication of pRN1 which has remained unknown despite successful utilization of the plasmid in developing *Sufolobus* shuttle vectors. The focus of Chapter 4 is to explore the reason for absence of cellobiose metabolism in *S. acidocaldarius* and identify means to establish this metabolism that is critical for catabolism of cellulose. The chapter is also focused on developing genetic tools for studying the organism by exploring strategy for chromosomal mutagenesis and homologous recombination in the organism. The dissertation was summarized in Chapter 5; the chapter also described future direction for the study.

References

1. **Aagaard, C., H. Phan, S. Trevisanato, and R. A. Garrett.** 1994. A spontaneous point mutation in the single 23S rRNA gene of the thermophilic arachaeon *Sulfolobus acidocaldarius* confers multiple drug resistance. Journal of Bacteriology **176:**7744-7747.
2. **Ahmed, H., T. J. G. Ettema, B. Tjaden, A. C. M. Geerling, J. van der Oost, and B. Siebers.** 2005. The semi-phosphorylative Entner-Doudoroff pathway in hyperthermophilic archaea: a re-evaluation. Biochemical Journal. **390:**529-540.

3. **Albers, S.-V., S. M. Koning, W. N. Konings, and A. J. M. Driessen.** 2004. Insights into ABC Transport in Archaea. Journal of Bioenergetics and Biomembranes **36:**5-15.
4. **Berkner, S., D. Grogan, S.-V. Albers, and G. Lipps.** 2007. Small multicopy, non-integrative shuttle vectors based on the plasmid pRN1 for *Sulfolobus acidocaldarius* and *Sulfolobus solfataricus*, model organisms of the (cren-)archaea. Nucleic Acids Research. **35:**e88-.
5. **Berkner, S., and G. Lipps.** 2008. Genetic tools for *Sulfolobus* spp.: vectors and first applications. Archives of Microbiology **190:**217-230.
6. **Brock, T. D., K. M. Brock, R. T. Belly, and R. L. Weiss.** 1972. *Sulfolobus*: A new genus of sulfur-oxidizing bacteria living at low pH and high temperature. Archives of Microbiology **84:**54-68.
7. **Burdette, D., and J. G. Zeikus.** 1994. Purification of acetaldehyde dehydrogenase and alcohol dehydrogenases from *Thermoanaerobacter ethanolicus* 39E and characterization of the secondary-alcohol dehydrogenase (2° adh) as a bifunctional alcohol-dehydrogenase acetyl-coa reductive thioesterase. Biochemical Journal **302:**163-170.
8. **Cammarano, P., A. Teichner, P. Londei, M. Acca, B. Nicolaus, J. L. Sanz, and R. Amils.** 1985. Insensitivity of archaebacterial ribosomes to protein synthesis inhibitors. Evolutionary implications. The EMBO journal **4:**811-6.
9. **Chen, L., K. Brugger, M. Skovgaard, P. Redder, Q. She, E. Torarinsson, B. Greve, M. Awayez, A. Zibat, H.-P. Klenk, and R. A. Garrett.** 2005. The Genome of *Sulfolobus acidocaldarius*, a Model Organism of the Crenarchaeota. Journal of Bacteriology **187:**4992-4999.
10. **Chong, P. K., A. M. Burja, H. Radianingtyas, A. Fazeli, and P. C. Wright.** 2007. Proteome and Transcriptional Analysis of Ethanol-Grown *Sulfolobus solfataricus* P2 Reveals ADH2, a Potential Alcohol Dehydrogenase. Journal of Proteome Research **6:**3985-3994.
11. **Chundawat, S. P. S., G. T. Beckham, M. E. Himmel, and B. E. Dale.** 2011. Deconstruction of Lignocellulosic Biomass to Fuels and Chemicals. Annual Review of Chemical and Biomolecular Engineering **2:**121-145.
12. **De Rosa, M., A. Gambacorta, B. Nicolaus, P. Giardina, E. Poerio, and V. Buonocore.** 1984. Glucose metabolism in the extreme thermoacidophilic archaebacterium *Sulfolobus solfataricus*. Biochemical Journal. **224:**407-0.
13. **Deng, L., H. Zhu, Z. Chen, Y. Liang, and Q. She.** 2009. Unmarked gene deletion and host–vector system for the hyperthermophilic crenarchaeon *Sulfolobus islandicus*. Extremophiles **13:**735-746.
14. **Dien, B. S., M. A. Cotta, and T. W. Jeffries.** 2003. Bacteria engineered for fuel ethanol production: current status. Applied Microbiology and Biotechnology **63:**258-266.
15. **Gorke, B., and J. Stulke.** 2008. Carbon catabolite repression in bacteria: many ways to make the most out of nutrients. Nature Reviews Microbiology **6:**613-624.
16. **Grogan, D. W.** 1989. Phenotypic characterization of the archaebacterial genus *Sulfolobus* : comparison of five wild-type strains. Journal of Bacteriology **171:**6710 - 6719.
17. **Harvey W, B.** Bioprocessing for biofuels. Current Opinion in Biotechnology.

18. **Haseltine, C., R. Montalvo-Rodriguez, A. Carl, E. Bini, and P. Blum.** 1999. Extragenic Pleiotropic Mutations That Repress Glycosyl Hydrolase Expression in the Hyperthermophilic Archaeon *Sulfolobus solfataricus*. Genetics **152:**1353-1361.
19. **Hatzinikolaou, D. G., E. Kalogeris, P. Christakopoulos, D. Kekos, and B. J. Macris.** 2001. Comparative growth studies of the extreme thermophile *Sulfolobus acidocaldarius* in submerged and solidified substrate cultures. World Journal of Microbiology and Biotechnology **17:**229-234.
20. **Huang, Y., G. Krauss, S. Cottaz, H. Driguez, and G. Lipps.** 2005. A highly acid-stable and thermostable endo-beta-glucanase from the thermoacidophilic archaeon *Sulfolobus solfataricus*. Biochem. J. **385:**581-588.
21. **Jonuscheit, M., E. Martusewitsch, K. M. Stedman, and C. Schleper.** 2003. A reporter gene system for the hyperthermophilic archaeon *Sulfolobus solfataricus* based on a selectable and integrative shuttle vector. Molecular Microbiology **48:**1241-1252.
22. **Joshua, C. J., R. Dahl, P. I. Benke, and J. D. Keasling.** 2011. Absence of Diauxie during Simultaneous Utilization of Glucose and Xylose by *Sulfolobus acidocaldarius*. J. Bacteriol. **193:**1293-1301.
23. **Karlin, S., L. Brocchieri, A. Campbell, M. Cyert, and J. MrÃzek.** 2005. Genomic and proteomic comparisons between bacterial and archaeal genomes and related comparisons with the yeast and fly genomes. Proceedings of the National Academy of Sciences of the United States of America **102:**7309-7314.
24. **Keasling, J. D., and H. Chou.** 2008. Metabolic engineering delivers next-generation biofuels. Nat Biotech **26:**298-299.
25. **Keeling, P. J., H.-P. Klenk, R. K. Singh, O. Feeley, C. Schleper, W. Zillig, W. F. Doolittle, and C. W. Sensen.** 1996. Complete nucleotide sequence of the *Sulfolobus islandicus* multicopy plasmid pRN1. Plasmid **35:**141-144.
26. **Keeling, P. J., H. P. Klenk, R. K. Singh, M. E. Schenk, C. W. Sensen, W. Zillig, and W. F. Doolittle.** 1998. *Sulfolobus islandicus* plasmids pRN1 and pRN2 share distant but common evolutionary ancestry. Extremophiles **2:**391-393.
27. **Koga, Y., and H. Morii.** 2007. Biosynthesis of Ether-Type Polar Lipids in Archaea and Evolutionary Considerations. Microbiology and Molecular Biology Reviews **71:**97-120.
28. **Kumar, R., S. Singh, and O. Singh.** 2008. Bioconversion of lignocellulosic biomass: biochemical and molecular perspectives. Journal of Industrial Microbiology and Biotechnology **35:**377-391.
29. **Kurosawa, N., and D. W. Grogan.** 2005. Homologous recombination of exogenous DNA with the *Sulfolobus acidocaldarius* genome: Properties and uses. FEMS Microbiology Letters **253:**141-149.
30. **Lee, S.-J., A. Böhm, M. Krug, and W. Boos.** 2007. The ABC of binding-protein-dependent transport in Archaea. Trends in Microbiology **15:**389-397.
31. **Leonard, E., P. K. Ajikumar, K. Thayer, W.-H. Xiao, J. D. Mo, B. Tidor, G. Stephanopoulos, and K. L. J. Prather.** 2010. Combining metabolic and protein engineering of a terpenoid biosynthetic pathway for overproduction and selectivity control. Proceedings of the National Academy of Sciences **107:**13654-13659.

32. **Limauro, D., R. Cannio, G. Fiorentino, M. Rossi, and S. Bartolucci.** 2001. Identification and molecular characterization of an endoglucanase gene, celS, from the extremely thermophilic archaeon *Sulfolobus solfataricus*. Extremophiles **5:**213-219.
33. **Lynd, L. R., P. J. Weimer, W. H. van Zyl, and I. S. Pretorius.** 2002. Microbial Cellulose Utilization: Fundamentals and Biotechnology. Microbiol. Mol. Biol. Rev. **66:**506-577.
34. **Magasanik, B.** 1961. Catabolite repression. Cold Spring Harbor Symposia on Quantitative Biology **26:**249-56.
35. **Martin, V. J. J., D. J. Pitera, S. T. Withers, J. D. Newman, and J. D. Keasling.** 2003. Engineering a mevalonate pathway in *Escherichia coli* for production of terpenoids. Nat Biotech **21:**796-802.
36. **Maurelli, L., A. Giovane, A. Esposito, M. Moracci, I. Fiume, M. Rossi, and A. Morana.** 2008. Evidence that the xylanase activity from *Sulfolobus solfataricus* Oα is encoded by the endoglucanase precursor gene (sso1354) and characterization of the associated cellulase activity. Extremophiles **12:**689-700.
37. **Moll, R., and G. Schäfer.** 1988. Chemiosmotic H+ cycling across the plasma membrane of the thermoacidophilic archaebacterium *Sulfolobus acidocaldarius*. FEBS Letters **232:**359-363.
38. **Monod, J.** 1942. Recherches sur la croissance des cultures bactâeriennes. Hermann, Paris,.
39. **Mosier, N., C. Wyman, B. Dale, R. Elander, Y. Y. Lee, M. Holtzapple, and M. Ladisch.** 2005. Features of promising technologies for pretreatment of lignocellulosic biomass. Bioresource Technology **96:**673-686.
40. **Munasinghe, P. C., and S. K. Khanal.** 2010. Biomass-derived syngas fermentation into biofuels: Opportunities and challenges. Bioresource Technology **101:**5013-5022.
41. **Nawabi, P., S. Bauer, N. Kyrpides, and A. Lykidis.** 2011. Engineering *Escherichia coli* for Biodiesel Production Utilizing a Bacterial Fatty Acid Methyltransferase. Applied and Environmental Microbiology **77:**8052-8061.
42. **Park, A.-R., H.-J. Kim, J.-K. Lee, and D.-K. Oh.** 2010. Hydrolysis and Transglycosylation Activity of a Thermostable Recombinant β-Glycosidase from *Sulfolobus acidocaldarius*. Applied Biochemistry and Biotechnology **160:**2236-2247.
43. **Peralta-Yahya, P. P., M. Ouellet, R. Chan, A. Mukhopadhyay, J. D. Keasling, and T. S. Lee.** 2011. Identification and microbial production of a terpene-based advanced biofuel. Nat Commun **2:**483.
44. **Reilly, M. S., and D. W. Grogan.** 2001. Characterization of Intragenic Recombination in a Hyperthermophilic Archaeon via Conjugational DNA Exchange. J. Bacteriol. **183:**2943-2946.
45. **Schelert, J., V. Dixit, V. Hoang, J. Simbahan, M. Drozda, and P. Blum.** 2004. Occurrence and Characterization of Mercury Resistance in the Hyperthermophilic Archaeon *Sulfolobus solfataricus* by Use of Gene Disruption. J. Bacteriol. **186:**427-437.
46. **Schleper, C., K. Kubo, and W. Zillig.** 1992. The particle SSV1 from the extremely thermophilic archaeon *Sulfolobus* is a virus: demonstration of

infectivity and of transfection with viral DNA. Proceedings of the National Academy of Sciences of the United States of America **89:**7645-7649.

47. **Shaw, A. J., K. K. Podkaminer, S. G. Desai, J. S. Bardsley, S. R. Rogers, P. G. Thorne, D. A. Hogsett, and L. R. Lynd.** 2008. Metabolic engineering of a thermophilic bacterium to produce ethanol at high yield. Proceedings of the National Academy of Sciences **105:**13769-13774.
48. **She, Q., R. K. Singh, F. Confalonieri, Y. Zivanovic, G. Allard, M. J. Awayez, C. C. Chan-Weiher, I. G. Clausen, B. A. Curtis, and A. De Moors.** 2001. The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2. Proceedings of the National Academy of Sciences of the United States of America **98:**7835 - 7840.
49. **Snijders, A. P. L., J. Walther, S. Peter, I. Kinnman, M. G. J. d. Vos, H. J. G. v. d. Werken, S. J. J. Brouns, J. v. d. Oost, and P. C. Wright.** 2006. Reconstruction of central carbon metabolism in *Sulfolobus solfataricus* using a two-dimensional gel electrophoresis map, stable isotope labelling and DNA microarray analysis. PROTEOMICS **6:**1518-1529.
50. **Stedman, K. M., C. Schleper, E. Rumpf, and W. Zillig.** 1999. Genetic Requirements for the Function of the Archaeal Virus SSV1 in *Sulfolobus solfataricus*: Construction and Testing of Viral Shuttle Vectors. Genetics **152:**1397-1405.
51. **Steen, E. J., Y. Kang, G. Bokinsky, Z. Hu, A. Schirmer, A. McClure, S. B. del Cardayre, and J. D. Keasling.** 2010. Microbial production of fatty-acid-derived fuels and chemicals from plant biomass. Nature **463:**559-562.
52. **Taylor, M. P., K. L. Eley, S. Martin, M. I. Tuffin, S. G. Burton, and D. A. Cowan.** 2009. Thermophilic ethanologenesis: future prospects for second-generation bioethanol production. Trends in Biotechnology **27:**398-405.
53. **Verhees, C. H., S. W. M. Kengen, J. E. Tuininga, G. J. Schut, M. W. W. Adams, W. M. De Vos, and J. Van der Oost.** 2004. The unique features of glycolytic pathways in Archaea (vol 375, pg 231, 2003). Biochemical Journal **377:**819-822.
54. **Wagner, M., S. Berkner, M. Ajon, A. J. M. Driessen, G. Lipps, and S.-V. Albers.** 2009. Expanding and understanding the genetic toolbox of the hyperthermophilic genus *Sulfolobus*. Biochemical Society Transactions **37:**97-101.
55. **Worthington, P., V. Hoang, F. Perez-Pomares, and P. Blum.** 2003. Targeted Disruption of the a-Amylase Gene in the Hyperthermophilic Archaeon Sulfolobus solfataricus. J. Bacteriol. **185:**482-488.
56. **Wyman, C. E., B. E. Dale, R. T. Elander, M. Holtzapple, M. R. Ladisch, and Y. Y. Lee.** 2005. Comparative sugar recovery data from laboratory scale application of leading pretreatment technologies to corn stover. Bioresource Technology **96:**2026-2032.
57. **Yeats, S., P. McWilliam, and W. Zillig.** 1982. A plasmid in the archaebacterium *Sulfolobus acidocaldarius*. Embo Journal **1:**1035-1038.
58. **Zillig, W., K. O. Stetter, S. Wunderl, W. Schulz, H. Priess, and I. Scholz.** 1980. The Sulfolobus-"Caldariella" group: Taxonomy on the basis of the structure of DNA-dependent RNA polymerases. Archives of Microbiology **125:**259-269.

Chapter Two: Absence of Diauxie during Simultaneous Utilization of Glucose and Xylose by *Sulfolobus acidocaldarius*

This chapter was published in the Journal of Bacteriology under the reference

Joshua, C. J., R. Dahl, P. I. Benke and J. D. Keasling. 2011. Absence of Diauxie during Simultaneous Utilization of Glucose and Xylose by *Sulfolobus acidocaldarius* . *Journal of Bacteriology*. 193:1293 - 1301

ABSTRACT

Sulfolobus acidocaldarius utilizes glucose and xylose as sole carbon sources, but its ability to metabolize these sugars simultaneously is not known. We report the absence of diauxie during growth of *S. acidocaldarius* on glucose and xylose as co-carbon sources. The presence of glucose did not repress xylose utilization. The organism utilized a mixture of 1 g/L of each sugar simultaneously with a specific growth rate of 0.079 h^{-1} and showed no preference for the order with which it utilized each sugar. The organism grew faster on 2 g/L xylose (0.074 h^{-1}) as a sole carbon source than on an equal amount of glucose (0.022 h^{-1}). When grown on a mixture of the two carbon sources, the growth rate of the organism increased from 0.052 h^{-1} to 0.085 h^{-1} as the ratio of xylose to glucose increased from 0.25 to 4. *S. acidocaldarius* appeared to utilize a mixture of glucose and xylose at a rate roughly proportional to their concentration in the medium, resulting in complete utilization of both sugars at about the same time. Gene expression in cells grown on xylose alone was very similar to that in cells grown on a mixture of xylose and glucose and substantially different from that in cells grown on glucose alone. The mechanism by which the organism utilized a mixture of sugars has yet to be elucidated.

INTRODUCTION

Sulfolobus acidocaldarius is a hyperthermophilic archaeon that grows optimally at 75ºC and pH 2.0 – 3.0 (10, 20, 24) and has been shown to utilize a broad range of sugars (20, 24). Numerous studies have shown that bacteria as well as eukaryotes sequentially utilize individual sugars when grown on a mixture of sugars. These organisms preferentially utilize the sugar that best supports their growth (mostly glucose) by repressing the utilization of others sugars in the growth medium until the preferential sugar is completely consumed. This phenomenon termed "carbon catabolite repression (CCR)" or "diauxie" is characterized by a di-phasic growth pattern when an organism is grown on a mixture of glucose and other sugars. Most studies on CCR have focused mainly on bacterial or eukaryotic systems. However, a few studies have reported that the presence of glucose in a growth medium represses the metabolism of other sugars in species closely related to *S. solfataricus* via a mechanism that is similar to CCR (22-23, 25, 34).

S. acidocaldarius metabolizes the smallest number of sugars of all known *Sulfolobus* species; but the sugars metabolized by the organism include glucose and xylose, the key constituents of ligno-cellulose (cellulosic) biomass (20, 24, 37). The lack of diauxie on 5- and 6-carbon sugars and the ability to grow at high temperature and low pH are excellent characteristics for a host that produces biofuel from cellulosic biomass deconstructed using acidic and/or high temperature pre-treatment methods. The production of biofuels from cellulosic biomass as an alternative to fossil fuels has received increased attention in the past few years. But the development of microbial system(s) that can efficiently and simultaneously convert glucose and xylose into biofuel has been challenging. It is not known if *S. acidocaldarius* can utilize glucose and xylose (or other sugars) simultaneously. The ability of the organism to simultaneously utilize glucose and xylose would reduce the cost of biofuel production by eliminating the need for separating the 5- and 6-carbon sugars during pre-treatment, and thus shortening the time required for fermentation of both sugars. *Saccharomyces cerevisiae* and *Escherichia coli*, two model systems for metabolic engineering, vary in their abilities to utilize glucose and xylose. *S. cerevisiae* lacks a native mechanism for utilizing xylose (26), while *E. coli* utilizes both glucose and xylose but not simultaneously due to CCR (36).

Central to CCR in bacteria is the multi-component phospho-transferase system (PTS) that mediates global or operon-specific regulation of the process (14, 19). Unlike in bacteria, PTS systems are apparently absent in most if not all archaea, and sugar transport is primarily mediated by ATP-binding cassette (ABC) transporters (5, 28, 32, 42). However, reports of the existence of a CCR-like system in *S. solfataricus* (22-23, 25, 34) might indicate that the organism would be able to sequentially metabolize a mixture of sugars in their order of preference. The absence of PTS systems in *Sulfolobus* species (10, 43) suggests that the mechanism of CCR in these organisms might be different from the one observed in bacteria. Our goal in this study was to determine if *S. acidocaldarius* can simultaneously utilize glucose, xylose, and other sugars that might be derived from cellulosic biomass in an effort to develop a system for cellulosic biofuel production. The genome of *S. acidocaldarius* is relatively stable because of the absence of active transposable elements that characterize other species of *Sulfolobus* (10, 43), therefore making it a suitable system for engineering. In addition, a number of genetic tools are

available for manipulating the genome of this organism (6, 30). As such, *S. acidocaldarius* may find use in the production of advanced biofuel.

In this study, we report experiments describing growth of *S. acidocaldarius* on different combinations of glucose, xylose, arabinose, and galactose. We also conducted micro-array studies to determine gene expression patterns during growth on glucose and xylose as sole and co-carbon sources. We present evidence that *S. acidocaldarius* utilizes various mixtures of glucose, xylose, arabinose, or galactose simultaneously and synergistically as co-carbon sources and showed no preference for any of the sugars. Our results indicate the absence of diauxie during growth of this organism on multiple sugars and that the regulation of multiple sugar metabolism in the organism is different from the bacterial CCR. To our knowledge, this is the first report of such behavior in a *Sulfolobus* species or any archaea. We also identified a cluster of ABC-transport genes that are likely involved in glucose transport.

MATERIALS AND METHODS

Strains and growth media. *Sulfolobus acidocaldarius* (DSM 639) was obtained from Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ – Braunschweig, Germany). The cells were grown aerobically at 75ºC in ATCC #1723 (YT) medium, which contained the following components (per liter): 1.3 g $(NH_4)_2SO_4$, 0.28 g $KH_2.PO_4$, 0.25 g $MgSO_4{\cdot}7H_2O$, 0.07 g $CaCl_2{\cdot}2H_2O$, 0.02 g $FeCl_3{\cdot}6H_2O$, 1.8 mg $MnCl_2 \cdot 4H_2O$, 4.5 mg $Na_2B_4O_7$, and 1 ml of a 1000X stock solution of trace elements containing (per liter): 220 mg $ZnSO_4{\cdot}7H_2O$, 50 mg $CuCl_2{\cdot}2H_2O$, 30 mg $Na_2MoO_4{\cdot}2H_2O$, 30 mg $VOSO_4{\cdot}3\text{-}5H_2O$ and 10 mg $CoSo_4{\cdot}7H_2O$. The pH of the medium was adjusted to 2.5 with 5 M H_2SO_4 and filter sterilized with 0.2-µM pore size membrane filters. Sugar utilization assays were carried out in minimal medium (NYT) that was composed of ATCC #1723 medium without yeast extract and tryptone. The sugars used in these studies were aseptically added to the desired final concentrations as sterile 5 – 40% stock solutions. Growth was monitored by measuring optical density at 600 nm (OD_{600}) using a Beckman-Coulter DU-800 UV/Visible spectrophotometer. Sugars and other chemicals used in this study were obtained from Sigma and Fisher scientific.

***S. acidocaldarius* adaptation to growth on sugars as sole carbon sources.** *S. acidocaldarius* 639 was adapted to growth on each sugar as sole source of carbon and energy by repeated sub-culturing in YT medium supplemented with 2 g/L of each sugar. The cells were then transferred to minimal medium (NYT) supplemented with 2 g/L each sugar once the OD_{600} value in the YT medium containing 2 g/L of the sugar exceeded 1.0. The maximum attainable OD_{600} on YT medium alone was about 0.6 – 0.8. The adapted cells were subsequently used to study sugar utilization in minimal medium.

Co-utilization of glucose and xylose by *S. acidocaldarius*. The organism was grown to late-exponential or stationary growth phase in minimal medium supplemented with 2 g/L each of glucose and xylose. The cells were harvested by centrifugation at 6,000 × g for 5 min and washed twice with 50 ml NYT medium. The cell pellets were re-suspended in the same medium and OD_{600} was measured to determine the volume of cell suspension required to obtain an initial OD_{600} of 0.04 – 0.07 in the final medium. The suspended cells were inoculated in triplicate into 50 ml NYT medium in 250-ml screw-cap flasks that were supplemented with various ratios of glucose and xylose. In one set of experiments, the concentration of xylose was kept constant at 1 g/L while the glucose

concentration was varied by 1, 2, and 4 g/L. In the second set, the glucose concentration was kept constant at 1 g/L while xylose concentration was varied by 2 and 4 g/L. Two sets of experiments in triplicate containing 2 g/L of each sugar as a sole carbon source served as controls. At various intervals, 1.0-ml samples were withdrawn to measure growth (OD_{600}). Second 1.0-ml samples were withdrawn and pelleted by centrifugation at 17,900 × g for 1 min, and the supernatants were filtered through 0.2-µm syringe filters and analyzed for residual sugars using an Agilent 1100 Series Binary LC System equipped with an Agilent 1200 series refractive index detector. The samples were eluted isocratically using Bio-Rad Aminex HPX-87H 300x7.8 column (with Micro-Guard Cation H Refill Cartridges, 30 x 4.6 mm) with a mobile phase consisting of 4 mM H_2SO_4 in HPLC grade purity water (Honeywell Burdick & Jackson, Morristown, NJ). The flow-rate and temperature were maintained at 0.6 ml/min and 50°C, respectively.

Analysis of gene expression in *S. acidocaldarius* grown on glucose and xylose as co-carbon sources. *S. acidocaldarius* was inoculated in triplicate into 70 ml NYT medium supplemented with 4 g/L xylose, 4 g/L glucose, or a mixture 2 g/L xylose and 2 g/L glucose, respectively. Prior to inoculation, the cells were pre-adapted to each experimental condition by sub-culturing on the same sugar(s) at least three times. The total RNA was extracted from 50 ml of each culture at mid-exponential phase (OD_{600} = 0.2 – 0.4) using RNeasy mini kit (Qiagen). The cells were pelleted at 8,000 × g for 5 min at 4ºC and then re-suspended in 0.5 ml of NYT medium. 1.0 ml of RNA*later* (Qiagen) was added to the cell suspension to stabilize the RNA transcripts. The cell suspension was mixed briefly for 5 seconds and incubated at room temperature for 5 min. Each stabilized cell suspension was pelleted in 2-ml tubes at 5,000 × g for 5 min at room temperature and then re-suspended in 200 µl of TE buffer containing 15 mg/ml lysozyme and 20 µl of a 20 mg/ml solution of proteinase K (Qiagen). The total RNA was then extracted as described by the manufacturer. Genomic DNA was removed from each sample using an on-column DNaseI protocol (Qiagen) as described by the manufacturer.

cDNA synthesis and micro-array analysis. The first strand of cDNA was synthesized from each extracted total RNA sample using SuperScript™ Plus Indirect cDNA labeling kit (Invitrogen). Approximately 0.5 µg of random hexamer primers was added to 17 – 22 µg of total RNA, and the annealing reaction volume adjusted to 18 µl with sterile RNase-free water. The mixture was incubated at 70°C for 5 min and then cooled on ice for at least 1 min. The following reagents were added to the annealing reaction to a final volume of 30 µl: 5X First-Strand buffer, 0.15 M DTT, dNTP mix including aminoallyl and aminohexyl-modified nucleotides, 40 U RNaseOUT™ and 800 U SuperScript™ III Reverse Transcriptase (RT). The reaction mixtures were incubated at 46°C for 3 hours, and the synthesized first-strand cDNA was hydrolyzed and purified as described by the manufacturer. The purified cDNA was dried in a speedvac to 5 µl, and labeled with Alexa Fluor 555 dye (Invitrogen) in 10-µl reactions for 2 hours. The labeled cDNA was purified on QiaQuick PCR clean-up columns, eluted 2 times into 30 µl DEPC-water, and dried in a speedvac. The samples were processed according to Nimblegen protocols and hybridized on four-plex arrays at 42°C for 17 hours.

Microarray data analysis. Arrays were scanned using a GenePix 4200A scanner (Molecular Devices) and processed with NimbleScan software. NimbleScan was used to create pair reports and normalize the data using quantile normalization and robust multi-array analysis (RMA) (8). Normalized data were analyzed and clustered using Genesis

(45). Data are stored at gene expression omnibus (GEO) under Accession number GSE16973 and at microbesonline.org (http://microbesonline.org/cgi-bin/microarray/viewExp.cgi?expId=1723). Comparative genomic and protein sequence analyses were carried out using www.microbesonline.org, http://www.ncbi.nlm.nih.gov. Protein sequences were aligned using ClustalW2 (http://www.ebi.ac.uk/Tools/clustalw2/index.html) and the alignments were visualized using Jalview (11, 48).

RESULTS

Growth of *S. acidocaldarius* on sugars as exclusive carbon sources. We analyzed the ability of *S. acidocaldarius* to utilize glucose, xylose, and arabinose as sole carbon sources. The organism grew faster on xylose and arabinose as sole carbon sources with growth rates of 0.070 and 0.082 h^{-1}, respectively, compared to 0.041 h^{-1} on glucose (Table 1) in agreement with previous reports (24). Initial adaptation of the original strain of *S. acidocaldarius* (from DSMZ) to glucose did not result in any significant growth on xylose or arabinose as sole carbon sources (Supplementary Figure 1). The glucose-adapted strain was able to grow immediately on sucrose and galactose as sole carbon sources (Supplementary Figure 1), even though studies have previously shown that the organism does not metabolize galactose (20, 24). Growth of the organism on xylose or arabinose as sole carbon sources was established by repeatedly sub-culturing the organism in complex medium supplemented with either xylose or arabinose. Similar behavior was also observed with *S. solfataricus*, suggesting that this adaptative behavior might be common among *Sulfolobus* species. Efforts to adapt the organism to cellobiose and lactose were not successful, in agreement with previous reports (10, 20, 24).

No diauxie during growth on a mixture of glucose and xylose. To determine the existence of glucose-induced CCR in *S. acidocaldarius* that is similar to a previous report of its existence in *S. solfataricus* (7, 23, 34), we conducted classical diauxie growth experiments (36). Our results revealed the absence of diauxie during growth of *S. acidocaldarius* on different combinations of glucose and xylose (Figure 1). The organism utilized both sugars simultaneously without an apparent di-phasic growth pattern (Figure 1). The initial ratios of the sugars in the growth media were maintained during growth of the organism on both sugars (Supplementary Figure 2) and in most cases, resulting in the complete utilization of the sugars at about the same time (Figure 1). The growth rate of the organism on a mixture of equal amounts of glucose and xylose (1 g/L each) was nearly four-fold faster than on 2 g/L glucose alone and slightly slower than on 2 g/L xylose alone (Figure 1). Increases in the ratio of xylose to glucose from 0.25 to 4.0 increased the growth rate of the organism from 0.052 to 0.085 h^{-1} (Figure 2).

The growth rate of the organism on a mixture of 4 g/L glucose and 1 g/L xylose (0.052 h^{-1}) was more than two-fold greater than on 2 g/L glucose alone (0.022 h^{-1}). We also observed that the organism stopped utilizing a mixture of 4 g/L glucose and 1 g/L xylose after a period of time (Figure 1C), even though it simultaneously metabolized both sugars at a ratio roughly proportional to their starting concentrations. The cause of this growth cessation is not clear at the moment, but this growth behavior was always observed when the glucose to xylose ratio was greater than two. This cessation in growth was not observed when the organism was grown on 4 g/L glucose alone (Supplementary Figure 3), indicating that the high concentration of glucose was not the cause. In addition,

changes in the ratio of glucose and xylose did not significantly impact the yield coefficient and the specific consumption rates of the sugars (Figure 2). The synergism observed during growth on a mixture of glucose and xylose contrasts the growth pattern expected if diauxie occurred (33, 35), thus indicating that none of these sugars inhibits utilization of the other.

***S. acidocaldarius* also metabolizes glucose, galactose, and arabinose simultaneously.** We evaluated the ability of *S. acidocaldarius* to simultaneously and synergistically utilize other sugars (pentoses and hexoses) in the presence of glucose to determine whether the growth effect that we observed was specific to glucose and xylose alone. The starter culture was grown in minimal medium supplemented with a mixture of D-glucose, L-arabinose, and D-galactose prior to inoculation. The pre-adapted cells were subsequently grown in minimal medium supplemented with a total of 3.0 g/L of sugars: 3.0 g/L of each sugar as sole carbon sources, 1.5 g/L of each sugar in binary combinations, and 1.0 g/L of each sugar in trinary combinations. In all cases, the organism utilized the sugars (including galactose) simultaneously (Figure 3). Irrespective of their combinations, the organism completely utilized different mixtures of the sugars at about the same time, thus confirming the absence of glucose-induced diauxie. The addition of equal amounts of arabinose to a growth medium containing glucose or galactose reduced the time taken by the organism to consume the total sugars by more than 50% (Figure 3C-D). The organism failed to completely utilize arabinose as a sole carbon source in this experiment (Figure 3A), possibly due to repeated sub-culturing of the cells on glucose and galactose to improve the utilization of these sugars. We observed throughout our studies that repeated growth of *S. acidocaldarius* on one sugar for an extended period of time gradually decreased its ability to utilize other sugars. The decreased ability of the organism to utilize any of these other sugars was temporary and was easily restored by simply re-adapting it to the same sugar in a complex medium. Despite the observed weakness in arabinose metabolism, the organism was able to efficiently utilize arabinose in the presence of glucose and/or galactose (Figure 3C-E), further highlighting the absence of diauxie and synergistic metabolism of these sugars.

Gene expression during growth on various carbon sources. To understand the regulation of simultaneous sugar catabolism in *S. acidocaldarius*, we analyzed gene expression at mid-exponential growth phase in cells that were grown on glucose and xylose either as sole or co-carbon sources. The results showed that approximately 843 genes (38%) did not change significantly ($\leq$ 1.5-fold; $p > 0.05$) under any of the growth conditions studied. However, about 465 genes (21%) were differentially expressed by two-fold or more ($p \leq 0.05$) in at least one of the growth conditions (Figure 4), of which a total of 383 genes changed by two-fold or more ($p \leq 0.05$) between cells that were grown on glucose and xylose as sole carbon sources (Figure 4). Surprisingly, the gene expression pattern in cells that were grown on both sugars was closer to that of xylose-grown cells than to that of cells grown on glucose. Only 63 genes changed by two-fold or more ($p \leq 0.05$) between cells grown on both sugars and those grown on xylose alone, as compared to a total of 312 genes differentially expressed genes between glucose-grown cells and those grown on both sugars (Figure 4).

We expected that genes involved in simultaneous metabolism of glucose and xylose would be differentially expressed in cells grown on both sugars, but remain relatively unchanged in cells grown on either sugar as a sole carbon source. But we were

surprised to see that of the genes whose expression was relatively unchanged in cells grown on either sugar alone only ten of these genes changed significantly by a $\log_2$ ratio of 1.0 and above ($p \leq 0.05$) when the cells were grown on both sugars as co-carbon sources (Table 2). Most of these differentially expressed genes encoded for hypothetical genes that were mainly conserved in *Sulfolobus* species or archaea. Only one of the ten genes, Saci_0385, was down-regulated in cells that were grown on both sugars; the others were up-regulated. Saci_0385 encodes for 3-hydroxy-isobutyrate (HIB) dehydrogenase, which is involved in valine degradation (38, 41). The annotated, up-regulated genes include Saci_1497 encoding endonuclease III, which is involved in DNA repair (15), and Saci_1494 encoding a type II/IV secretion system protein suspected to be involved in archaeal flagella synthesis (Table 2). These differentially expressed genes provided little information about the likely mechanism or regulation of simultaneous metabolism of glucose and xylose.

Furthermore, about 85% of the genes in *S. acidocaldarius* (1891 genes) did not change significantly (≤ 1.5-fold; $p > 0.05$) between xylose-grown cells and those grown on both sugars, whereas 51% of the genes (1132 genes) remained unchanged between glucose-grown cells and those grown on both sugars (Figure 5), suggesting similarity in the physiology of xylose-grown cells and those grown on a mixture of glucose and xylose. This physiological similarity is further highlighted by the similar growth patterns observed when cells that were pre-adapted to either xylose or a mixture of glucose and xylose were grown on each sugar as a sole carbon source (Supplementary Figure 5). This observation suggests that xylose influences simultaneous metabolism of both sugars than glucose, but the mechanism and the extent of this influence remained to be elucidated. Most house-keeping genes—especially the chaperonins and those involved in the tricarboxylic acid (TCA) cycle, replication, and translation—remained fairly unchanged under any of the conditions analyzed (Supplementary Table 1). Saci_2032 and 2033, which encode for glycerol-3-phosphate dehydrogenase and glycerol kinase respectively, were always up-regulated in the presence of glucose, whereas their paralogs (Saci_1117 and 1118) did not respond to the presence of glucose in the growth medium and were induced by xylose (Table 4). These genes are involved in the metabolism of glycerol and glyceraldehyde via the glycerol-3-phosphate pathway (21).

Identification of a putative glucose ABC-transport system. Transporters for glucose and xylose have been identified in *S. solfataricus* (2, 16, 43), but there has been no report of transporters for these sugars in *S. acidocaldarius* (10). We compared gene expression in cells grown on glucose against that in cells grown on xylose as sole carbon sources to elucidate the sugar transporters. This comparison suggests that Saci_1163 – 1166, previously annotated as a maltose transporter (10), may be the ABC transporter for glucose. This cluster of genes was always up-regulated by the presence of glucose in the growth medium (Figure 6A) and shares about 14 – 38% homology with the amino acid sequence of Sso2847 – 2850 (GlcSTUV), the glucose ABC transporter in *S. solfataricus* (Supplementary Figure 6). Sso2847 – 2850 and Saci_1163 – 1166 differ mostly in the order and direction of some of the genes (Figure 6B). We propose that Saci_1163 – 1164 encode for the putative trans-membrane permeases (GlcTU, respectively), We also propose that Saci_1165 encodes for the sugar-binding protein (GlcS) and Saci_1166 encodes for the ATPase component (GlcV). The permeases (GlcTU) were expressed at a lower level in cells grown on glucose and xylose as co-carbon sources compared to cells

grown on glucose alone, while the expression levels of the genes encoding the ATPase and sugar-binding protein (GlcSV) under both growth conditions were only slightly altered (Figure 6A). It was difficult to identify a putative xylose transporter, but our results suggest that Saci_0880 – 0883 might be the putative xylose transporter. The ABC-transport gene-cluster appeared to be up-regulated in response to xylose (Table 3). Saci_0883 was previously annotated as a permease component of an ammonia ABC-transporter (10). Other genes such as Saci_0946 and 1707, which encode different components of an ABC transporter were also up-regulated by approximately two-fold in response to xylose (Table 3). Some putative permeases, such as Saci_1731, belonging to the major facilitator superfamily (MFS) were also up-regulated during the growth on xylose (Table 3). Up-regulation of components of ABC and MFS transport systems could be an indication that either or both systems are involved in xylose transport in *S. acidocaldarius*.

DISCUSSION

In this study, we report the absence of glucose-induced diauxie (17, 33, 36) in *S. acidocaldarius*. The organism utilized different combinations of glucose, xylose, arabinose and galactose simultaneously, while maintaining the initial ratios of the sugars in the growth media. Simultaneous metabolism of these sugars indicates that this organism lacks a glucose-induced CCR-like system that was previously reported in *S. solfataricus* (7, 23, 34). The organism grew faster on xylose as sole carbon sources than on glucose alone, but glucose utilization was improved by the presence of xylose or arabinose in the growth medium (Table 1). Contrary to the report of glucose-induced inhibition of arabinose uptake in *S. solfataricus* (34), uptake and metabolism of arabinose by *S. acidocaldarius* was actually improved by the presence of glucose in the growth medium (Figure 3). Bacterial catabolite repression is usually induced by the sugar that best supports the growth of an organism as sole carbon source; this sugar normally represses utilization of the other sugars (35). In *S. acidocaldarius*, more rapidly metabolized xylose or arabinose did not inhibit glucose metabolism; different combinations of these sugars were utilized simultaneously suggesting that the organism does not prefer any of these sugars over the others. The absence of a bacterial CCR mechanism in this archaeon might be primarily due to the absence of a PTS system (5, 42). Sugar transport in archaea is generally mediated by ABC transport systems (28, 32). The transporters of glucose, xylose, and arabinose have been identified in *S. solfataricus* (2-3, 16), but have not been characterized in *S. acidocaldarius* (10).

To our knowledge this is the first report of simultaneous sugar metabolism in any *Sulfolobus* species. Similar metabolisms have also been reported in *Corynebacterium glutamicum* (18, 50) and *Caldicellulosiruptor saccharolyticus* (47). *C. glutamicum* utilized glucose and carbon sources such as gluconate (18, 31) or acetate (50) simultaneously without an apparent diauxie; in other instances, glucose inhibited the metabolism of carbon sources like ethanol (4) or glutamate (29) suggesting the existence of more than one regulatory mechanism. Similarly, the anaerobic hyperthermophilic bacterium, *C. saccharolyticus*, utilized multiple sugars independent of each other and in particular different combinations of hexose and pentose sugars with no evidence of catabolite repression (27, 46-47). While the simultaneous utilization of multiple sugars is ideal for biofuel production from cellulosic biomass, the very long doubling time of this

organism would seem to be an impediment to its industrial application. The relatively shorter doubling time of *S. acidocaldarius* and its ability to grow on multiple sugars simultaneously make this organism a good candidate for biofuel production.

We evaluated gene expression during growth of *S. acidocaldarius* on glucose and xylose to elucidate the mechanism and possible regulation of simultaneous sugar metabolism in the organism. Only few genes were specifically expressed in response to growth of the organism on glucose and xylose as co-carbon sources at mid-exponential growth phase (Table 2). Of these few differentially express genes, only Saci_0385 was down-regulated. The others that were up-regulated mostly encoded hypothetical proteins, many of which are conserved only in *Sulfolobus* species or archaea. Some of the few annotated genes encoded for proteins like endonuclease III (Saci_1497), which is involved in DNA repair (15), and typeII/IV secretion system protein (Saci_1494) that is likely involved in flagella biosynthesis. The relevance of these proteins to simultaneous catabolism remains unknown and the absence of well annotated orthologs of the differentially expressed genes made it difficult to assign function to them.

The microarray data revealed similarity in gene expression between cells that were grown on xylose alone and those grown on a mixture of glucose and xylose, suggesting physiological similarity under these growth conditions. This similarity was further confirmed by similar growth patterns observed when cells that were pre-adapted to either xylose alone or a mixture of glucose and xylose were grown on each sugar as a sole carbon source (Figure 5B). This observed physiological similarity indicates that xylose strongly influences simultaneous metabolism of a mixture of glucose and xylose. The pathway for xylose metabolism in *Sulfolobus* species was only recently uncovered by the work of Nunn, *et. al.*, (40) in which they revealed that *Sulfolobus* species metabolize pentoses by splitting them into glycolaldehyde and pyruvate via Dahms pathway (1, 12-13, 44), in addition to the Weimberg pathway (49) that was previously described by Brouns *et al.* (9). Simultaneous operation of these two pentose pathways, however, does not explain the observed rapid growth of *S. acidocaldarius* on the pentoses. We also observed the up-regulation of genes that encoded for some components of the glycerol-3-phosphate shuttle pathway (21) in response to glucose. The expression of these glycerol-3-phosphate shuttle pathway genes was likely induced by the glyceraldehyde (or glyceraldehyde-3-phosphate) produced during glycolysis. This pathway might be actively involved in glucose metabolism alongside the reported non- and semi-phosphorylated Entner-Doudoroff pathway (1, 13, 44) or in generating glycerol for the synthesis of the organism's tetraether cell membrane (39); the extent of this involvement is not clear.

Although it has long been established that *S. acidocaldarius* metabolizes both glucose and xylose (20), very little is known about how these sugars are transported into the cell (10). The ABC transporters for most sugars have been identified and characterized in *S. solfataricus* (2-3, 16). We were able to identify Saci_1163 – 1166 as a putative glucose ABC transport cluster because these genes were always up-regulated in the presence of glucose (Figure 6). Saci_1163 – 1166 were previously annotated as components of a maltose ABC-transporter (10). We suggest that Saci_0880 – 0883, a complete cluster of an ABC transport system that was previously annotated as a putative ammonia ABC-transport, might be involved in xylose transport. Other genes encoding for individual components of ABC transporter systems and permeases belonging to the

major facilitator superfamily were also up-regulated in response to xylose, suggesting possible involvement of either or both of ABC and MFS transporters (Table 3).

We have demonstrated in this study that *S. acidocaldarius* metabolizes multiple sugars simultaneously without apparent diauxie, thus suggesting the absence of glucose-induced CCR as previously reported. Although, the ecological and evolutionary significance of this mechanism is not very clear at the moment, we believe that this mechanism evolved as an adaptive mechanism for survival in a nutrient-poor environment. The ability of *S. acidocaldarius* to simultaneously utilize five- and six-carbon sugars has important implications for cellulosic biofuel production by eliminating the need for engineering special pathways or regulation for simultaneous metabolism of both sugars. *S. acidocaldarius* has long been believed to have the least metabolic capabilities of all *Sulfolobus* species, but our observation that this species does metabolize galactose strongly suggests that we still do not know much about this organism. Further studies need to be carried out to determine the mechanism and regulation of simultaneous metabolism of multiple sugars in *S. acidocaldarius* as well as the cause of growth cessation at high glucose concentration relative to xylose.

Acknowledgments

We thank Edward Baidoo and J. L. Fortman for helpful discussions. This work was part of the DOE Joint BioEnergy Institute (http://www.jbei.org) supported by the U. S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U. S. Department of Energy.

REFERENCES

1. **Ahmed, H., T. J. G. Ettema, B. Tjaden, A. C. M. Geerling, J. van der Oost, and B. Siebers.** 2005. The semi-phosphorylative Entner-Doudoroff pathway in hyperthermophilic archaea: a re-evaluation. Biochemical Journal. **390:**529-540.
2. **Albers, S.-V., M. G. L. Elferink, R. L. Charlebois, C. W. Sensen, A. J. M. Driessen, and W. N. Konings.** 1999. Glucose Transport in the Extremely Thermoacidophilic *Sulfolobus solfataricus* Involves a High-Affinity Membrane-Integrated Binding Protein. Journal of Bacteriology **181:**4285-4291.
3. **Albers, S.-V., S. M. Koning, W. N. Konings, and A. J. M. Driessen.** 2004. Insights into ABC Transport in Archaea. Journal of Bioenergetics and Biomembranes **36:**5-15.
4. **Arndt, A., M. Auchter, T. Ishige, V. F. Wendisch, and B. J. Eikmanns.** 2008. Ethanol catabolism in *Corynebacterium glutamicum*. Journal of Molecular Microbiology and Biotechnology **15:**222-33.
5. **Barabote, R. D., and M. H. Saier, Jr.** 2005. Comparative Genomic Analyses of the Bacterial Phosphotransferase System. Microbiology and Molecular Biology Review. **69:**608-634.
6. **Berkner, S., D. Grogan, S.-V. Albers, and G. Lipps.** 2007. Small multicopy, non-integrative shuttle vectors based on the plasmid pRN1 for *Sulfolobus*

acidocaldarius and *Sulfolobus solfataricus*, model organisms of the (cren-)archaea. Nucleic Acids Research. **35:**e88-.

7. **Bini, E., and P. Blum.** 2001. Archaeal catabolite repression: a gene regulatory paradigm. Advances in Applied Microbiology **50:**339-66.
8. **Bolstad, B. M., R. A. Irizarry, M. Åstrand, and T. P. Speed.** 2003. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics **19:**185-193.
9. **Brouns, S. J. J., J. Walther, A. P. L. Snijders, H. J. G. van de Werken, H. L. D. M. Willemen, P. Worm, M. G. J. de Vos, A. Andersson, M. Lundgren, H. F. M. Mazon, R. H. H. van den Heuvel, P. Nilsson, L. Salmon, W. M. de Vos, P. C. Wright, R. Bernander, and J. van der Oost.** 2006. Identification of the Missing Links in Prokaryotic Pentose Oxidation Pathways: evidence for enzyme recruitment. Journal of Biological Chemistry **281:**27378-27388.
10. **Chen, L., K. Brugger, M. Skovgaard, P. Redder, Q. She, E. Torarinsson, B. Greve, M. Awayez, A. Zibat, H.-P. Klenk, and R. A. Garrett.** 2005. The Genome of *Sulfolobus acidocaldarius*, a Model Organism of the Crenarchaeota. Journal of Bacteriology **187:**4992-4999.
11. **Clamp, M., J. Cuff, S. M. Searle, and G. J. Barton.** 2004. The Jalview Java alignment editor. Bioinformatics **20:**426-427.
12. **Dahms, A. S.** 1974. 3-Deoxy-D-pentulosonic acid aldolase and its role in a new pathway of D-xylose degradation. Biochemical and Biophysical Research Communications **60:**1433-1439.
13. **De Rosa, M., A. Gambacorta, B. Nicolaus, P. Giardina, E. Poerio, and V. Buonocore.** 1984. Glucose metabolism in the extreme thermoacidophilic archaebacterium *Sulfolobus solfataricus*. Biochemical Journal. **224:**407-0.
14. **Deutscher, J.** 2008. The mechanisms of carbon catabolite repression in bacteria. Current Opinion in Microbiology **11:**87-93.
15. **Eide, L., E. Fosberg, E. Hoff, and E. Seeberg.** 2001. Overexpression of endonuclease III protects *Escherichia coli* mutants defective in alkylation repair against lethal effects of methylmethanesulphonate. FEBS Letters **491:**59-62.
16. **Elferink, M. G. L., S.-V. Albers, W. N. Konings, and A. J. M. Driessen.** 2001. Sugar transport in *Sulfolobus solfataricus* is mediated by two families of binding protein-dependent ABC transporters. Molecular Microbiology **39:**1494-1503.
17. **Englesberg, E.** 1959. Glucose Inhibition and the Diauxie Phenomenon. Proceedings of the National Academy of Sciences of the United States of America **45:**1494-507.
18. **Frunzke, J., V. Engels, S. Hasenbein, C. Gatgens, and M. Bott.** 2008. Co-ordinated regulation of gluconate catabolism and glucose uptake in *Corynebacterium glutamicum* by two functionally equivalent transcriptional regulators, GntR1 and GntR2. Molecular Microbiology **67:**305-22.
19. **Gorke, B., and J. Stulke.** 2008. Carbon catabolite repression in bacteria: many ways to make the most out of nutrients. Nature Reviews Microbiology **6:**613-624.
20. **Grogan, D. W.** 1989. Phenotypic characterization of the archaebacterial genus *Sulfolobus* : comparison of five wild-type strains. Journal of Bacteriology **171:**6710 - 6719.

21. **Hagopian, K., J. J. Ramsey, and R. Weindruch.** 2008. Enzymes of glycerol and glyceraldehyde metabolism in mouse liver: effects of caloric restriction and age on activities. Biosience Reports **028:**107-115.
22. **Haseltine, C., R. Montalvo-Rodriguez, A. Carl, E. Bini, and P. Blum.** 1999. Extragenic Pleiotropic Mutations That Repress Glycosyl Hydrolase Expression in the Hyperthermophilic Archaeon *Sulfolobus solfataricus*. Genetics **152:**1353-1361.
23. **Haseltine, C., M. Rolfsmeier, and P. Blum.** 1996. The glucose effect and regulation of alpha-amylase synthesis in the hyperthermophilic archaeon *Sulfolobus solfataricus*. Journal of Bacteriology **178:**945-50.
24. **Hatzinikolaou, D. G., E. Kalogeris, P. Christakopoulos, D. Kekos, and B. J. Macris.** 2001. Comparative growth studies of the extreme thermophile *Sulfolobus acidocaldarius* in submerged and solidified substrate cultures. World Journal of Microbiology and Biotechnology **17:**229-234.
25. **Hoang, V., E. Bini, V. Dixit, M. Drozda, and P. Blum.** 2004. The Role of cis-acting Sequences Governing Catabolite Repression Control of lacS Expression in the Archaeon *Sulfolobus solfataricus*. Genetics **167:**1563-1572.
26. **Jeffries, T. W.** 2006. Engineering yeasts for xylose metabolism. Current Opinion in Biotechnology **17:**320-326.
27. **Kádár, Z., T. de Vrije, G. van Noorden, M. Budde, Z. Szengyel, K. Réczey, and P. Claassen.** 2004. Yields from glucose, xylose, and paper sludge hydrolysate during hydrogen production by the extreme thermophile *Caldicellulosiruptor saccharolyticus*. Applied Biochemistry and Biotechnology **114:**497-508.
28. **Karlin, S., L. Brocchieri, A. Campbell, M. Cyert, and J. MrÃzek.** 2005. Genomic and proteomic comparisons between bacterial and archaeal genomes and related comparisons with the yeast and fly genomes. Proceedings of the National Academy of Sciences of the United States of America **102:**7309-7314.
29. **Kramer, R., C. Lambert, C. Hoischen, and H. Ebbighausen.** 1990. Uptake of glutamate in *Corynebacterium glutamicum*. 1. Kinetic properties and regulation by internal pH and potassium. European Journal of Biochemistry **194:**929-35.
30. **Kurosawa, N., and D. W. Grogan.** 2005. Homologous recombination of exogenous DNA with the *Sulfolobus acidocaldarius* genome: Properties and uses. FEMS Microbiology Letters **253:**141-149.
31. **Lee, H. W., J. G. Pan, and J. M. Lebeault.** 1998. Enhanced l -lysine production in threonine-limited continuous culture of *Corynebacterium glutamicum* by using gluconate as a secondary carbon source with glucose. Applied Microbiology and Biotechnology **49:**9-15.
32. **Lee, S.-J., A. Böhm, M. Krug, and W. Boos.** 2007. The ABC of binding-protein-dependent transport in Archaea. Trends in Microbiology **15:**389-397.
33. **Loomis, W. F., Jr., and B. Magasanik.** 1967. Glucose-Lactose Diauxie in *Escherichia coli*. Journal of Bacteriology **93:**1397-1401.
34. **Lubelska, J., M. Jonuscheit, C. Schleper, S.-V. Albers, and A. Driessen.** 2006. Regulation of expression of the arabinose and glucose transporter genes in the thermophilic archaeon *Sulfolobus solfataricus*. Extremophiles **10:**383-391.

35. **Magasanik, B.** 1961. Catabolite repression. Cold Spring Harbor Symposia on Quantitative Biology **26:**249-56.
36. **Monod, J.** 1942. Recherches sur la croissance des cultures bactâeriennes. Hermann, Paris.
37. **Mosier, N., C. Wyman, B. Dale, R. Elander, Y. Y. Lee, M. Holtzapple, and M. Ladisch.** 2005. Features of promising technologies for pretreatment of lignocellulosic biomass. Bioresource Technology **96:**673-686.
38. **Murin, R., A. Schaer, B. S. Kowtharapu, S. Verleysdonk, and B. Hamprecht.** 2008. Expression of 3-hydroxyisobutyrate dehydrogenase in cultured neural cells. Journal of Neurochemistry **105:**1176-86.
39. **Nicolaus, B., A. Trincone, E. Esposito, M. R. Vaccaro, A. Gambacorta, and M. Derosa.** 1990. Calditol tetraether lipids of the archaebacterium *Sulfolobus solfataricus* - Biosynthetic studies. . Biochemical Journal **266:**785-791.
40. **Nunn, C. E. M., U. Johnsen, P. Schönheit, T. Fuhrer, U. Sauer, D. W. Hough, and M. J. Danson.** 2010. Metabolism of Pentose Sugars in the Hyperthermophilic Archaea *Sulfolobus solfataricus* and *Sulfolobus acidocaldarius*. Journal of Biological Chemistry **285:**33701-33709.
41. **Rougraff, P. M., R. Paxton, M. J. Kuntz, D. W. Crabb, and R. A. Harris.** 1988. Purification and characterization of 3-hydroxyisobutyrate dehydrogenase from rabbit liver. Journal of Biological Chemistry **263:**327-331.
42. **Sensen, C. W., R. L. Charlebois, C. Chow, I. G. Clausen, B. Curtis, W. F. Doolittle, M. Duguet, G. Erauso, T. Gaasterland, R. A. Garrett, P. Gordon, I. H. de Jong, A. C. Jeffries, C. Kozera, N. Medina, A. De Moors, J. van der Oost, H. Phan, M. A. Ragan, M. E. Schenk, Q. She, R. K. Singh, and N. Tolstrup.** 1998. Completing the sequence of the *Sulfolobus solfataricus* P2 genome. Extremophiles **2:**305-12.
43. **She, Q., R. K. Singh, F. Confalonieri, Y. Zivanovic, G. Allard, M. J. Awayez, C. C. Chan-Weiher, I. G. Clausen, B. A. Curtis, and A. De Moors.** 2001. The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2. Proceedings of the National Academy of Sciences of the United States of America **98:**7835 - 7840.
44. **Snijders, A. P. L., J. Walther, S. Peter, I. Kinnman, M. G. J. d. Vos, H. J. G. v. d. Werken, S. J. J. Brouns, J. v. d. Oost, and P. C. Wright.** 2006. Reconstruction of central carbon metabolism in *Sulfolobus solfataricus* using a two-dimensional gel electrophoresis map, stable isotope labelling and DNA microarray analysis. PROTEOMICS **6:**1518-1529.
45. **Sturn, A., J. Quackenbush, and Z. Trajanoski.** 2002. Genesis: cluster analysis of microarray data. Bioinformatics **18:**207-208.
46. **van de Werken, H. J. G., M. R. A. Verhaart, A. L. VanFossen, K. Willquist, D. L. Lewis, J. D. Nichols, H. P. Goorissen, E. F. Mongodin, K. E. Nelson, E. W. J. van Niel, A. J. M. Stams, D. E. Ward, W. M. de Vos, J. van der Oost, R. M. Kelly, and S. W. M. Kengen.** 2008. Hydrogenomics of the Extremely Thermophilic Bacterium *Caldicellulosiruptor saccharolyticus*. Applied and Environmental Microbiology **74:**6720-6729.
47. **VanFossen, A. L., M. R. A. Verhaart, S. M. W. Kengen, and R. M. Kelly.** 2009. Carbohydrate Utilization Patterns for the Extremely Thermophilic

Bacterium Caldicellulosiruptor saccharolyticus Reveal Broad Growth Substrate Preferences. Appl. Environ. Microbiol. **75:**7718-7724.

48. **Waterhouse, A. M., J. B. Procter, D. M. A. Martin, M. Clamp, and G. J. Barton.** 2009. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. Bioinformatics **25:**1189-1191.
49. **Weimberg, R.** 1961. Pentose Oxidation by *Pseudomonas fragi*. Journal of Biological Chemistry **236:**629-635.
50. **Wendisch, V. F., A. A. de Graaf, H. Sahm, and B. J. Eikmanns.** 2000. Quantitative determination of metabolic fluxes during coutilization of two carbon sources: comparative analyses with *Corynebacterium glutamicum* during growth on acetate and/or glucose. Journal of Bacteriology **182:**3088-96.

TABLES

Table 1. Growth kinetics of *S. acidocaldarius* on glucose, arabinose and xylose as sole or co-carbon sources

Growth Condition	Carbon Source(s)	G/X ratio	Specific growth rate (h^{-1})	Specific consumption rate (g/OD/h)[c#]	Yield Coefficient (OD/g)[d]
[a]Multiple sugars	2 g/L Xylose	n.a.	0.074	0.052	0.66
	2 g/L Glucose	n.a.	0.022	0.015	0.79
	1 g/L Glucose + 4 g/L Xylose	0.25	0.085	0.046	0.78
	1 g/L Glucose + 2 g/L Xylose	0.50	0.082	0.051	0.68
	1 g/L Glucose + 1 g/L Xylose	1.00	0.079	0.049	0.69
	2 g/L Glucose + 1 g/L Xylose	2.00	0.062	0.044	0.73
	4 g/L Glucose + 1 g/L Xylose	4.00	0.052	0.042	0.71
[b]Single sugar	2 g/L Glucose	n.a.	0.041	n.d	n.d
	2 g/L Xylose	n.a.	0.070	n.d	n.d
	2 g/L Arabinose	n.a.	0.082	n.d	n.d

n.d = not determine; n.a. = not applicable; G/X = glucose/xylose; OD = OD_{600}

[a]Inoculum was pre-adapted to glucose + xylose as co-carbon source

[b]Inoculum was pre-adapted to each sugar as sole carbon source

[c]Specific substrate consumption rate (q_s) = $1/X \cdot \Delta S/\Delta t$ (S = substrate; X = Growth, t = time)

[d]Yield coefficient ($Y_{X/S}$) = $\Delta OD/\Delta S$

[#]S = Total sugar concentration during simultaneous utilization of glucose and xylose

Table 2. Gene expression in *S. acidocaldarius* grown on glucose and xylose as sole and co-carbon sources

GENE ID	Description	Growth conditions*					
		G/X[b]		X/XG[a]		G/XG[a]	
		log_2 ratio	**P value**	**log_2 ratio[a]**	**P value**	**log_2 ratio[a]**	**P value**
Saci_0385	3-hydroxyisobutyrate dehydrogenase	0.32	0.17	4.45	0.02	4.77	0.01
Saci_0480	hypothetical protein	-0.39	0.17	-1.52	0.02	-1.91	0.01
Saci_0667	hypothetical protein	0.06	0.88	-1.61	0.03	-1.55	0.03
Saci_0748	hypothetical protein	0.25	0.38	-1.58	0.03	-1.33	0.01
Saci_0951	hypothetical protein	-0.23	0.43	-1.96	0.02	-2.19	0.01
Saci_1225	hypothetical protein	0.14	0.27	-2.05	0.02	-1.91	0.01
Saci_1302	hypothetical protein	0.11	0.61	-1.37	0.01	-1.26	0.01
Saci_1494	type II/IV secretion system protein	0.20	0.34	-1.47	0.03	-1.27	0.00
Saci_1496	hypothetical protein	0.36	0.09	-1.59	0.03	-1.23	0.03
Saci_1497	endonuclease III	-0.15	0.63	-1.52	0.02	-1.67	0.01

*G = cells grown on glucose; X = cells grown on xylose; XG = cells grown on a mixture of glucose and xylose

[a]Filtered genes with log_2 ratio of $|1.0|$ and above ($p \leq 0.05$) in G/XG and X/XG,that remained unchanged in G/X

[b]G/X must have a log_2 ratio below $|0.5|$ ($p \geq \approx 0.05$) to be considered unchanged

Table 3. List of potential xylose transporters in *S. acidocaldarius*.

GENE ID	Description	Transporter Superfamily[a]	Log_2 Ratio[b]		
			G/X	X/XG	G/XG
Saci_0880	ATP binding protein	ABC	-1.53	0.38	-1.15
Saci_0881	Permease	ABC	-2.00	0.32	-1.68
Saci_0882	Regulator	ABC	-1.00	0.56	-0.44
Saci_0883	Permease	ABC	-1.71	1.98	0.28
Saci_0946	ATP binding protein	ABC	-1.99	0.60	-1.39
Saci_1707	Sugar binding protein	ABC	-1.43	0.16	-1.27
Saci_1003	Permease	MFS	-3.04	1.69	-1.35
Saci_1004	Hypothetical Protein	unknown	-2.06	1.89	-0.17
Saci_1731	Permease	MFS	-1.67	1.47	-0.20

[a]ABC = ATP-binding cassette; MFS = Major facilitator superfamily

[b]Log_2 ratios of cell grown on 4 g/L glucose (G), 4 g/L xylose (X) or equal mixture (2 g/L each) of glucose and xylose (XG)

Table 4. Glycerol-3-phosphate shuttle pathway genes

GENE ID	Protein	Log_2 Ratio[a]		
		G/X	X/XG	G/XG
Glycerol-3-phosphate shuttle pathway genes				
Saci_2033	glycerol kinase	1.950	-0.405	1.545
Saci_2032	glycerol-3-phosphate dehydrogenase	1.972	-0.736	1.237
Saci_1117	glycerol kinase	-0.192	-0.233	-0.426
Saci_1118	glycerol-3-phosphate dehydrogenase	-0.840	0.002	-0.838

[a]Log_2 ratios of cell grown on 4 g/L glucose (G), 4 g/L xylose (X) or equal mixture (2 g/L each) of glucose and xylose (XG)

FIGURES

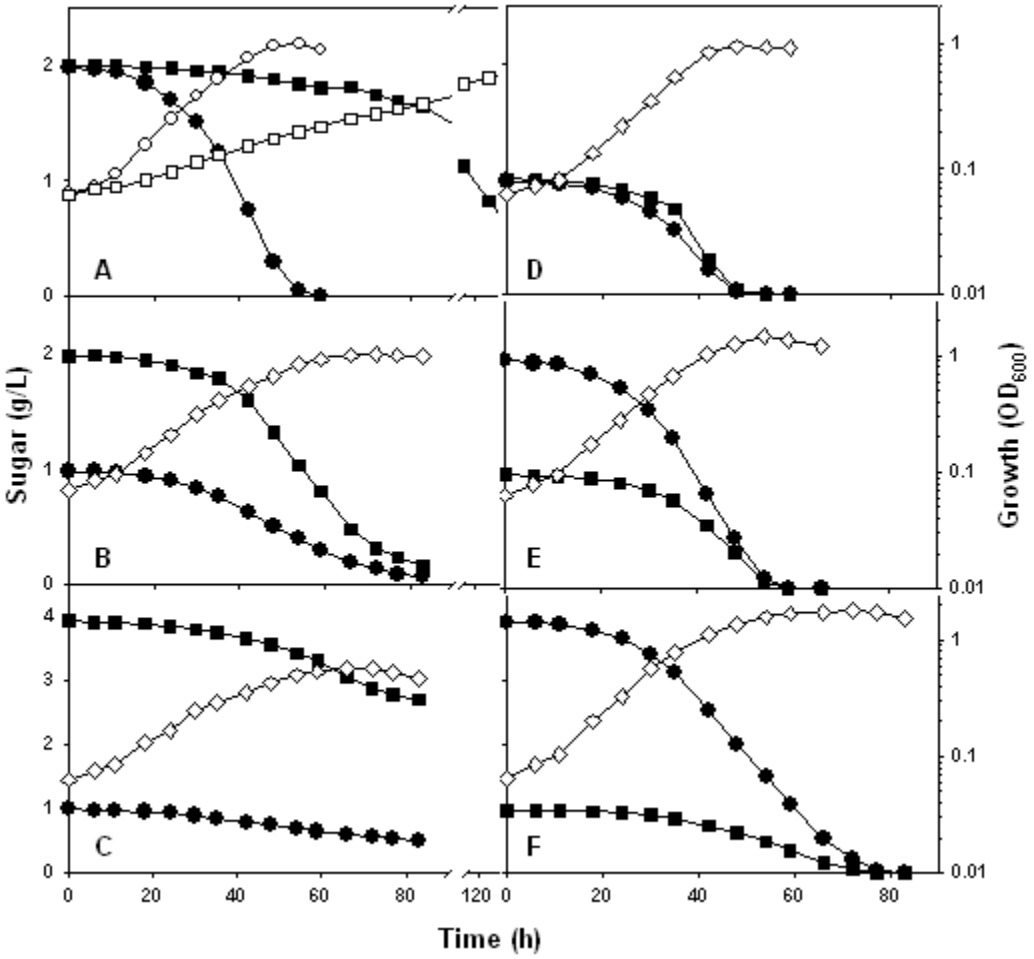


Figure 1. Growth and simultaneous utilization of glucose and xylose by *S. acidocaldarius*. The organism was grown at 75°C in minimal medium supplemented with glucose (squares) and xylose (circles) as sole or co-carbon sources. (A) Utilization of 2 g/L glucose or 2 g/L xylose as a sole carbon source; sugars (filled symbols), growth (open symbol). Simultaneous utilization of 2 g/L of glucose and 1 g/L xylose (B); 4 g/L glucose and 1 g/L xylose (C); 1 g/L of each of glucose and xylose (D); 1 g/L glucose and 2 g/L xylose (E); and 1 g/L glucose and 4 g/L xylose. Growth during simultaneous sugar metabolism is represented by open diamonds. Growth was monitored at OD_{600}.

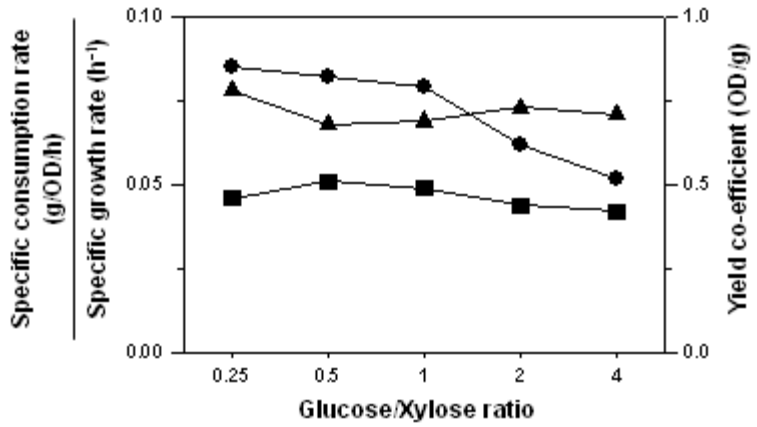


Figure 2. Kinetics of simultaneous utilization of glucose and xylose by *S. acidocaldarius*. Specific growth rate (circles); specific consumption rate (squares); and yield coefficient (triangles) were determined from cells grown on varying ratios of glucose and xylose.

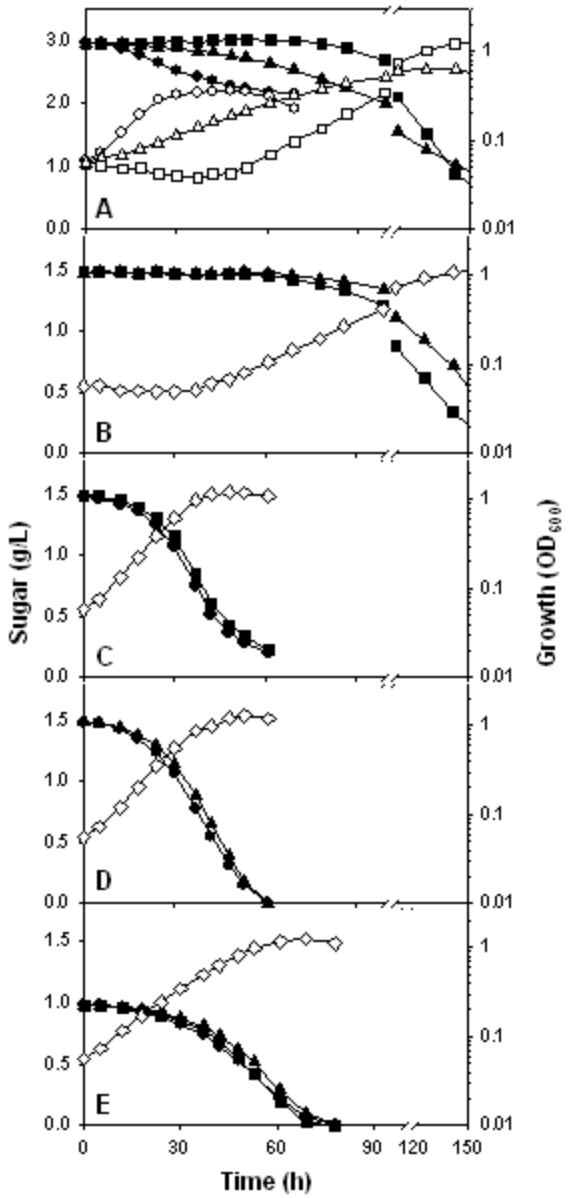


Figure 3. Growth of *S. acidocaldarius* on arabinose, glucose and galactose. The organism was grown at 75°C in minimal medium supplemented with glucose (squares) galactose (triangles) or arabinose (circles) as sole or co-carbon sources. (A) Growth on (open symbols) and utilization of (filled symbols) 3 g/L of each of arabinose, glucose or galactose as a sole carbon source. Simultaneous metabolism of 1.5 g/L binary combination of each of glucose and galactose (B); glucose and arabinose (C); galactose and arabinose (D). Simultaneous metabolism of 1 g/L trinary combination of each of glucose, galactose and arabinose (E). Growth during simultaneous sugar metabolism is represented by open diamonds. Growth was monitored at OD_{600}.

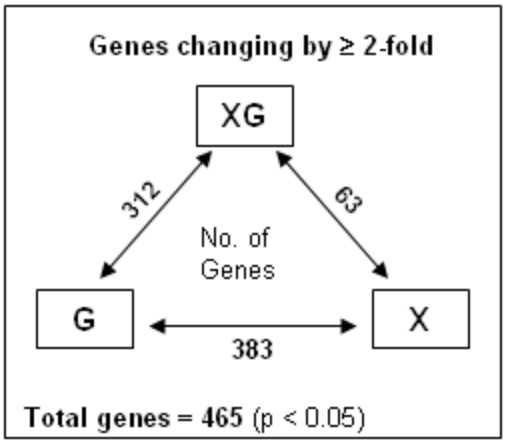


Figure 4. A comparison of the number of differentially expressed genes during growth of *S. acidocaldarius* on glucose (G), xylose (X), and a mixture of glucose and xylose (XG). Double headed arrows compares result from two growth conditions, highlighting the total number of genes changing by ≥ 2-fold ($p \leq 0.05$). The number at the bottom shows total number of differentially expressed genes (≥ 2-fold; $p \leq 0.05$).

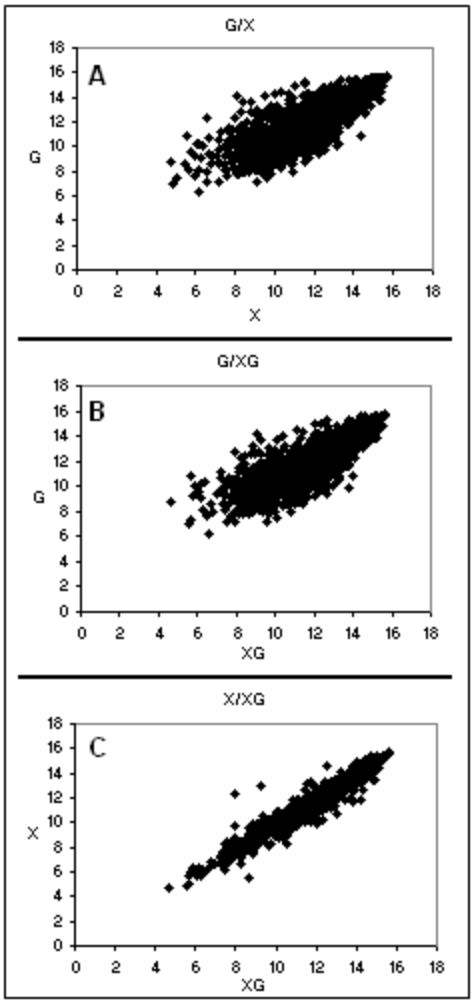


Figure 5. Scatter plots of total gene expression patterns between cells grown on G vs X (A), G vs XG (B), and X vs XG (C). Total RNA used for gene expression profile was extracted at OD_{600} = 0.2 – 0.4.

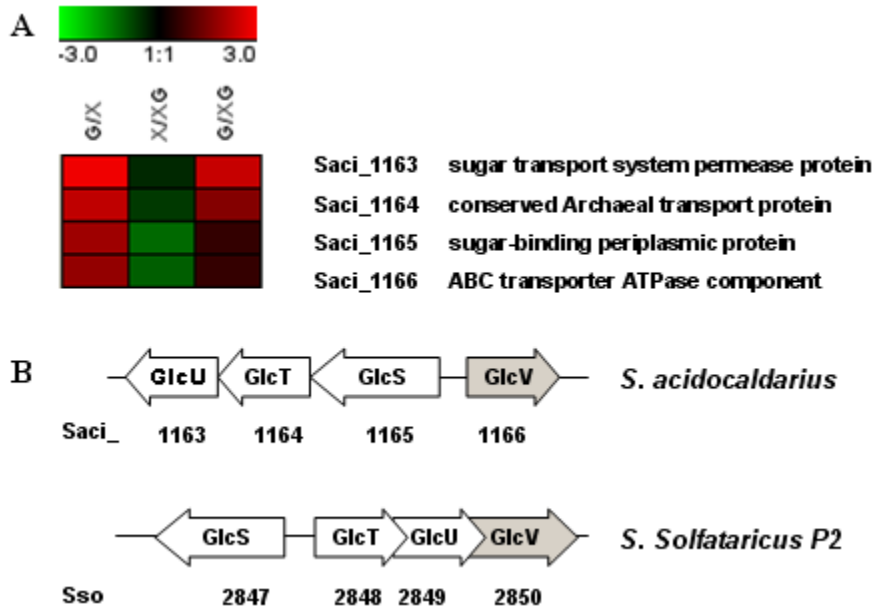


Figure 6. Identification of putative glucose ABC transporter in *S. acidocaldarius*. (A) Heat map showing up-regulation of Saci_1163 – 1166 encoding GlcSTUV, the putative glucose ABC transporter, in the presence of glucose. (B) Schematic representation of genomic arrangement of *glcSTUV* from *S. acidocaldarius* (putative) and *S. solfataricus* P2.

SUPPLEMENTARY MATERIALS

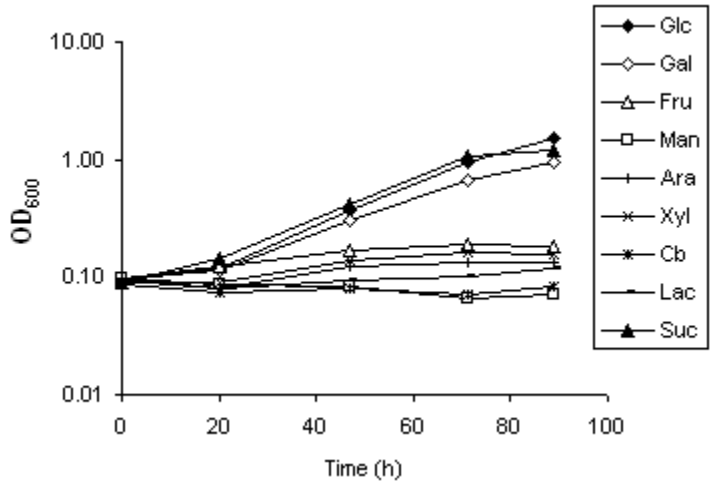


Figure 1. Growth of glucose-adapted *S. acidocaldarius 639* on sugars

The original strain of *S. acidocaldarius 639* (DSMZ) was adapted to glucose by repeatedly sub-culturing the organism on ATCC #1723 (YT) supplemented with 0.4% glucose. The profile of sugar utilization of this glucose-adapted cells was determined by growing the cells different sugars as sole carbon sources. The organism was grown at 75°C in minimal medium (NYT) supplemented with 0.4% of glucose (glc), galactose (gal), sucrose (suc), arabinose (ara), xylose (xyl), mannose (man), fructose (fru), lactose (lac) and cellobiose (cb) as sole carbon sources. The glucose-adapted cells were able to efficiently utilize glucose, galactose and sucrose as sole carbons sources. But the organism barely utilized fru, man, cb, ara, xyl and lac as sole carbon sources.

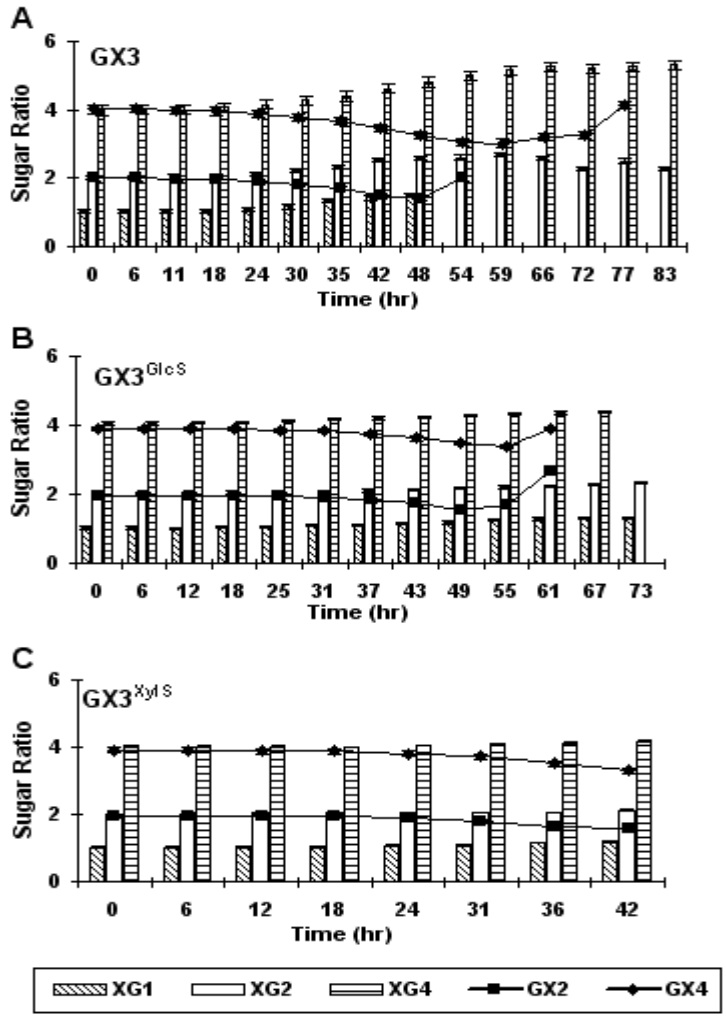


Figure 2. Ratio of residual sugars during growth of different phenotypes of *S. acidocaldarius* on glucose and xylose. Residual ratio of glucose and xylose during simultaneous utilization of both sugars as co-carbon sources by different phenotypes of *S. acidocaldarius*. Bars highlight ratio of glucose to xylose; while line graphs highlight ratio of xylose to glucose (A) GX3 – wild-type (*S. acidocaldarius 639*), (B) $GX3^{GlcS}$ – phenotype with reduced glucose metabolism and (C) $GX3^{XylS}$ – phenotype with reduced xylose metabolism. Each phenotype was grown at 75°C in minimal media containing varying concentrations of glucose and xylose; XG1 = 0.1 % xylose + 0.1 % glucose; XG2 = 0.1 % xylose + 0.2 % glucose; XG4 = 0.1 % xylose + 0.4 % glucose; GX2 = 0.1 % glucose + 0.2 % xylose and GX4 = 0.1 % glucose + 0.4 % xylose.

$GX3^{GlcS}$ and $GX3^{XylS}$ are phenotypes of *S. acidocaldarius* generated by repeatedly growing the organism on xylose and glucose as sole carbon sources. $GX3^{GlcS}$ grows normally on xylose and shows weak growth on glucose, while $GX3^{XylS}$ has impaired xylose metabolism but grows normally on glucose. Each phenotype is easily reverted to the WT by growing the cells on YT supplemented with the sugar that it can longer metabolize efficiently.

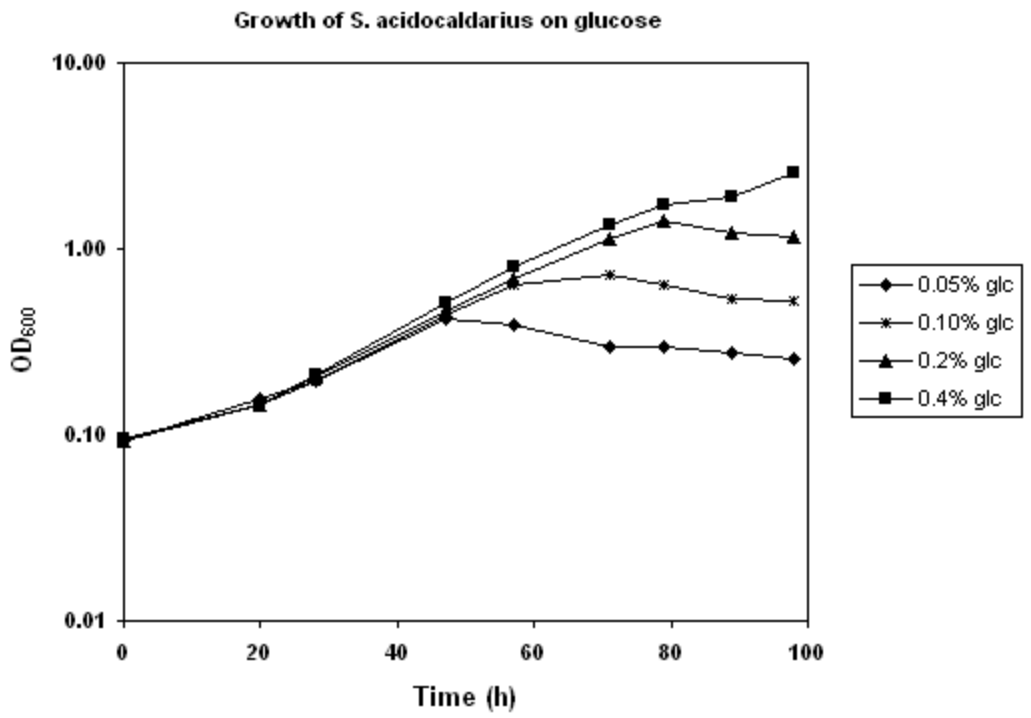


Figure 3. Growth of *S. acidocaldarius* on different glucose concentrations. *S. acidocaldarius* was grown in minimal medium (NYT) supplemented with 0.5, 1, 2 and 4 g/L as sole carbon source. Each experiment was carried out in triplicate and the organism was growth at 75°C. The maximum growth yield doubled as the sugar concentration doubled.

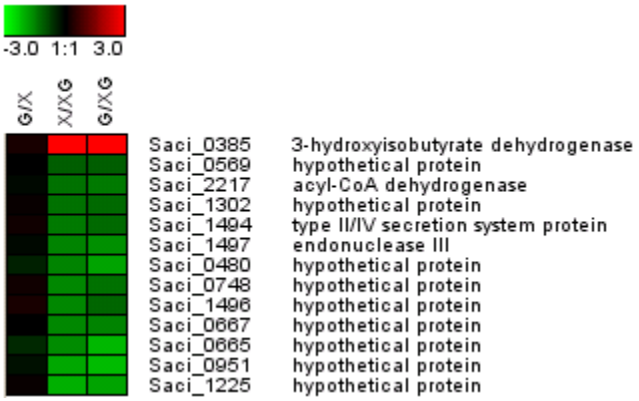


Figure 4. Heat map of genes expressed in response to go the growth of *S. acidocaldarius* on a mixture of glucose and xylose.

Comparison of gene expression during growth on glucose alone, xylose alone, or glucose and xylose together. The heat map highlights genes that were differentially expressed only during co-utilization of glucose and xylose that remained relatively unchanged during growth on each sugar as a sole carbon source. The organism was grown in minimal media supplemented with 0.4 % glucose (G), 0.4 % xylose (X), or 0.2 % glucose + 0.2 % xylose (XG); and the gene expression level was determined at mid-exponential growth phase (OD_{600} = 0.2 – 0.4). The column represents $\log_2$ ratios of gene expression in cells grown on glucose versus xylose (G/X), xylose versus xylose + glucose (X/XG), and glucose versus xylose + glucose (G/XG).

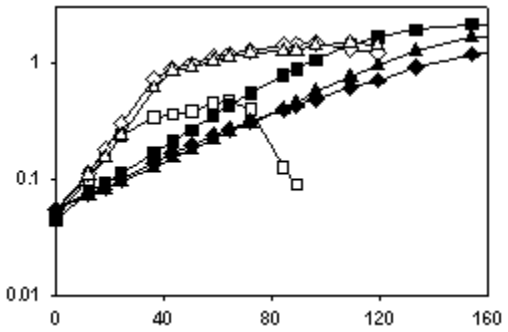


Figure 5. Effect of adaptation to growth of *S. acidocaldarius* on glucose and xylose. Growth cells pre-adapted to glucose (squares), xylose (diamond) or a mixture of glucose and xylose (triangle) on 4 g/L glucose (solid symbols) or 4 g/L xylose (open symbols) as sole carbon sources in minimal medium, at 75°C. All experiments were conducted in triplicates.

We investigated the physiological similarities between xylose-grown cells of *S. acidocaldarius* and those grown on a mixture of glucose and xylose by adapting the cells to each growth condition. The cells were sub-cultured at least five-times on glucose, xylose or a mixture of both sugars respectively. The pre-adapted cells were then grown on glucose and xylose as sole carbon sources. We observed that cells that were pre-adapted to xylose or a mixture of glucose and xylose showed similar growth characteristics when grown on either glucose or xylose, compared to glucose-adapted cells. Glucose-adapted cells grew normally on glucose; but their growth on xylose diminished greatly after 24 h and stopped completely after 64 h; but the growth rate at exponential phase was practically similar to that of xylose adapted cells. This result suggests physiological similarities in cells grown xylose or a mixture of glucose and xylose. The strong decrease in xylose metabolism observed with the glucose-adapted cells further confirmed our observation that repeated sub-culturing of the organism on one sugar gradually decreases its ability to utilize other sugars, although this growth retardation is easily reversed.

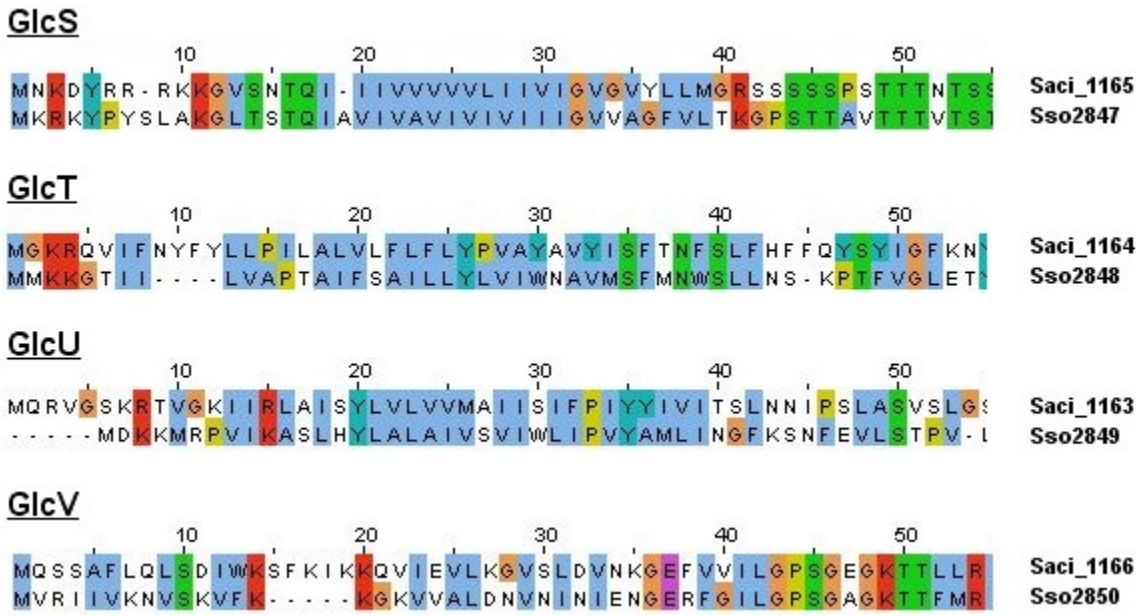


Figure 6. Comparison of N-terminal amino acids of *glcSTUV* from *S. acidocaldarius* and *S. solfataricus*. A comparison of N-terminal amino acid sequence of glcSTUV (glucose ABC transporter) from *S. acidocaldarius* (Saci_) and *S. solfataricus* (Sso) revealed some similarities. ClustalW2 alignments showing similarities in the of N-terminal amino acid sequence of GlcSTUV from *S. acidocaldarius* and *S. solfataricus*. Alignment is displayed in Jalview using clustalX color scheme.

Table 1. Expression of TCA and some house-keeping genes during growth of *S. acidocaldarius* on glucose and xylose

GENE ID	Protein	Log_2 Ratio[a]		
		G/X	X/XG	G/XG
TCA Cycle				
Saci_0243	Citrate synthase	-0.107	0.031	-0.076
Saci_1214	Aconitate hydratase	0.054	0.034	0.088
Saci_2375	Isocitrate dehydrogenase	-0.517	-0.160	-0.677
Saci_2306	2-oxoglutarate synthase alpha subunit	0.068	-0.049	0.018
Saci_2307	2-oxoglutarate synthase beta subunit	0.047	0.007	0.054
Saci_0208	2-oxoglutarate synthase alpha subunit	-0.234	0.211	-0.023
Saci_0209	2-oxoglutarate synthase beta subunit	-0.130	0.133	0.003
Saci_1265	succinyl-CoA synthetase beta chain	-0.305	-0.070	-0.374
Saci_1266	succinyl-CoA ligase, alpha chain	-0.425	-0.038	-0.463
Saci_0979	Succinate dehydrogenase Subunit D	0.155	-0.017	0.138
Saci_0980	Succinate dehydrogenase Subunit C	0.093	-0.034	0.060
Saci_0981	Succinate dehydrogenase Subunit B	-0.009	0.107	0.097
Saci_0982	Succinate dehydrogenase Subunit A	0.127	0.030	0.157
Saci_0122	fumarate hydratase	-0.809	0.180	-0.629
Saci_0246	Malate dehydrogenase	-0.406	-0.051	-0.456
Glycerol-3-phosphate shuttle pathway genes				
Saci_2033	glycerol kinase	1.950	-0.405	1.545
Saci_2032	glycerol-3-phosphate dehydrogenase	1.972	-0.736	1.237
Saci_1117	glycerol kinase	-0.192	-0.233	-0.426
Saci_1118	glycerol-3-phosphate dehydrogenase	-0.840	0.002	-0.838
Chaperonin				
Saci_1401	thermosome alpha subunit	-0.146	0.141	-0.005
Saci_0666	thermosome beta subunit	-0.054	0.038	-0.016
Saci_1203	thermosome gamma subunit	-0.115	0.012	-0.103
Replication/ Translation				
Saci_0074	DNA polymerase II	0.105	0.179	0.284
Saci_0075	DNA primase	-0.031	0.008	-0.023
Saci_0076	tyrosyl-tRNA synthetase	0.348	-0.109	0.239
Saci_0080	30S ribosomal protein S13P	0.105	-0.082	0.024
Saci_0081	30S ribosomal protein S4	0.031	-0.058	-0.027
Saci_0082	30S ribosomal protein S11P	-0.067	0.051	-0.016
Saci_0083	DNA-directed RNA polymerase subunit D	0.142	0.050	0.193
Saci_0084	50S ribosomal protein L18e	0.000	-0.039	-0.039
Saci_0085	50S ribosomal protein L13P	0.097	-0.039	0.058

[a]Log_2 ratios of cell grown on 4 g/L glucose (G), 4 g/L xylose (X) or equal mixture (2 g/L each) of glucose and xylose (XG)

This table highlights some of the genes that remained relatively unchanged in *S. acidocaldarius* that was grown on glucose and xylose as sole carbon sources or on a mixture of both sugars. The level of gene expression was determined at mid-exponential growth phase (OD_{600} = 0.2 – 0.4). There in normally little or no variation in the level of expression of these house-keeping genes in organisms regardless of the growth condition.

Chapter Three: Identifying the Origin of Replication of pRN1 from *Sulfolobus islandicus* REN1H1

This chapter will be submitted as a manuscript to the Journal of Bacteriology.

ABSTRACT

Sulfolobus islandicus REN1H1 harbors a cryptic plasmid (pRN1) that has been used for heterologous gene expression in *S. acidocaldarius*. This plasmid is known to replicate by a rolling circle mechanism, but its origin of replication has not been fully elucidated. We evaluated the role of a 241-bp region downstream of the *orf904* and found that it was critical for pRN1 replication. We therefore propose this 241-bp region to be the putative origin of replication of pRN1. This putative origin and the associated *orf56* and *orf904* were identified as the minimum replicon of pRN1. Deletion of any of these three pRN1 features from pRSP1 prevented the resulting plasmid from replicating in *S. acidocaldarius*. A 100-bp stem-loop structure within this putative origin is proposed to be the double-strand origin of replication of the plasmid. The loop of the 100-bp structure contains a GTG tri-nucleotide motif that was previously reported to be important for the primase activity of Orf904. Regions flanking the 5' and 3' ends of the 100-bp stem-loop are proposed to be the single-strand origin and terminator region of the plasmid, based on the phenotypes observed with cells transformed with plasmid lacking these regions. We showed that *S. acidocaldarius* could be successfully transformed with a combination of pRSP9 and pRSP10b that harbor WT *orf56* and *orf904*, respectively. The co-transformation results suggest that the mutation on each plasmid can be complemented by the WT gene that is present on the other plasmid. The results also indicated that the 75-bp upstream of the start codon of *orf904* codon could be essential for plasmid replication because a combination of pRSP9 and pRSP10, which lacks this 75-bp did not yield transformants. Detailed knowledge of pRN1 origin of replication would broaden the application of the plasmid in the development of genetic tool for *Sulfolobus* species.

INTRODUCTION

The multi-copy plasmid pRN1 (5350 bp) is a self-replicating cryptic plasmid from *Sulfolobus islandicus* REN1H1 (35). The plasmid belongs to the pRN family of plasmids that replicate via the rolling circle mechanism, but its origin of replication is not clearly known (3-4, 19-20, 22, 25, 29, 31). The plasmid contains six genes, two of these genes (*orf56* and *orf904*) are highly conserved within the pRN family of plasmids, while a third gene (*orf80*) is less conserved (19, 22, 29); the other genes are not conserved. Two of the genes on pRN1 (*orf56* and *orf904*) are co-transcribed, and both genes have been shown to be critical for pRN1 replication (4-5, 19). The protein encoded by *orf56* (CopG) is proposed to regulate plasmid copy number and its own level by binding to its own promoter (25-26). Orf904 is a large, multi-functional, replication protein with helicase, primase, and polymerase domains (25). Orf904 is believed to initiate replication by binding to the plasmid's origin of replication; its primase domain requires a specific trinucleotide (GTG) motif for activity (2). Comparative analysis of pRN1 and other members of the pRN family of plasmids suggested that the single-strand origin (*sso*) and double-strand origin (*dso*) of replication of pRN1 are located 3' of *orf904*, but the exact origin of replication has not been confirmed experimentally (22, 29).

More than a third of pRN1 has been suggested to be non-essential for the replication of the plasmid (25), but it very challenging to eliminate DNA sequence when the origin of replication is not known. Knowledge of the origin of replication of pRN1 would aid in eliminating non-essential DNA from the plasmid, which would in turn make it easier to use for gene expression studies. A number of shuttle vectors that replicate in other *Sulfolobus* species have been constructed from pRN1, making the plasmid a great genetic/molecular biology tool for the study of gene function in *Sulfolobus* species (4).

Genetic tools have been developed for *Sulfolobus* species over the years, but few of these have been independently used in other laboratories (6). One of the key challenges facing advancement of genetic tools for *Sulfolobus* is the absence of effective selectable markers. The commonly used selection strategies are based on uracil or lactose auxotrophy (4, 17, 24, 30, 32). The use of antibiotics for selection in *Sulfolobus* species has been reported (9) but this strategy has not been very successful largely because of low susceptibility of archaea to antibiotics (1, 8). Genes have been heterologously expressed in *Sulfolobus* species either by means of homologous recombination or by the use of shuttle vectors (3, 20, 28). In general, it is very difficult to achieve homologous recombination, thus making the use of plasmids a more practical option. A number of plasmid systems for *Sulfolobus* species have been developed (6); some of the most reproducible *Sulfolobus* shuttle vectors are those based on pRN1 such as pC, pJ-lacS and pCmalLacS, but these shuttle vectors are large, usually in the excess of 8 kb (3, 6).

Sulfolobus acidocaldarius is a hyperthermophilic, aerobic crenarchaea and that grows optimally at 75°C and pH 2 – 4 (10). *S. acidocaldarius* is capable of utilizing mixtures of sugars such as glucose and xylose simultaneously as sources of carbon and energy without an apparent diauxie, indicating the absence carbon catabolite repression (14, 16, 18, 27). The genomes of most *Sulfolobus* species, with the exception of *S. acidocaldarius*, are fairly unstable due the presence of large numbers of active insertion elements (10, 33). These organisms are great models for understanding physiology and functions in archaea (10, 33). Our knowledge of this third domain of life would further be

enhanced with the availability of reliable genetic tools. Our goal was to identify the origin of replication of pRN1 in effort to develop shuttle vectors that are comparable to commonly used bacterial plasmids in terms of size and stability.

In this study, we identified and characterized the origin of replication of pRN1 located 3' of *orf904*. We generated shuttle vectors of various sizes (pRSP1 – 7), all of which harbored the pRN1 origin. Deletion of the origin from the pRSP1 or pRSP3 terminated replication of the vectors. The minimum replication unit of pRN1 was confirmed to consist of *orf56*, *orf904*, and the newly identified origin of replication. We also present evidence that deletion within the 100-bp stem-loop structure prevents replication of pRN1-based vectors.

MATERIALS AND METHODS

Strains and growth media. Wild-type *S. acidocaldarius* 639 used in this study was obtained from DSMZ (Braunschweig, Germany). The strains were cultivated aerobically in liquid ATCC #1723 (YT) medium at 75ºC (200 rpm) and pH 2.5. The pH of the growth medium was adjusted with 5 M H_2SO_4 and filter sterilized using 0.2-µm pore size membrane filters. Growth on solid medium was carried out at 70ºC on YT or T (YT without yeast extract) medium solidified with 0.8% gellan gum (Spectrum and Sigma). Chemicals used in this study were obtained from Sigma, Spectrum, and Fisher scientific.

Generation of *S. acidocaldarius E4-39*, a spontaneous uracil auxotroph. *S. acidocaldarius* (DSMZ 639) was grown in liquid YTFU medium (YT medium supplemented with 5 µg/ml uracil and 50 µg/ml 5-fluoroorotic acid [5-FOA]) at 75ºC (200 rpm) for 6 days. The culture was spread onto solid media: YT, YTFU (20 µg/ml uracil, 50 µg/ml 5-FOA), YTF (no uracil) and YTU (no 5-FOA) to select for spontaneous uracil auxotrophs. Genomic DNA was extracted from selected mutants using the DNeasy blood and tissue kit (Qiagen) and analyzed for mutations in *pyrEF*. *S. acidocaldarius* strain E4-39 with a 17-bp duplication of a semi-palindromic sequence within *pyrE* was selected as the candidate uracil auxotrophic host (Supplementary Figure 1).

Construction of shuttle vectors. We constructed seven shuttle vectors (pRSP1 – 7) using three PCR-amplified DNA fragments (Table 1 and Figure 1.). First we amplified a 1750-bp fragment from pUC19 (2550 – 801) that includes the *bla* gene (Amp^R) and the pMB1 origin of replication (Figure 1) using forward primer P/N128 and reverse primer P/N129. Primer P/N128 has a 33-bp overhang (Supplementary Table 1, underlined) that overlaps the DNA sequence downstream of *S. solfataricus lacS* (terminator element), while P/N129 a 30-bp overhang (Supplementary Table 1, underlined) that overlaps *orf56*.

The second fragment, a 2239-bp *pyrE*::*lacS* cassette (Figure 1B) consisting of Sso *pyrE* and *lacS*, was amplified from pGlcSTUV_ko (unpublished) using forward primer P/N120 and reverse primer P/N117. The *pyrE-lacS* operon on the cassette is transcribed under the control of the *thsB* (Sso0282) promoter, and the *lacS* terminator element was included as part of an 80-bp sequence 3' of the *lacS* stop codon. Translation of *lacS* in the *pyrE-lacS* operon is directed by the RBS of Sso *pyrF* and with a single adenine (A) nucleotide between its start codon and the stop codon of *pyrE*.

Varying length of pRN1 were amplified from shuttle vector pC (4) using primers described in Supplementary Table 1. Each of the fragments contains *orfs56/904* and a 241 bp (5022 – 5262) A-T region immediately downstream of *orf904* in common. Some fragments contain additional contiguous stretch of DNA (Figure 1B). The 241 bp (5022 – 5262) A-T region was identified as the putative origin of replication of the plasmid by visual inspection of un-annotated DNA sequence of pRN1.

The first set of shuttle vectors was assembled by combining three fragments: pUC19, *pyrE::lacS* cassette and pRN1 to construct pRSP1, 3, 5, and 7, using a modified SLIC cloning strategy (12, 28). The fragments used in the plasmid construction overlapped one another by at least 30-bp and were used in approximately 1:1 molar ratio; the pUC19 fragment served as a reference (50 – 100 ng). Each set of fragments was added to a 30-µl chew-back and annealing (CBA) reaction mixture consisting of T4 ligase buffer (Fermentas), 0.1 mg/mL BSA, 0.75 µL T4 polymerase (Fermentas), and water. The CBA reaction was carried out in a thermocycler at 37ºC for 7 min, 75ºC for 20 min, cooled to 60ºC at 0.1ºC /s, and held at 60ºC for 30 min, and then cooled to 4ºC at 0.1ºC /s (12). A second fill-in/ligation (PL) reaction was carried out by adding 0.2 µl T4 polymerase, 0.2 µl 10 mM dNTP's, 0.2 µl 50 mM ATP and 0.5 µl T4 ligase to 10 µl of the CBA reaction and incubating at 37ºC for 30 min and then at 75ºC for 5 min to inactivate the enzyme. Samples from each cloning reaction (CBA and PL) were transformed into the *E. coli* cloning strains DH10B and TOP10.

We eliminated *lacS* from pRSP1, 3 and 5 to construct pRSP2, 4 and 6 respectively, using P/N135 and P/N 136 as forward and reverse primers respectively (Supplementary Table 1). Primer P/N135 has a 30-bp overhang at its 5' end that overlaps the 3' end of *pyrE*, resulting in the amplification of circularized plasmids.

Mapping the putative origin of replication of pRN1. The putative origin of replication of pRN1was mapped by deleting various regions within the putative origin from pRSP1 or pRSP3 (Table 1). The resulting plasmids were tested for loss of plasmid replication *in vivo*. The first set of plasmids (pRSP1a – d) had 52- or 60-bp deletions within the putative origin from pRSP1 (Table 1) that covers the entire origin region 3' of the *orf904* stop codon (Figure 2A). A 12-bp sequence that includes a GTG tri-nucleotide motif was also deleted from pRSP1 to construct pRSP1e (Table 1) to determine the role of the motif in plasmid replication.

Next, we mapped the 58-bp loop (4755 – 4812) of a stem-loop (SL) secondary structure within the putative origin in pRSP1 (4734 – 4833) to determine its role in the initiation of plasmid replication. The loop (L) was mapped using three plasmids (pRSP1-L1 – 3) by deleting 18- or 20-bp segments from the loop; a fourth plasmid (pRSP1-L4) was constructed by deleting the entire loop (Table 1 and Figure 2A). The third set of plasmids (pRSP1-G1 – 8) was constructed by deleting tri-nucleotide segments from a stretch of DNA (4773 – 4796) located within the loop of the SL in pRSP1 to evaluate the role of the GTG motif and adjacent sequence in the initiation of plasmid replication. The primers used in the construction of these plasmids are listed in Supplementary Table 1.

With the exception pRSP1b and pRSP1d, each plasmid described above was amplified directly from pRSP1 as a linear blunt-end PCR fragment without the target DNA. Approximately 0.2 – 0.3 pmol of the 5'-terminus of each linear amplicon was

phosphorylated at 37ºC for 20 min with T4 polymerase kinase (PNK) and ligated overnight at ~22ºC using T4 ligase as described by the manufacturer (Fermentas). The plasmids pRSP1b and pRSP1d were amplified as circular plasmids using primers with overlapping sequence (Supplementary Table 1). Each construct was transformed into cloning strains of *E. coli* such as DH10B and TOP10.

Plasmid complementation assay. To verify the role of the co-transcribed *orf56* and *orf904* in pRN1 replication, we constructed pRSP9 and pRSP10 by making in-frame deletions of the entire *orf904* and *orf56*, respectively, from pRSP1 (Table 1). A third plasmid, pRSP10b, was constructed by retaining 75-bp sequence upstream of the *orf904* start codon. The plasmids were constructed by PCR amplification of entire pRSP1 without the gene segment targeted for deletion using primers described in Supplementary Table 1 and the protocol described above. Each construct was methylated and transformed into *S. acidocaldarius E4-39*; pRSP9 was paired with pRSP10 or pRSP10b and transformed into the host to test for gene complementation.

Methylation and transformation of *S. acidocaldarius*. Each assembled shuttle vector was methylated to protect it against SuaI, a restriction endo-nuclease present in *S. acidocaldarius* (13) by transformation into *E. coli* 1821.pM.EsaBC4I (New England Biolabs) kindly provided by Dr. D. Grogan. The methylated constructs were transformed into 80 µl of electro-competent uracil auxotrophic *S. acidocaldarius* E4-39 cell suspension (OD_{600} = 5 – 10) diluted with approximately 5 parts of ice-cold 1% sucrose solution. The competent cells were transferred into a 1-mm cuvette and electroporated at 1.25 KV, 50 µF, and 750 Ω and recovered with 80 µl of 2X recovery buffer (1% sucrose, 20 mM β-alanine/1.5 mM malate buffer [pH = 4.5] and 10 mM $MgSO_4$). The recovered cells were incubated at 70ºC for 30 min, cooled on ice for 2 – 5 min, spread on YT or T plates and incubated at 70ºC in humidified Sterilite™ plastic containers. The competent cells were prepared following "G" procedure as described by Kurosawa and Grogan (23).

X-gal staining and plasmid recovery from *S. acidocaldarius*. Positive transformants were screened qualitatively for LacS activity by spraying each transformation plate with 5 mg/ml X-gal (5-bromo-4-chloro-3-indolyl-b-D-galactoside) dissolved in dimethyl formamide (DMF). Blue colonies were selected for further characterization. For transformants from plasmids without *lacS*, small to medium sized colonies were selected for further characterization. Plasmid DNA was recovered from about 50 ml culture of *S. acidocaldarius* transformants (OD600 ≈ 1.0) using modified Qiagen miniprep protocol. The cells were pelleted and re-suspended in 500 µl of P1 buffer supplemented with 15 mg/ml lysozyme. The re-suspended cells were incubated at 37ºC for 10 min, then treated with 500 µl of buffer P2 and 700 µl of buffer N3 following the protocol described by the manufacturer. The recovered plasmids were treated with EcoRI to determine the restriction digestion fingerprint.

Bioinformatics analyses. The plasmids used in this study were designed and analyzed using Invitrogen's Vector NTI 11. Plasmid DNA sequence analyses and alignments were carried out using: http://www.ncbi.nlm.nih.gov and ClustalW2 (http://www.ebi.ac.uk/Tools/clustalw2/index.html).

RESULTS

Putative 'origin' of pRN1 is critical for plasmid replication. We evaluated the role of a 241-bp stretch of DNA (5022 – 5262) on pRN1 as the origin of replication using *E. coli-Sulfolobus* shuttle vectors pRSP1 – 7 (Table 1). Each of these shuttle vectors yielded uracil prototrophic colonies upon transformation into *S. acidocaldarius E4-39*, suggesting that all of the plasmids replicated (Table 2 and supplementary Figure 2). Replication of pRSP1 and pRSP2 indicates that *orf56/904* and the identified putative origin of replication (Figure 1B), is the minimum replication unit of pRN1. We conclude that the region of pRN1 that was included in pRSP1 and pRSP2 was not essential for plasmid replication. In addition, more transformants were obtained with pRSP1 and pRSP2 than with any of the other plasmids. Colonies from pRSP1, 3, 5 and 7 transformations stained blue on X-gal plates due to β-glycosidase activity encoded by *lacS* from *S. solfataricus* (Supplementary Figure 2b, d, f and h). Because the host strain can revert to WT, we added *lacS* from *S. solfataricus* to the plasmids to enhance screening of the transformants. Colonies from pRSP2, 4 and 6 did not turn blue when exposed to X-gal due to the absence of *lacS* (Supplementary Figure 2c, e and g). We recovered and digested each plasmid from the transformed *S. acidocaldarius* strains (Table 1) to verify that they were freely replicating and not integrated into the host's chromosome (Figure 3). The EcoRI fingerprints of the recovered plasmids were consistent with the original constructs.

Replication of pRSP1 and pRSP2 in *S. acidocaldarius* confirms that the origin of pRN1 was present in the each of these plasmids (Table 2). The origin of replication of plasmids that replicate by rolling circle mechanism is usually located within or adjacent to the 3'end of the gene encoding the replication protein (20). Deletion of the putative origin of replication from pRSP1 and pRSP3 prevented recovery of transformants (Table 2 and Supplementary Figure 3). The inability of these plasmids to replicate suggests that there is no additional origin of replication within *orfs56/904* or the sequence adjacent to the 3' end of the deleted putative origin in pRSP3-NO. The region 3' end of pRN1 putative origin in pRSP3 was previously described as the double-strand origin of the plasmid (22, 29). Failure of pRSP1-NO or pRSP3-NO to replicate verifies that the deleted *cis*-acting element is critical for pRN1 replication and might potentially contain the plasmid's double-strand origin.

Replication initiation stem-loop is located within the pRN1 putative origin. We have shown that a *cis*-acting region downstream of *orf904* was critical for pRN1 replication in *S. acidocaldarius* and might contain the double-strand origin of replication, despite its previous description as the single-strand origin (22, 29). We mapped this region by dividing it into four segments of 52 or 60 bp in length and deleted each segment of from pRSP1 generating pRSP1a – d (Figure B). Two of these plasmids, pRSP1b and pRSP1c failed to produce transformants when transformed into *S. acidocaldarius E4-39*, while transformants of pRSP1a and pRSP1d produced tiny colonies on solid medium Supplementary Figure 3). Colonies from pRSP1a transformation grew poorly in liquid medium, while those from pRSP1d transformation did not grow in liquid medium. The inability of pRSP1b and pRSP1c to replicate suggests that deleted regions harbor critical *cis*-acting elements for pRN1 replication. We also deleted a 12-bp sequence that contains a GTG tri-nucleotide motif, but the resulting

plasmid (pRSP1e) replicated normally suggesting that the motif was not essential for plasmid replication (Table 2). The GTG motif, which was identified by aligning the pRN1 putative origin with the sequences of pRN2, pHEN7 and pDL10 (Supplementary Figure 4), has been shown to be important for primase activity of Orf904 (2).

The 241-bp putative origin contained two stem-loop structures (determined at 70°C, Na^+ = 1.0 M, Mg^{2+} = 0.0 M using mFold (36)) located at 5043 – 5061 (19-bp) and 5084 – 5183 (100-bp) (Figure 4). Approximately one-half of the 100-bp stem-loop was deleted in pRSP1b and pRSP1c (Figure 4). The inability of pRSP1b and pRSP1c to replicate in the host, suggests that part or the entire 100-bp stem-loop structure is critical for pRN1 replication. Therefore, we propose that the 100-bp stem-loop structure is part of the *cis* acting, double-strand origin of replication of pRN1 and possibly involved in the formation of the replication cruciform structure (Figure 4). The double-strand origins of rolling circle replicating plasmids are generally 100 bp or less (11, 20),

The much smaller, 19-bp stem-loop structure appeared to be important for replication but not critical; deletion of this structure (Figure 4A) and the adjacent DNA sequence in pRSP1d resulted in tiny colonies that did not grow in liquid medium (Supplementary Figure 3). Failure of pRSP1d transformants to grow in liquid, suggests that the plasmid is unstable. The instability of pRSP1d in the *S. acidocaldarius E4-39* suggests that the deleted DNA fragment could be a component of the pRN1 single-strand origin (Figure 4C), because the deletion of single-strain origin from a rolling circle plasmid usually results in plasmid instability (34). The possible role of the 19-bp stem-loop in pRSP1d as the pRN1 single-strand origin is further supported by its location upstream of the 100-bp stem-loop structure(15), and the possibility of two additional small hairpin structures (Supplementary Figure 5**)**. There are three potential small hairpin structures in the region deleted in pRSP1a which could function as single-strand origin, considering that the deletion also has an adverse effect on the growth of the transformants. However, we suspect that this region might be the terminator element responsible for the termination of pRN1 replication based on its location downstream of the 100-bp stem-loop structure (Figure 4).

Loop of the 100-bp stem-loop structure is critical for replication. The 100-bp stem-loop structure within the pRN1 putative origin of replication is similar to the cruciform structures of plasmids that replicate by rolling circle mechanism (11, 20). The loop of the cruciform structure has been reported by numerous studies to be critical for initiation of replication because it harbors the 'nick site' (11, 20). To determine if this loop is important for pRN1 replication, we deleted the 58 bp G-C rich loop of the 100-bp stem-loop structure from pRSP1; the resulting plasmid (pRSP1-L4) failed to yield colonies when transformed into *S. acidocaldarius E4-39*, indicating that the loop is critical for plasmid replication (Table 2). We further deleted 18 – 20 bp segments (pRSP1-L1 – 3) from the loop, and all of these plasmids failed to replicate in the host as evidenced by absence of transformants (Table 2). Inability of pRSP1 to tolerate deletions within the loop of the 100-bp stem-loop structure further emphasized its role as the potential *cis*-acting double-strand origin of pRN1 (Figure 4). We constructed eight additional plasmids by deleting only 3-bp for loop of the 100-bp stem-loop structure in pRSP1 (including a 'GTG' motif); none of these plasmids (pRSP1-G1 – 8) replicated in the host (Table 2). The inability of any of these three base-pair deletion mutants to

replicate in *S. acidocaldarius* further highlights the importance of the100-bp stem-loop structure in pRN1 replication. Therefore, we propose that the 100-bp stem-loop structure is the double-strand origin of pRN1 (Figure 4).

The *trans*-acting Orf56 and Orf904 can be complemented. Orf56 and Orf904 have been shown to be the only essential genes for pRN1 replication (4, 25). We verified the importance of these co-transcribed *orf56* and *orf904* (5, 19) in pRN1 replication by making in-frame deletions of each in pRSP1 to generate pRSP10 and pRSP9 (Table 2). Both plasmids failed to yield colonies when transformed into in *S. acidocaldarius E4-39* (Table 2), confirming that each of these genes is important for the replication of pRN1 as previously reported (4). A third plasmid, pRSP10b (Table 2), which differs from pRSP10 by the presence of a 75-bp DNA segment upstream of the start codon of *orf904* (Figure 5), also failed to replicate in the host.

We transformed a combination of pRSP9 and pRSP10 or pRSP9 and pRSP10b into *S. acidocaldarius E4-39* to determine if deletions of *orf56* and *orf904* could be complemented by *trans*-acting Orf56 and Orf904 borne on separate plasmids *in vivo*. Our result showed that the pRSP9 – pRSP10b pair yielded colonies indicating complementation of the proteins, while the pRSP9/pRSP10 pair failed to produce transformants (Table 2). The result was consistent on each of the three occasions that the experiment was repeated, although the pRSP9/pRSP10b transformants were generally few in number and grew poorly in liquid medium. However, complementation of pRSP9 and pRSP10b suggests that the 75-bp upstream element of *orf904* (Figure 5) might contain feature(s) that could be important for pRN1 replication.

Translation of the *orf904* could not be achieved by the machinery of *orf56* because of the presence of two in-frame stop codons within 75-bp upstream of its start codon (Figure 5). However, the transcript of *orf904* generated by the *orf56* promoter could be translated by the RBS of *orf904* located within the 75-bp upstream sequence. It is not clear to why the pRSP9/pRSP10 pair did not replicate in the host; we are not ruling out the possibility of a critical *cis*-acting element(s) within the 75-bp sequence upstream of *orf904* (Figure 5). We were not able to identify inverted or tandem repeats within this 75-bp sequence or a potential binding sequence for Orf56.

DISCUSSION

In this study, we sought to identify and map the origin of replication of pRN1 in an effort to improve its utility. pRN1 is one of a few plasmids that have been reproducibly used for *Sulfolobus* transformation (4, 7), but the details of its origin of replication have remained elusive (25). We evaluated the role of a region (241 bp) immediately downstream of *orf904* as the putative origin of replication of pRN1 using a number of shuttle vectors, including pRSP1 and pRSP2 (Figure 1). All of these shuttle vectors replicated in *S. acidocaldarius* (Table 2 and Supplementary Figure 2); however pRSP1 and pRSP3 failed to replicate when the putative of origin was deleted (Table 2), indicating that this region was critical for plasmid replication. The features 241-bp region is consistent with the expected features of the *cis*-acting origin of replication in a plasmid that replicates via a rolling circle mechanism (11, 20), and therefore could function as

pRN1 origin of replication (Figure 4). The putative origin of replication describe in the study was previously reported to be the single-strand origin of pRN1, while the double-strand origin was reported to be located immediately downstream of this structure (22, 29). The ability of pRSP1 and pRSP2 to replicate in the host indicated that *orf56*, *orf904* and the proposed putative origin (Figure 1) were sufficient for pRN1 replication (Table 2). Deletion of any of the three essential components from pRSP1 prevented replication of the plasmid (Table 2), indicating that *orf56*, *orf904* and the putative origin represent the minimum replication unit of pRN1. Our results assert the role of the *cis*-acting origin and affirm previous report of the importance of *orf56* and *orf904* in pRN1 replication (4). Previous reports have suggested that the origin of pRN1 is located downstream of *orf904* (22, 25, 29), but none of these studies experimentally verified the location of the origin. The region that was previously described as the critical double-strand origin (22, 29) was absent in pRSP1 or pRSP2, but each plasmid replicated (Table 2). In addition, the presence of the annotated, double-strand origin in pRSP3-NO, which lacks the proposed *cis-acting* origin of replication, did not replicate the plasmid in *S. acidocaldarius* (Table 2), further confirming that the feature was not essential for pRN1 replication.

The proposed pRN1 origin of replication contains various *cis*-acting elements such as inverted repeats (Figure 4) that are involved in the formation of stem-loop structures associated with a rolling circle replication mechanism (20-21). We identified a 100-bp stem-loop structure within the origin as the double-strand origin of replication (Figure 4) that might be involved in the formation of the replication initiation 'cruciform' (20). Deletions of 3 – 60 bp within the 100-bp stem-loop structure in pRSP1 were not tolerated (Figure 2), whereas deletion of 12 bp outside of this structure did not alter the replication of the plasmid (Table 2). The 58-bp loop of the 100-bp stem-loop structure appeared to be the initiation site for pRN1 replication, because deletions of three nucleotides from this loop in pRSP1 knocked-out the ability of plasmid to replicate in the host (Table 2). The length of the loop appeared to be critical for the replication of the plasmid because single base-pair substitution (A → G) within the loop of the putative double-strand origin in pRSP1 did not alter the replication of the resulting plasmid (data not shown). The loop of the 100-bp stem-loop structure also contains a tri-nucleotide GTG motif (Figure 4) that was reported to be required for primase activity of Orf904 (2). We suspect that the 'nick site' for initiating pRN1 replication could be located a few base pairs 3' of the GTG motif, but we are yet to identify the specific 'nick site'.

The region immediately 5' of the 100-bp stem-loop structure harbors three potential hairpin structures that were deleted in pRSP1d (Supplementary Figure 5). Cells transformed with pRSP1d grew poorly on solid medium and failed to grow in liquid medium, suggesting that the plasmid could not replicate (Supplementary Figure 3). Plasmid instability has been associated with the deletion of single-strand origin from rolling circle replicating plasmids (34). Similarly, the region 3' of the 100-bp stem-loop also contains three potential hair-pin structures that were deleted in pRSP1a; but transformants of this plasmid grew on both solid and liquid media, although poorly when compared to pRSP1 transformants (Supplementary Figure 3). The adjacent regions of the 100-bp stem-loop have features of single-strand origin, but we proposed that the region upstream of the stem-loop to be the single-strand origin, based on its location (20-21) and the severe impact of deleting this region (Supplementary Figure 3). We are not ruling out

the possibility that pRN1 contains two single-strand origins, but we suspect that set of stem-loop structures downstream of the double-strand origin might be a terminator region (*ter*) involved in terminating plasmid replication (Figure 4).

We also showed that *S. acidocaldarius* could be transformed with a combination of pRSP9 and pRSP10b, but not with pRSP9 – pRSP10 pair. We expected to recover transformants with the pRSP9 – pRSP10 pair because the transcription and translation of *orf904* was directed by *orf56* regulatory elements in pRSP10 (Figure 5 and Table 2). However, growth of pRSP9 – pRSP10b transformants were weaker compared to pRSP1 transformants. But the presence of transformants with the pRSP9 – pRSP10b pair showed that Orf56 and Orf904 were functionally expressed, since both genes are essential for replication of the plasmid (4, 25). We suspect that the *trans*-acting Orf56 and Orf904 produced by pRSP9 and pRSP10b, respectively, acted on both plasmids. In addition, the ability of pRSP9 and pRSP10 to produce transformants, suggests that the 75-bp upstream of the start codon of *orf904* might be important for pRN1 replication. The nature or role of this element(s) within this upstream sequence is not fully known at the moment. We suspect that *orf904* was transcribed by the promoter of *orf56* and translated by the putative RBS of *orf904* (Figure 5). However, we are not ruling out that the 75-bp upstream of *orf904* contains *cis*-acting element(s) that might be essential for initiating replication of the plasmid (Figure 5). Transformation of *S. acidocaldarius* with pRSP9 and pRSP10b also highlighted the possibility of transforming *S. acidocaldarius* with multiple plasmids, even those with same origin of replication. The ability of pRSP9 and pRSP10b to replicate in the host strain expands the application of pRN1-based plasmid. However, further work needs to be carried out to improve these tools.

In summary, we have identified and mapped the origin of replication of pRN1, the cryptic plasmid from *S. islandicus REN1H1*. The proposed double-strand origin of the plasmid was identified as the most critical *cis*-acting element for replication of the plasmid. We also proposed a potential single-strand origin for the replication of the lagging strand, but we did not identify the 'nick site' within the loop of the double-strand origin. We also demonstrated that two pRN1-based shuttle vectors can be co-transformed in *S. acidocaldarius* and selected on a single auxotrophic marker by splitting the essential *trans*-acting Orfs56/904. Ability to complement pRN1-based vectors in *S. acidocaldarius* makes the plasmid a powerful tool for studying the *Sulfolobus* species. To enhance the utility of these shuttle vectors, we constructed two additional plasmids pRSP1-CL and pRSP2-CL by adding multiple cloning sites to pRSP1 and pRSP2 respectively (Supplementary Figure 6). Further studies need to be performed to fully understand the detailed mechanism of replication of this plasmid. The role of the regions up- and downstream of the 100-bp stem-loop in plasmid replication and their impact on growth of deletion mutants in liquid media need to be investigated.

Acknowledgments
We thank Dr. Dennis Grogan at the University of Cincinnati for providing the *E. coli* strain ER1821-pM.EsaBC4I and for providing us with the know-how on *S. acidocaldarius* transformation. We also thank Dr Georg Lipps at the University of

Applied Sciences, Northwestern Switzerland for providing the shuttle vector pC. We thank Malik Tahir for his contribution at the early part of this work. This work was part of the DOE Joint BioEnergy Institute (http://www.jbei.org) supported by the U. S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U. S. Department of Energy.

REFERENCES

1. **Aagaard, C., H. Phan, S. Trevisanato, and R. A. Garrett.** 1994. A spontaneous point mutation in the single 23S rRNA gene of the thermophilic arachaeon Sulfolobus acidocaldarius confers multiple drug resistance. Journal of Bacteriology **176:**7744-7747.
2. **Beck, K., and G. Lipps.** 2007. Properties of an unusual DNA primase from an archaeal plasmid. Nucleic Acids Research **35:**5635-5645.
3. **Beck, K., A. Vannini, P. Cramer, and G. Lipps.** 2010. The archaeo-eukaryotic primase of plasmid pRN1 requires a helix bundle domain for faithful primer synthesis. Nucleic Acids Research **38:**6707-6718.
4. **Berkner, S., D. Grogan, S.-V. Albers, and G. Lipps.** 2007. Small multicopy, non-integrative shuttle vectors based on the plasmid pRN1 for *Sulfolobus acidocaldarius* and *Sulfolobus solfataricus*, model organisms of the (cren-)archaea. Nucleic Acids Research. **35:**e88-.
5. **Berkner, S., and G. Lipps.** 2007. Characterization of the Transcriptional Activity of the Cryptic Plasmid pRN1 from Sulfolobus islandicus REN1H1 and Regulation of Its Replication Operon. J. Bacteriol. **189:**1711-1721.
6. **Berkner, S., and G. Lipps.** 2008. Genetic tools for Sulfolobus spp.: vectors and first applications. Archives of Microbiology **190:**217-230.
7. **Berkner, S., A. Wlodkowski, S.-V. Albers, and G. Lipps.** 2010. Inducible and constitutive promoters for genetic systems in <i>Sulfolobus acidocaldarius</i&gt. Extremophiles **14:**249-259.
8. **Cammarano, P., A. Teichner, P. Londei, M. Acca, B. Nicolaus, J. L. Sanz, and R. Amils.** 1985. Insensitivity of archaebacterial ribosomes to protein synthesis inhibitors. Evolutionary implications. The EMBO journal **4:**811-6.
9. **Cannio, R., P. Contursi, M. Rossi, and S. Bartolucci.** 1998. An Autonomously Replicating Transforming Vector for Sulfolobus solfataricus. J. Bacteriol. **180:**3237-3240.
10. **Chen, L., K. Brugger, M. Skovgaard, P. Redder, Q. She, E. Torarinsson, B. Greve, M. Awayez, A. Zibat, H.-P. Klenk, and R. A. Garrett.** 2005. The Genome of *Sulfolobus acidocaldarius*, a Model Organism of the Crenarchaeota. Journal of Bacteriology **187:**4992-4999.
11. **del Solar, G., R. Giraldo, M. J. Ruiz-Echevarría, M. Espinosa, and R. Díaz-Orejas.** 1998. Replication and Control of Circular Bacterial Plasmids. Microbiology and Molecular Biology Reviews **62:**434-464.
12. **Gibson, D. G., L. Young, R.-Y. Chuang, J. C. Venter, C. A. Hutchison, and H. O. Smith.** 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Meth **6:**343-345.

13. **Grogan, D.** 2003. Cytosine methylation by the SuaI restriction-modification system: implications for genetic fidelity in a hyperthermophilic archaeon. Journal of Bacteriology **185:**4657-4661.
14. **Grogan, D. W.** 1989. Phenotypic characterization of the archaebacterial genus *Sulfolobus* : comparison of five wild-type strains. Journal of Bacteriology **171:**6710 - 6719.
15. **Gruss, A. D., H. F. Ross, and R. P. Novick.** 1987. Functional analysis of a palindromic sequence required for normal replication of several staphylococcal plasmids. Proceedings of the National Academy of Sciences **84:**2165-2169.
16. **Hatzinikolaou, D. G., E. Kalogeris, P. Christakopoulos, D. Kekos, and B. J. Macris.** 2001. Comparative growth studies of the extreme thermophile *Sulfolobus acidocaldarius* in submerged and solidified substrate cultures. World Journal of Microbiology and Biotechnology **17:**229-234.
17. **Jonuscheit, M., E. Martusewitsch, K. M. Stedman, and C. Schleper.** 2003. A reporter gene system for the hyperthermophilic archaeon Sulfolobus solfataricus based on a selectable and integrative shuttle vector. Molecular Microbiology **48:**1241-1252.
18. **Joshua, C. J., R. Dahl, P. I. Benke, and J. D. Keasling.** 2011. Absence of Diauxie during Simultaneous Utilization of Glucose and Xylose by Sulfolobus acidocaldarius. J. Bacteriol. **193:**1293-1301.
19. **Keeling, P. J., H.-P. Klenk, R. K. Singh, O. Feeley, C. Schleper, W. Zillig, W. F. Doolittle, and C. W. Sensen.** 1996. Complete Nucleotide Sequence of theSulfolobus islandicusMulticopy Plasmid pRN1. Plasmid **35:**141-144.
20. **Khan, S.** 1997. Rolling-circle replication of bacterial plasmids. Microbiol. Mol. Biol. Rev. **61:**442-455.
21. **Khan, S. A.** 2005. Plasmid rolling-circle replication: highlights of two decades of research. Plasmid **53:**126-136.
22. **Kletzin, A., A. Lieke, T. Urich, R. L. Charlebois, and C. W. Sensen.** 1999. Molecular Analysis of pDL10 From Acidianus ambivalens Reveals a Family of Related Plasmids From Extremely Thermophilic and Acidophilic Archaea. Genetics **152:**1307-1314.
23. **Kumar, R., S. Singh, and O. Singh.** 2008. Bioconversion of lignocellulosic biomass: biochemical and molecular perspectives. Journal of Industrial Microbiology and Biotechnology **35:**377-391.
24. **Kurosawa, N., and D. W. Grogan.** 2005. Homologous recombination of exogenous DNA with the *Sulfolobus acidocaldarius* genome: Properties and uses. FEMS Microbiology Letters **253:**141-149.
25. **Lipps, G.** 2009. Molecular biology of the pRN1 plasmid from Sulfolobus islandicus. Bichemical Society Transactions **37:**42 - 45.
26. **Lipps, G., M. Stegert, and G. Krauss.** 2001. Thermostable and site-specific DNA binding of the gene product ORF56 from the Sulfolobus islandicus plasmid pRN1, a putative archael plasmid copy control protein. Nucleic Acids Research **29:**904-913.
27. **Magasanik, B.** 1961. Catabolite repression. Cold Spring Harbor Symposia on Quantitative Biology **26:**249-56.

28. **Mamie Z, L., and E. Stephen J.** 2007. SLIC sub-cloning using T4 DNA polymerase treated inserts with RecA.
29. **Peng, X., I. Holz, W. Zillig, R. A. Garrett, and Q. She.** 2000. Evolution of the family of pRN plasmids and their integrase-mediated insertion into the chromosome of the crenarchaeon Sulfolobus solfataricus. Journal of Molecular Biology **303:**449-454.
30. **Reilly, M. S., and D. W. Grogan.** 2001. Characterization of Intragenic Recombination in a Hyperthermophilic Archaeon via Conjugational DNA Exchange. J. Bacteriol. **183:**2943-2946.
31. **Sanchez, M., M. Drechsler, H. Stark, and G. Lipps.** 2009. DNA translocation activity of the multifunctional replication protein ORF904 from the archaeal plasmid pRN1. Nucleic Acids Research **37:**6831-6848.
32. **Schelert, J., V. Dixit, V. Hoang, J. Simbahan, M. Drozda, and P. Blum.** 2004. Occurrence and Characterization of Mercury Resistance in the Hyperthermophilic Archaeon Sulfolobus solfataricus by Use of Gene Disruption. J. Bacteriol. **186:**427-437.
33. **She, Q., R. K. Singh, F. Confalonieri, Y. Zivanovic, G. Allard, M. J. Awayez, C. C. Chan-Weiher, I. G. Clausen, B. A. Curtis, and A. De Moors.** 2001. The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2. Proceedings of the National Academy of Sciences of the United States of America **98:**7835 - 7840.
34. **Solar, G., G. Kramer, S. Ballester, and M. Espinosa.** 1993. Replication of the promiscuous plasmid pLSI: a region encompassing the minus origin of replication is associated with stable plasmid inheritance. Molecular and General Genetics MGG **241:**97-105.
35. **Zillig, W., Kletzin, A., Schleper, C., Holtz, I., Janekovic, D., Hain, J., Lanzendorfer, M. and Kristiansson. J. K.** 1994. Screening for Sulfolobales, their plasmids, and their viruses in Islandic solfataras. . Syst. Appl. Microbiol. **16:**606-628.
36. **Zuker, M.** 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Research **31:**3406-3415.

TABLES

Table 1. Description of plasmids and strains

Plasmid/Strains	Description of important features*	Source/ Reference
Plasmids		
pUC19	pMB1 origin, AmpR, (2.7 kb)	
pC	pRN1 (entire plasmid); pBluescript; *pyrEF* (9.0 kb)	3
pRSP1	pRN1 (*orfs56*/*904*), AmpR, *pyrE* and *lacS* (7.2 kb)	This study
pRSP1-NO	4681 – 4912 deleted from pRSP1 (6.9 bp)	This study
pRSP1a	4853 – 4912 deleted from pRSP1 (7.1 kb)	This study
pRSP1b	4793 – 4852 deleted from pRSP1 (7.1 kb)	This study
pRSP1c	4733 – 4792 deleted from pRSP1 (7.1 kb))	This study
pRSP1d	4681 – 4732 deleted from pRSP1 (7.1 kb)kb)	This study
pRSP1e	4844 – 4855 deleted from pRSP1 (7.2 kb))	This study
pRSP1-L1	4793 – 4812 deleted from pRSP1 (7.1 kb)	This study
pRSP1-L2	4773 – 4792 deleted from pRSP1 (7.1 kb)	This study
pRSP1-L3	4755 – 4772 deleted from pRSP1 (7.1 kb)	This study
pRSP1-L4	4755 – 4812 deleted from pRSP1 (7.1 kb)	This study
pRSP1-G1	4773 – 4775 deleted from pRSP1 (7.2 kb)	This study
pRSP1-G2	4776 – 4778 deleted from pRSP1 (7.2 kb)	This study
pRSP1-G3	4779 – 4781 deleted from pRSP1 (7.2 kb)	This study
pRSP1-G4	4782 – 4784 deleted from pRSP1 (7.2 kb)	This study
pRSP1-G5	4785 – 4787 deleted from pRSP1 (7.2 kb)	This study
pRSP1-G6	4788 – 4790 deleted from pRSP1 (7.2 kb)	This study
pRSP1-G7	4791 – 4793 deleted from pRSP1 (7.2 kb)	This study
pRSP1-G8	4794 – 4796 deleted from pRSP1 (7.2 kb)	This study
pRSP1-CL	pRSP1with XhoI-AatII-NruI-PacI-PstI-PvuII-NsiI site (7.2 kb)	This study
pRSP2	*lacS* deleted from pRSP1(5.7 kb)	This study
pRSP2--CL	pRSP2 with HindIII-XhoI-AatII-NruI-PacI-PstI-PvuII-NsiI sites (5.7 kb)	This study
pRSP3	pRN1 (*orfs56*/*904*), AmpR, *pyrE* and *lacS* (7.6 kb)	This study
pRSP3-NO	4707 – 4912 deleted from pRSP3 (7.4 kb)	This study

Table 1 continued. Description of plasmids and strains

Plasmid/Strains	Description of important features*	Source/ Reference
Plasmids		
pRSP3	RSP3fragment, AmpR, *pyrE* and *lacS* (7.6 kb)	This study
pRSP3-NO	4707 – 4912 deleted from pRSP3 (7.4 kb)	This study
pRSP4	*lacS* deleted from pRSP3 (6.1 kb)	This study
pRSP5	pRN1 (*orfs56/904/80*), AmpR, *pyrE* and *lacS* (8.1 kb)	This study
pRSP6	*lacS* deleted from pRSP5 (6.6 kb)	This study
pRSP7	pRN1 (*orfs56/904/80/90a/72*), AmpR, *pyrE* and *lacS* (8.8 kb)	This study
pRSP9	*orf904* (1986 – 4680) deleted from pRSP1(4.5 kb)	This study
pRSP10	*orf56* (1815 – 1965) deleted from pRSP1 (7.0 kb)	This study
pRSP10b	Portion of *orf56* (1815 – 1890) deleted from pRSP1 (7.1 kb)	This study
pGlcSTUV_ko	Source of pyrE::lacS cassette (6.9 kb)	Unpublished
Strains		
GX3	*S. acidocaldarius* (DSMZ 639) utilizes glucose and galactose)	14
E4-39	Uracil auxotroph with 17 bp duplication in pyrE	This study
E4-RSP1	Strain E4-39 harboring pRSP1	This study
E4-RSP1-CL	Strain E4-39 harboring pRSP1-CL	This study
E4-RSP2	Strain E4-39 harboring pRSP2	This study
E4-RSP2-CL	Strain E4-39 harboring pRSP2-CL	This study
E4-RSP3	Strain E4-39 harboring pRSP3	This study
E4-RSP4	Strain E4-39 harboring pRSP4	This study
E4-RSP5	Strain E4-39 harboring pRSP5	This study
E4-RSP6	Strain E4-39 harboring pRSP6	This study
E4-RSP7	Strain E4-39 harboring pRSP7	This study

**pyrE, pyrF* and *lacS* are from *S. solfataricus P2*; AmpR gene obtained from pUC19

Table 2. Transformation of *S. acidocaldarius E4-39*

Plasmid(s)	Deletion[#]	Affected feature	Transformants[a]
pRSP1	None		++++
pRSP2	None		+++*
pRSP3	None		+++
pRSP4	None		+++*
pRSP5	None		++
pRSP6	None		+++*
pRSP7	None		++
pRSP1a	4853 – 4912	Putative origin of replication in pRSP1	+++[b]
pRSP1b	4793 – 4852		No
pRSP1c	4733 – 4792		No
pRSP1d	4681 – 4732		+[b]
pRSP1e	4844 – 4855		+++
pRSP1-NO	4681 – 4912		No
pRSP3-NO	4707 – 4912	Putative origin of replication in pRSP3	No
pRSP1-L1	4793 – 4812	100-bp stem-loop of putative origin of replication in pRSP1	No
pRSP1-L2	4773 - 4792		No
pRSP1-L3	4755 - 4772		No
pRSP1-L4	4755 - 4812		No
pRSP1-G1	4773 - 4775	Loop of 100-bp stem-loop structure	No
pRSP1-G2	4776 - 4778		No
pRSP1-G3	4779 - 4781		No
pRSP1-G4	4782 - 4784		No
pRSP1-G5	4785 - 4787		No
pRSP1-G6	4788 - 4790		No
pRSP1-G7	4791 - 4793		No
pRSP1-G8	4794 - 4796		No

Table 2 continued. Transformation of *S. acidocaldarius E4-39*

Plasmid(s)	**Deletion**	**Affected feature**	**Transformants**[a]
pRSP9	Deleted 1986 - 4680	orf904	No
pRSP10	Deleted 1815 - 1965	orf56	No
pRSP10b	Deleted 1815 – 1890	orf56	No
pRSP9 + pRSP10			No
pRSP9 + pRSP10b			+

[a] Apparent transformation efficiency (+ – ++++) ranged from 10^2 - >10^4cfu/ug of plasmid (efficiency varies with experiments)

[b] Colonies are very tiny compared to transformants from pRSP1

* Transformants were classified as medium to small colonies; large colonies are revertants

[#] Numbers denote positions of nucleotides deleted from the plasmids

FIGURES

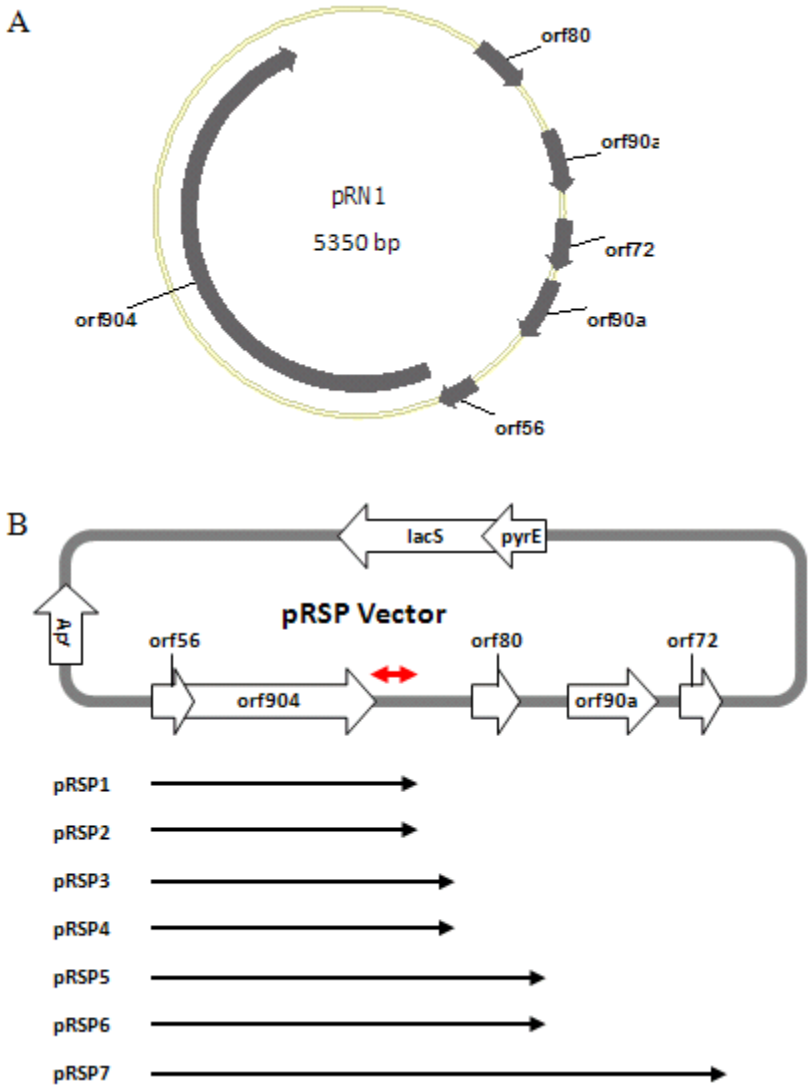
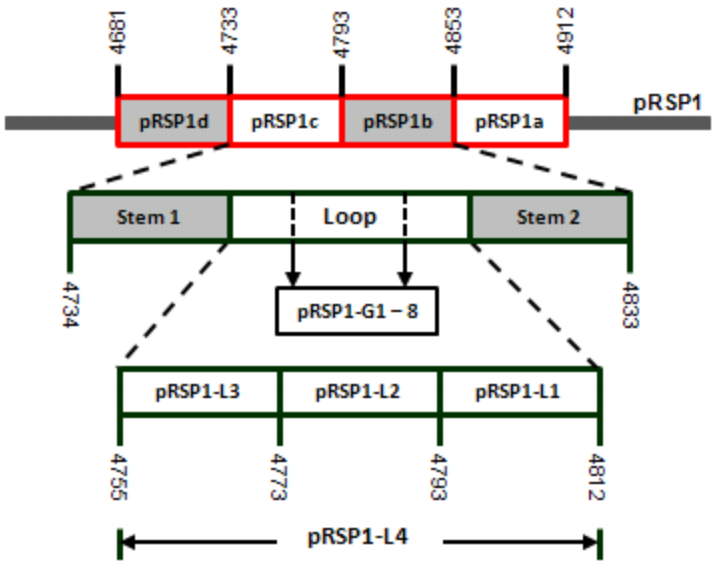


Figure 1. Strategy for shuttle vector construction. (A) Map of pRN1 showing features of pRN1. (B) Map of shuttle vector(s) used for evaluating the role of a 241 bp region downstream of *orf904* as pRN1 origin of replication (red double arrow). Shuttle vectors harbor pUC19 Ap^r gene and origin (red box); lacS-pyrE cassette; and a segment of pRN1 highlighted by black arrows; *lacS* was deleted from pRSP1, 3 and 5 to construct pRSP2, 4 and 6.

A



B

pRSP1	TACACGGTGGGACAATAATATAAT
pRSP1-G1	- - - ACGGTGGGACAATAATATAAT
pRSP1-G2	TAC - - - GTGGGACAATAATATAAT
pRSP1-G3	TACACG - - - GGACAATAATATAAT
pRSP1-G4	TACACGGTG - - - CAATAATATAAT
pRSP1-G5	TACACGGTGGGA- - - TAATATAAT
pRSP1-G6	TACACGGTGGGACAA- - -TATAAT
pRSP1-G7	TACACGGTGGGACAATAA- - -AAT
pRSP1-G8	TACACGGTGGGACAATAATAT-- -

Figure 2. Mapping pRN1 putative origin of replication. (A)Various deletions were made within the 'putative ori' in pRSP1 to generate plasmids highlighted in each rectangle box (i) to determine the role of the deleted regions pRN1 replication. (B) Three nucleotide deletion mutants of pRSP1 were constructed to map the loop of stem-loop structure within pRN1 origin. Each construct was transformed into *S. acidocaldarius* to determine plasmid replication.

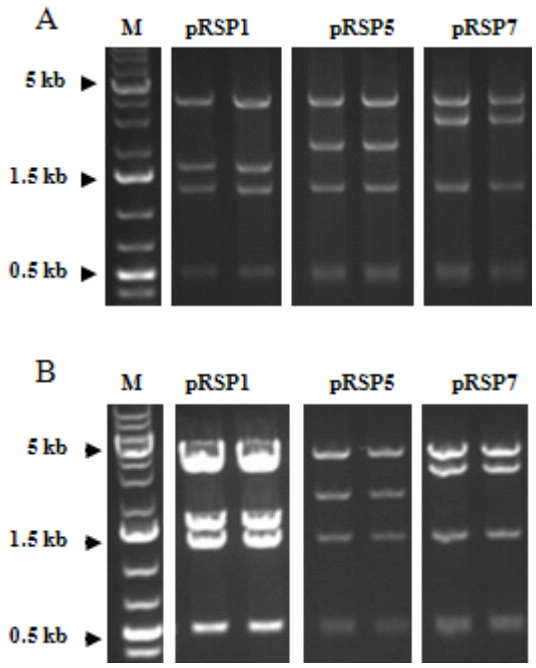


Figure 3. Plasmid extraction from *S. acidocaldarius* transformants. (A) Plasmids were extracted from *S. acidocaldarius* strains harboring pRSP1, 5 and 7 and digested with EcoRI to determine the DNA fingerprint of each plasmid. (B) EcoRI fingerprint was also determined after plasmids extracted from *S. acidocaldarius* were re-transformed into *E. coli*.

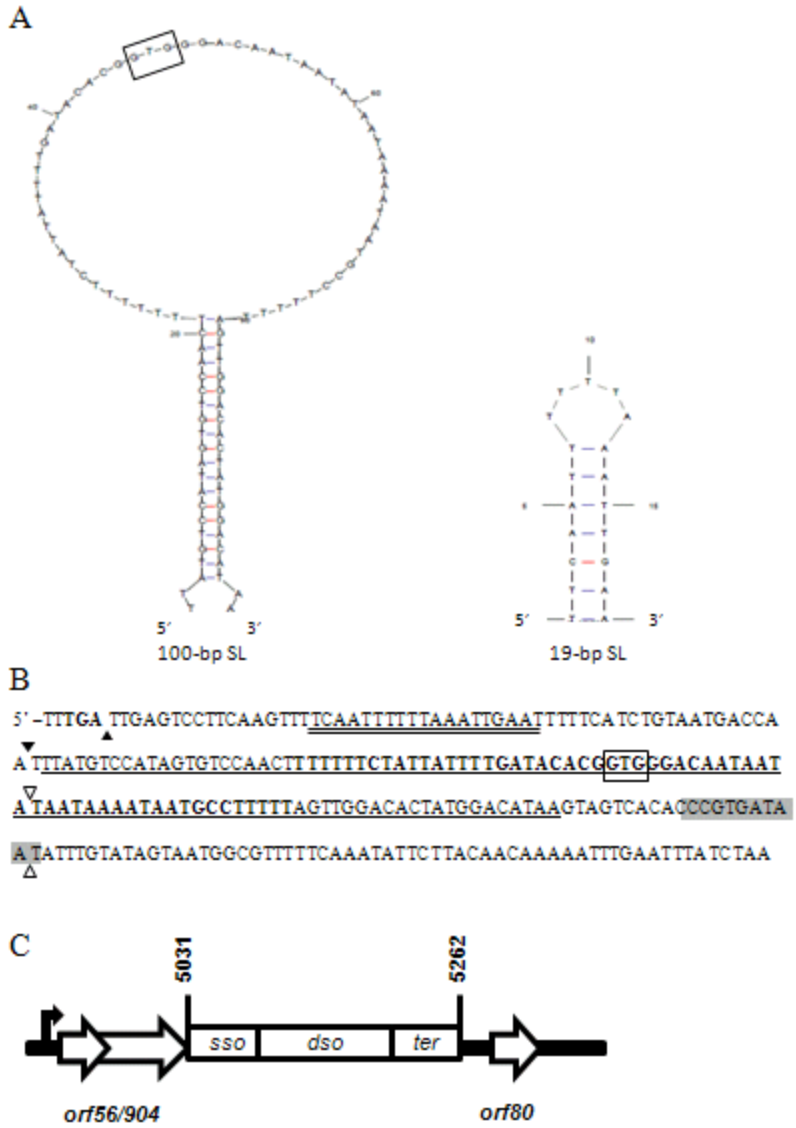


Figure 4. Stem-loop structures in pRN1 origin of replication. The secondary structure of pRN1 origin of replication revealed the presence of a 100-bp and 19-bp stem-loop (SL) structures at70ºC (A). The secondary structure was determine using DNA mFold. DNA sequence showing the 100-bp SL (single underline) and 19-bp SL (double underline) in pRN1 origin (B). The fragments deleted in pRSP1a (Δ), pRSP1b (▽), pRSP1c (▼) and pRSP1d (▲) are separated by space; the region deleted in pRSP1e is shaded. The 'GTG' motif (box) and the loop of the 100-bp SL (bolded) are also highlighted. The 100-bp SL is proposed to be the double-strand origin (*dso*) of pRN1 and is flanked by the proposed single-strand origin (*sso*) and replication termination site (*ter*) at the 5' and 3' ends respectively (C).

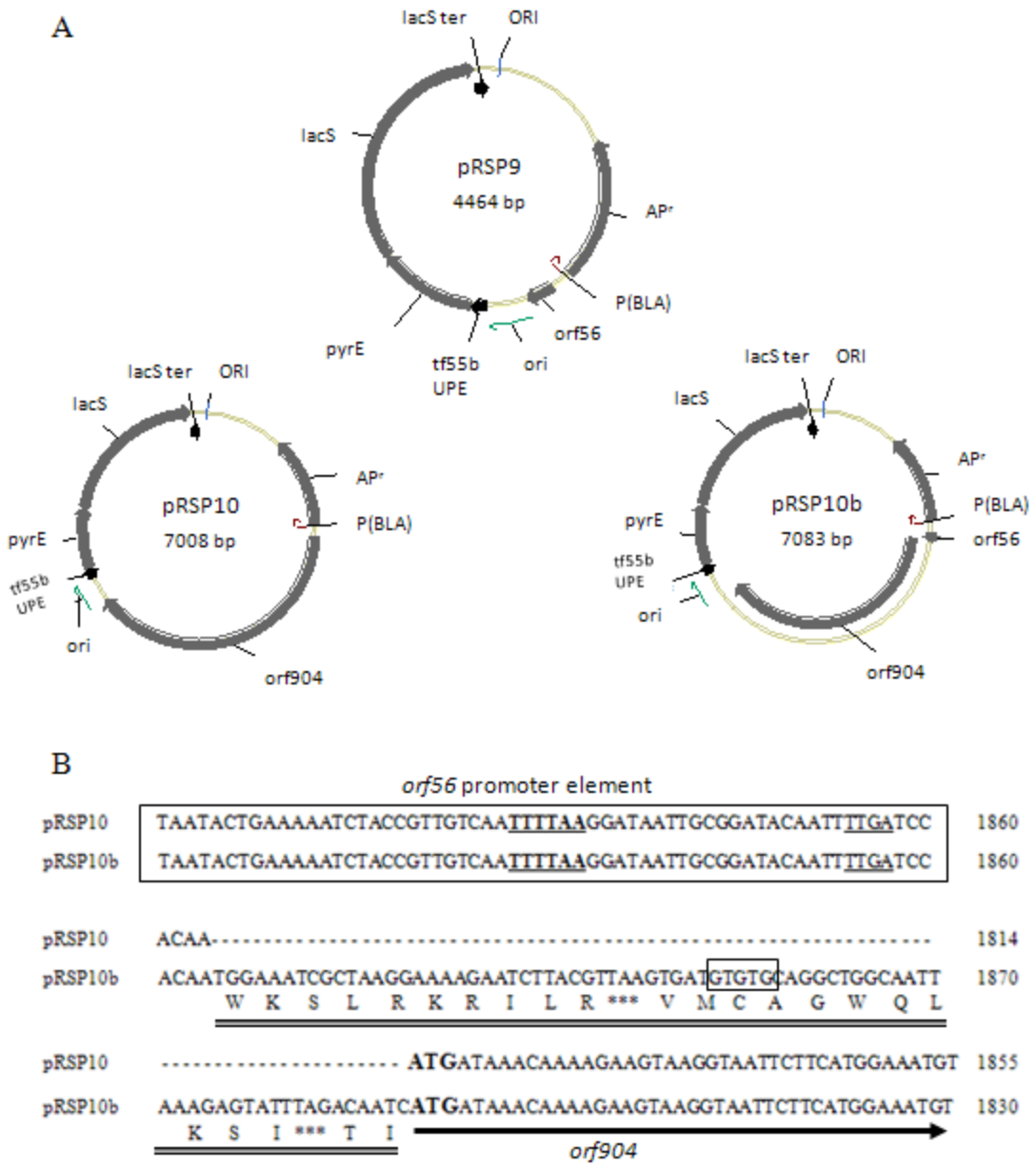


Figure 5. Effect of *orf56* and *orf904* in pRN1 replication. Two shuttle vectors, pRSP9 and pRSP10 constructed by deleting *orf904* and *orf56* respectively from pRN1 to determine the role of each in pRN1 replication; each gene was deleted from start to stop codon (A). A third pRSP10b was constructed by deleting only a portion of *orf56* leaving a 75 bp sequence upstream of *orf904* start codon. Each plasmid contains pUC19 and pRN1 origins of replication; pyrE-lacS cassette transcribed by *S. solfataricus thsB* promoter (tf55b-UPE) and *lacS* terminator (lacS-ter); and AP^r gene (ampicillin resistance). Transcription of *orf904* in pRSP10 and pRSP10b is directed by *orf56* promoter elements (B); box A (bold single underline) and box B (single underline). 75 bp upstream of *orf904* in pRSP10b (double underline) contains a putative RBS (box) and two *in-frame* stop codons. Start of *orf904* is bolded.

SUPPLEMENTARY MATERIALS

Table 1. Primers

P/N	Name	Length (bp)	Amplicon	Sequence
117	stuv-ss-lacS-RT	20	pyrE-lacS fragment	AAAGTATTGCAATCTAATGA
120	stuv-ss-tf55b-upe-F	18	pyrE-lacS fragment	AAACGCTTAACATTACTT
128	RUC-F	50	pUC19 fragment	TCTAATCTCATTTTCATTAGATTGCAATACTTTAAAGAACATGTGAGCAA
129	RUC-R	50	pUC19 fragment	AAATTGACAACGGTAGATTTTTCAGTATTATATGTATCCGCTCATGAGAC
130	RC-F	18	RSP1, 3, 5 and 7 fragments	TAATACTGAAAAATCTAC
131	RC-R1	50	RSP1 fragment	AAGTTTATCGAAAAGTAATGTTAAGCGTTTGAATGACGTTAGATAAATTC
132	RC-R3	50	RSP3 fragment	AAGTTTATCGAAAAGTAATGTTAAGCGTTTTTAATCATATCTGTGATTGT
133	RC-R5	48	RSP5 fragment	AAGTTTATCGAAAAGTAATGTTAAGCGTTTGTATCATTATGTATCGAT
134	RC-R7	50	RSP7 fragment	AAGTTTATCGAAAAGTAATGTTAAGCGTTTCGGCATAAGTAGGAATAGTT
135	RS-E-F	50	pRSP2, 4 and 6	GATTATTTGGTGAAGAATGTTGAAAAGTAGACTTTCTCAAGTCTCACTAT
136	RS-E-R	20	pRSP2, 4 and 6	CTACTTTTCAACATTCTTCA
137	pCN-F	50	pRSP3-NO	TTGAGTCCTTCAAGTTTTCAATTTTTTAAACGTCATTCTCTCTATAGAAC
138	pCN-R	28	pRSP3-NO	TTTAAAAAATTGAAAACTTGAAGGACTC
160	RSP9-F	17	pRSP9	GTCCTTCAAGTTTTCAA
161	RSP9-R	49	pRSP9	TCAATTTAAAAAATTGAAAACTTGAAGGACTTACTTCTTTTGTTTATCA
162	RSP10-F	50	pRSP10	GATAATTGCGGATACAATTTTGATCCACAAATGATAAACAAAAGAAGTAA
165	RSP1-CLF	58	pRSP1-CL/pRSP2-CL	AGTTTGACGTCGCGATTAATTAACTGCAGCTGATGCATAAACGCTTAACATTACTTTT
166	RSP1-CLR	60	pRSP1-CL/pRSP2-CL	GCTGCAGTTAATTAATCGCGACGTCAAACTCGAGAAGCTTGAATGACGTTAGATAAATTC
167	RSP1a-F	50	pRSP1a	TATGGACATAAGTAGTCACACCCGTGATAACGTCATTCAAACGCTTAACA

Table 1 continued…

P/N	Name	Length (bp)	Amplicon	Sequence
168	RSP1a-R	20	pRSP1a	TTATCACGGGTGTGACTACT
169	RSP1b-F	50	pRSP1b	ATTATTTTGATACACGGTGGGACAATAATA TATTTGTATAGTAATGGCGT
171	RSP1c-F	50	pRSP1c	TAAATTGAATTTTTCATCTGTAATGACCAA TAATAAAATAATGCCTTTTT
172	RSP1c-R	18	pRSP1c	TTGGTCATTACAGATGAA
173	RSP1d-F	50	pRSP1d	GATTGATGAAGAGCACGTGCGGTTTGATTG ATTTATGTCCATAGTGTCCA
179	RSP10-R2	20	pRSP10	GGATCAAAATTGTATCCGCA
181	RSP1e-R	20	pRSP1e	GTGTGACTACTTATGTCCAT
182	RSP10-F2	20	pRSP10	ATGATAAACAAAAGAAGTAA
184	RSP1-CLF2	20	pRSP1-CL/pRSP2-CL	AAACGCTTAACATTACTTTT
186	RSP1a-F2	20	pRSP1a/pRSP1-NO	CGTCATTCAAACGCTTAACA
187	RSP1c-F2	20	pRSP1c	TAATAAAATAATGCCTTTTT
189	RSP1e-F2	18	pRSP1e	TTGTATAGTAATGGCGTT
190	RSP1-NO-R	21	pRSP1-NO	TCAATCAAACCGCACGTGCTC
191	RSP1-L1F	22	pRSP1-L1/pRSP1-L4	AGTTGGACACTATGGACATAAG
192	RSP1-L1R	22	pRSP1-L1	TATTATTGTCCCACCGTGTATC
193	RSP1-L2F	25	pRSP1-L2	TAATAAAATAATGCCTTTTTAGTTG
194	RSP1-L2R	23	pRSP1-L2	TCAAAATAATAGAAAAAAAGTTG
195	RSP1-L3F	20	pRSP1-L3	TACACGGTGGGACAATAATA
196	RSP1-L3R	18	pRSP1-L3/pRSP1-L4	AGTTGGACACTATGGACA
197	RSP1-A1-F	15	pRSP1-A1	GGCAATAATATAATA
198	RSP1-A2-F	15	pRSP1-A2	GACGATAATATAATA
199	RSP1-A3-F	15	pRSP1-A3	GACAGTAATATAATA
203	RSP1-A7-F	15	pRSP1-A7	GACAATAATATGATA
205	RSP1-A1-8-R	15	pRSP1-A1 - pRSP1-A7	CCACCGTGTATCAAA
206	RSP1-G1-R	23	pRSP1-G1 and 2	TCAAAATAATAGAAAAAAAGTTG
207	RSP1-G1-F	18	pRSP1-G1	ACGGTGGGACAATAATAT

Table 1 continued…

P/N	Name	Length (bp)	Amplicon	Sequence
208	RSP1-G2-F	21	pRSP1-G2	TACGTGGGACAATAATATAAT
209	RSP1-G3-R	18	pRSP1-G3 and 4	CGTGTATCAAAATAATAG
210	RSP1-G3-F	21	pRSP1-G3	GGACAATAATATAATAAAATA
211	RSP1-G4-F	24	pRSP1-G4	GTGCAATAATATAATAAAATAATG
212	RSP1-G5-R	17	pRSP1-G5 and 6	TCCCACCGTGTATCAAA
213	RSP1-G5-F	20	pRSP1-G5	TAATATAATAAAATAATGCC
214	RSP1-G6-F	22	pRSP1-G6	CAATATAATAAAATAATGCCTT
215	RSP1-G7-R	20	pRSP1-G7 and 8	TTATTGTCCCACCGTGTATC
216	RSP1-G7-F	24	pRSP1-G7	AATAAAATAATGCCTTTTTAGTTG
217	RSP1-G8-F	24	pRSP1-G8	TATAAAATAATGCCTTTTTAGTTG

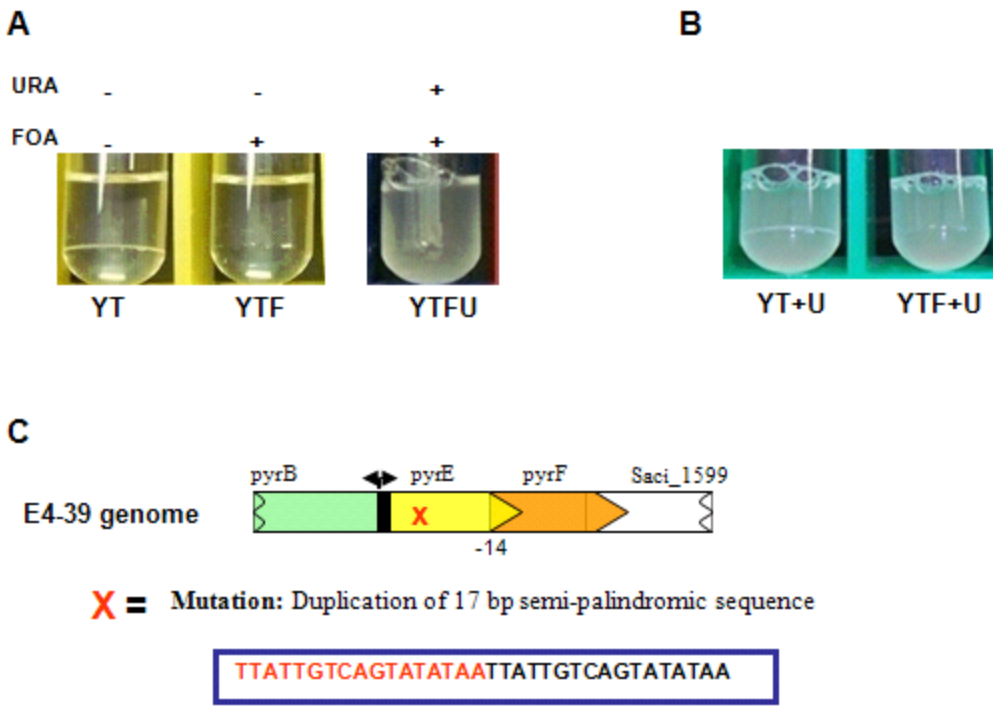


Figure 1. Selection of uracil auxotroph. *S. acidocaldarius* was grown on YT medium supplemented with 5 µg/ml uracil and 50 µg/ml 5-fluoroorotic acid [5-FOA]) at 75°C (200 rpm) and grown for 6 days. (A) Shows the spontaneous uracil auxotrophs from this 6-day culture grew on YTFU (YT with uracil and 5-FOA) after 48 h incubation but not on YT or YTF (YT with 5-FOA, but not uracil). (B) Cell grew on YT and YTF tubes after uracil was added and grown for additional 48 h. (C) Strain E4-39 with duplication of a 17-bp (blue and red) sequence within *pyrE* was selected as candidate after characterization of the mutants.

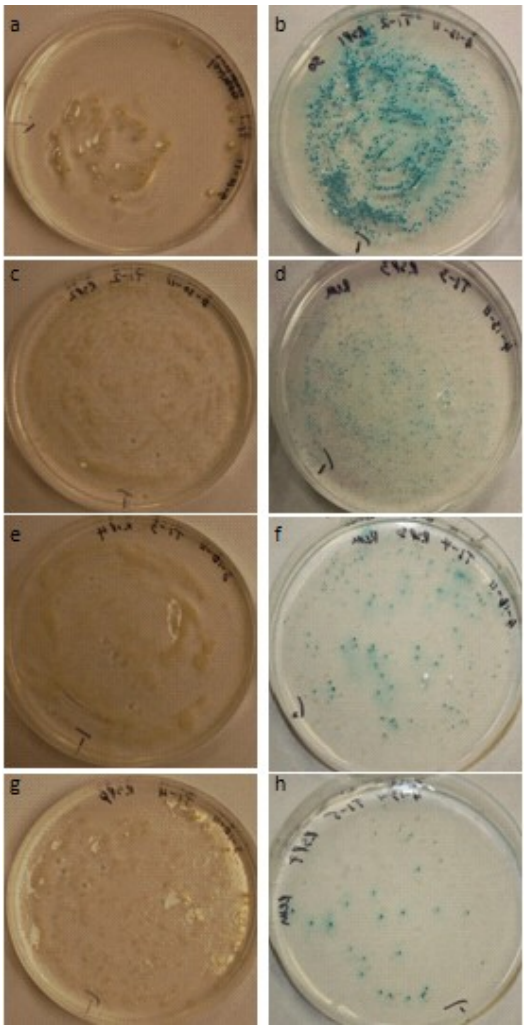


Figure 2. Transformation of pRSP1 – 7 into *S. acidocaldarius E4-39*. All of the shuttle vectors harboring pRN1 fragments replicated in *S. acidocaldarius E4-39* . Cells transformed with pRSP1 (b), pRSP3 (d), pRSP5 (f) and pRSP7 (h) yielded blue colonies because of the presence of LacS, while those pRSP2 (c), pRSP4 (e) and pRSP6 (g) transformants retained WT phenotype because of the absence of LacS. Large WT colonies seen on all plates represent revertants and are easily distinguished from the transformants based on their size. The cells were transformed by electroporation and no plasmid was added to the control sample (a).

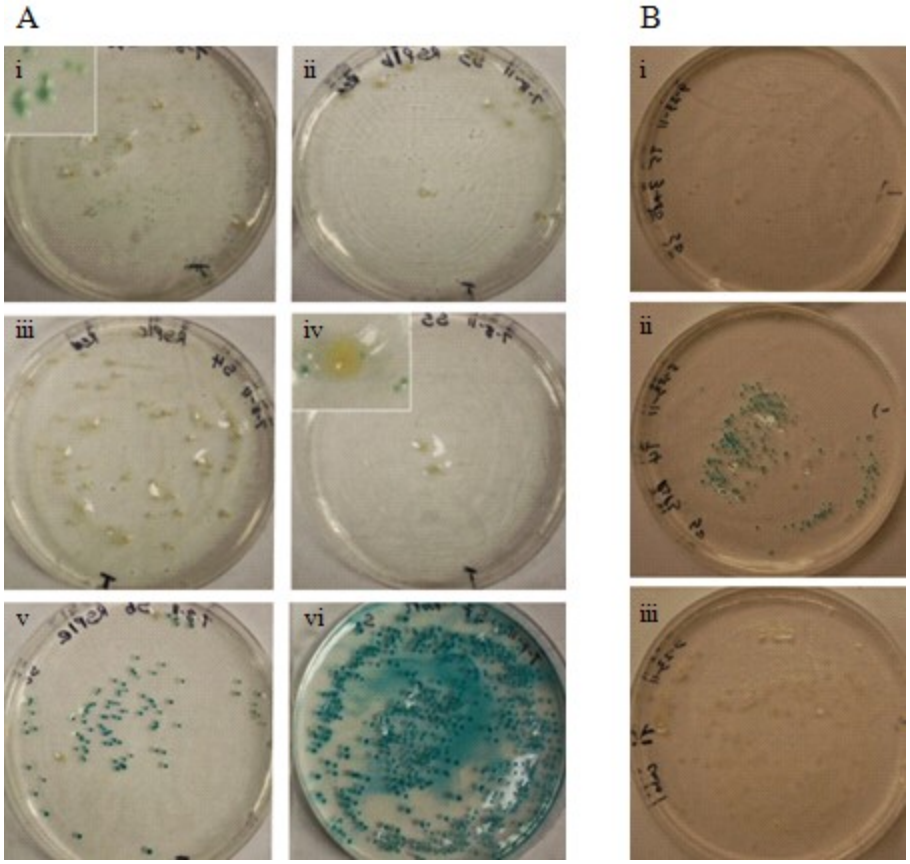


Figure 3. Analyzing pRN1 origin of replication. Regions of pRN1 putative origins were deleted from pRSP1 and transformed into *S. acidocaldarius E4-39* (A). Plates show cells with pRSP1a (i), pRSP1b (ii), pRSP1c (iii), pRSP1d (iv), pRSP1e (v) and pRSP1 (vi). Inserts in (i) and (ii) highlight enlarged tiny blue colonies from pRSP1a and pRSP1d respectively. The role pRN1 putative origin in pRSP3 replication in *S.acidocaldarius E4-39* (B) by transforming the strain with: no plasmid (i), pRSP3 (ii) and pRSP3-NO which lacks the entire putative origin except 35 bp at the 5' end (iii). Similar result was obtained pRSP1-NO, which lacks the entire putative origin except 9 bp at the 5' end (not shown).

```
pDL10          -----------------CACGCGGGG--------TTTTCAATTTTTTGAAATTCAATTTT
pRN2           --------------CAACACGCGGGGT-------TTTACAGTTTTTTGAAATTCAATTTT
pRN1-ori       ---------------TTTGATTGAGTCCTTCAAGTTTTCAATTTTTT-AAATTGAATTTT
pHEN7          TATATTAAACCTTTTTGTTTTCTAGATTGACATTTTTTCAATTTTTCTAAATTGAATTTT
                                         *          *** ** *****  ***** ******

pDL10          TAAACTGTAATGTACAAATAAGGTACCGAAAGTCCAAGAACATGGGGGATATA---TTTT
pRN2           TAATCTATAATGACCAATTATCGGACCGATAGGACAGGAA-------ATCATATATTTTT
pRN1-ori       TCATCTGTAATGACCAATTTATGTCCATAGTGTCC-----------------AACT
pHEN7          TCATCTGTCATTACCAATTATGTCACCACT-----------------GTCACAACTTTTT
               * * ** * **   *** *      *                        *   ****

pDL10          TAACACAATTTTTTTGCA------GACTTTTTTGGGGTAAAAATAAACCCTAAAACCTGG
pRN2           TAACACGGTTTTTGGGT-------GAGTTTTATAGGGGTAAAAAATAGTGTGTTTCCTGT
pRN1-ori                                                              AGTTGG
pHEN7          TTCTATTATATGGCTACACGCGTGTAGACAAAATGAATAAAATAATAAAAAAATAGTGGT
               *   *   * *               *              ***  *             *

pDL10          ACTTTCGATCCCTTATTCAAATATA---------CCGTGATAATATTACAATGCTCATG-
pRN2           CCTATCGGTCCGTTTTTTGCGTAAG---------CCGTGTTAATATCTGTATAGTAATGG
pRN1-ori       AC---------ACTATGGACATAAGTAGTCACACCCGTGATAATATTTGTATAGTAATGG
pHEN7          GTCAGTGGT---------GACTAAATAGTCACCACCGTGATAATATTTGTATAGTAATGG
                                     **            ***** ******     **  * ***

pDL10          ----TTACAA-----TTGAATGCAAAATACTGAATATGTCTAACGTATTTTCTCTTCATA
pRN2           CGTTTTTCGAATTTTTTTAGTAAGAATATTTGAATTTATCTATCGC-TTTTCTCTCTCTA
pRN1-ori       CGTTTTTCAAATATTCTTACAACAAAAATTTGAATTTATCTAA----------------
pHEN7          CGTTTTTCAAATATTCCTACAACAAAAAATTGAATTTGTCTATCGT-CATTCTCTTCATC
                   ** * *         *      **    ***** * ****
```

Figure 4. Alignment putative origin of replication of pRN1 with members of the pRN family of plasmid. The putative origin of pRN1 aligned with regions down of the replication protein in pRN2, pDL10 and pHEN7. The loop of the 100-bp stem-loop (underlined) is shaded gray. The consensus sequence with 'GTG' motif is highlighted (box). The putative origin region of pRN1 corresponds to the predicted single-strand origin of pHEN7 and pDL10.

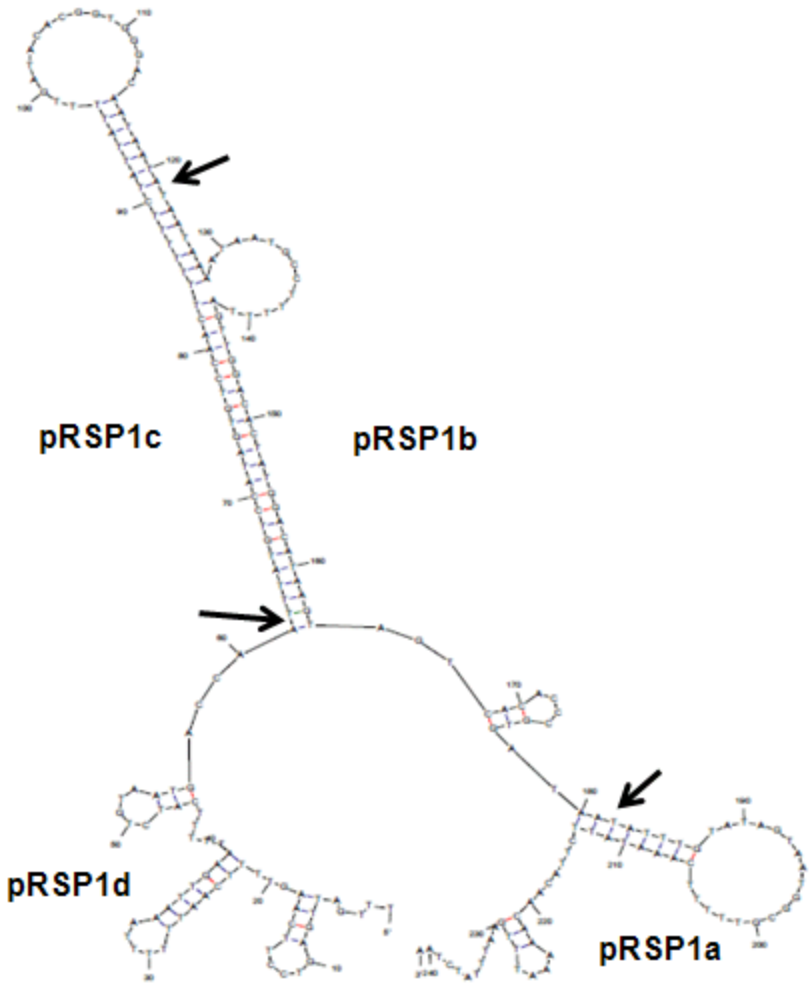


Figure 5. Secondary structure of the putative origin of replication at 37°C. The mFold structure highlights potential hair-pin structures at the 5' and 3' ends of the putative origin. The regions deleted in pRSP1a – pRSP1d is also highlighted (arrows).

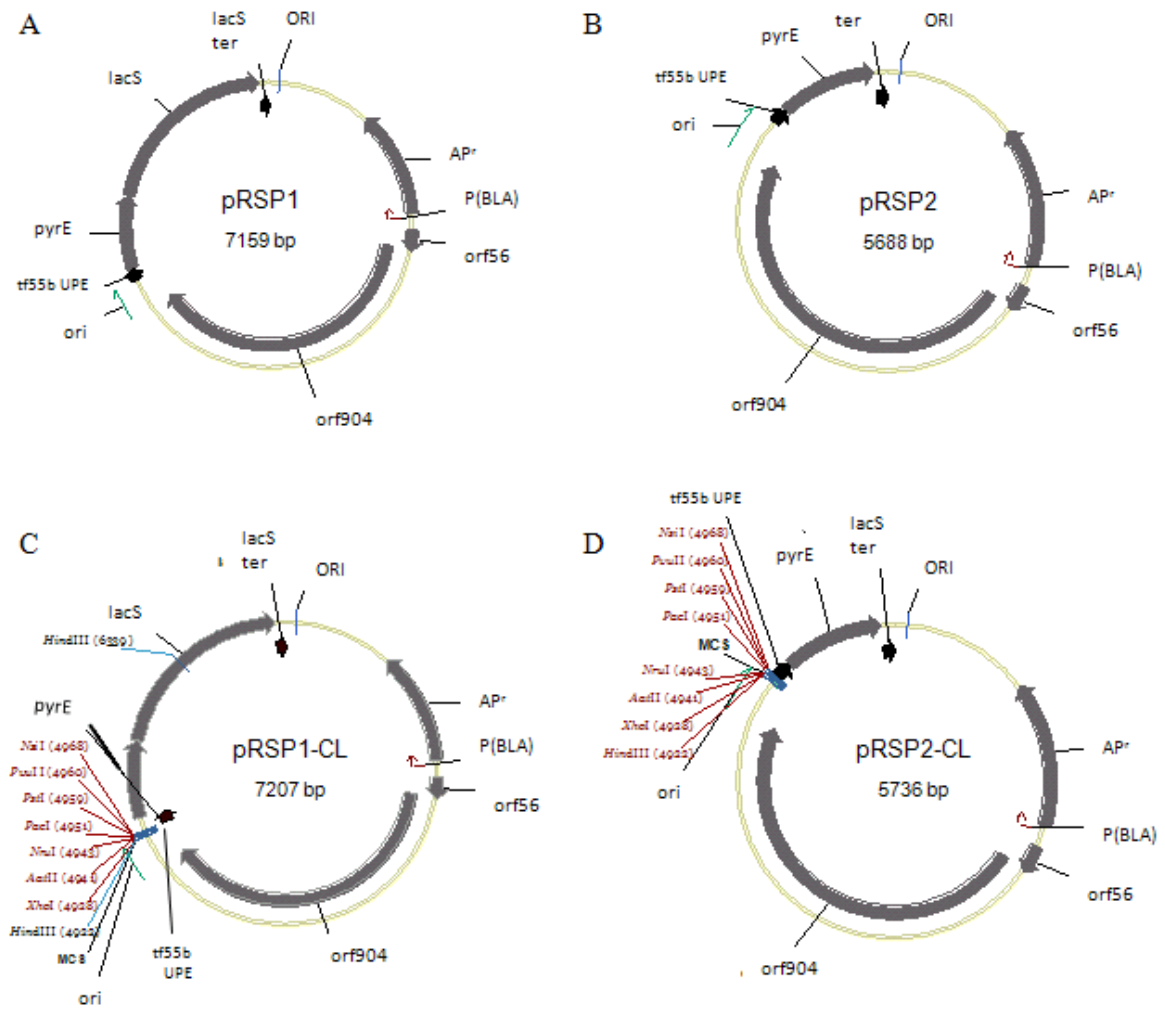


Figure 6. Map of pRSP1-CL and pRSP2-CL. Multiple cloning site made up of HindIII-XhoI-AatII-NruI-PacI-PstI-PvuII-NsiI (MCS) was added to pRSP1 (A) and pRSP2 (B) to generate pRSP1-CL (C) and pRSP2-CL (D) to enhance its utility for cloning.

Chapter Four: Effect of Heterologous Beta-glycosidase (LacS) on Cellobiose Metabolism in *Sulfolobus acidocaldarius*

ABSTRACT

Sulfolobus acidocaldarius lacks cellobiose and lactose metabolism which is present in most *Sulfolobus* species. Cellobiose metabolism in *S. solfataricus* is catalyzed by a β-glycosidase that is encoded by *lacS*, but this gene was not annotated in *S. acidocaldarius*. Although *S. acidocaldarius* harbors a homologue of *lacS* called *bgaS*, which encodes a β-galactosidase, it lacks the ability to cleave X-gal. To investigate if the absence of a functional b β-glycosidase was responsible for the absence of cellobiose metabolism in *S. acidocaldarius*, we heterologously expressed *lacS* from *S. solfataricus* in the organism. *S. acidocaldarius* expressing *lacS* was able to cleave X-gal but failed to grow on cellobiose or lactose. Our result suggested that absence of cellobiose metabolism in *S. acidocaldarius* could be due to absence of a transporter for the sugar. This study also provided us an opportunity to develop genetic tools such as auxotrophic strains and recombinant strains for the studying the organism.

INTRODUCTION

Sulfolobus acidocaldarius is an aerobic hyperthermo-acidophilic crenarchaea that grows optimally at 75 – 80°C and at pH of 2 – 3 ((6, 8). The organism is currently being considered a good candidate for the production of biofuel from ligno-cellulose (cellulosic) biomass, because of its ability to tolerate extreme growth conditions and simultaneously utilize sugars like glucose, xylose, galactose and arabinose simultaneously (23).

Sustainable development of ligno-cellulose-based (cellulosic) biofuels as alternative renewable energy sources requires a microbial system that can efficiently catabolize most cellulosic sugars, especially glucose and xylose. Current cellulosic biofuel technology in effective largely due to high cost of pre-treatment employed to release fermentable sugars and inefficient conversion of these sugars into the desired biofuels by the fermenting microbes (32, 39). Unfortunately, the metabolism of sugars like xylose and cellobiose is repressed by glucose-induced catabolite repression in common metabolic engineering strains (13, 26, 28, 31). Although, glucose-induced catabolite repression in absent in *S. acidocaldarius*, the organism does not metabolize cellobiose (18, 22), a major intermediate in cellulose hydrolysis (26, 32). Cellobiose is hydrolyzed into two glucose moieties that are subsequently metabolized for growth and biofuel production (24, 26) .

Lack of cellobiose metabolism is peculiar to *S. acidocaldarius* because the sugar is readily metabolized by other *Sulfolobus* species (18, 22). The initial step in the catabolism of cellobiose has been report to be catalyzed by beta glycosidase (EC 3.2.1.21) which is encoded by *lacS* (Sso3019) in *S. solfataricus* (1, 10). LacS has also been shown to exhibit similar activities against lactose and cellobiose (17); inactivation of *lacS* in *S. solfataricus* terminated the ability of the organism to utilize lactose and cellobiose (21). Surprisingly, the genome of *S. acidocaldarius* harbors a *lacS* homologue called *bgaS* (Saci_1849); the protein that was annotated as beta-galactosidase, share greater than 70% amino acid identity with LacS (11). *S. acidocaldarius* does not produce blue colonies with X-gal (5-bromo-4-chloro-indolyl-β-D-galactopyranoside) like other *Sulfolobus* species, despite the presence of *bgaS* in its genome. Expression of LacS by means of a plasmid resulted in heterogeneous population of *S. acidocaldarius* with respect to cleavage of X-gal (2). However, BgaS was recently shown to cleave lactose, cellobiose and a list of analogs *in vitro* (33), but the study did not address the inability of the organism to metabolize or cleave X-gal *in vivo* (33). Cellobiose metabolism is very important if *S. acidocaldarius* would be given serious consideration as a platform for cellulosic biofuel production.

Absence of cellobiose metabolism in *S. acidocaldarius* might be due to lack of active beta-glucosidase to hydrolyze cellobiose or inability to import the sugar. Unlike *S. solfataricus* which harbors genes encoding cellobiose transporter (8, 37), no cellobiose or lactose transporter was annotated in the genome of *S. acidocaldarius*. However, absence of cellobiose transporter does not necessarily imply that the organism cannot import the sugar because the organism metabolizes xylose and arabinose despite absence of annotated or confirmed transporters (8, 23). In this study, we shall attempt to verify that absence of cellobiose metabolism in the organism was not due to its inability to cleave cellobiose into glucose molecules. We would be evaluating the effect LacS from *S. solfataricus* on cellobiose metabolism in *S. acidocaldarius*. The *lacS* would be inserted into the genome of the organism to ensure stable expression of the gene and to avoid the heterogeneous population that was previously reported with expression of the gene via

plasmid (2). Stable expression of functional LacS in *S. acidocaldarius* could potentially provide valuable insight on the impact of the β-glycosidase on cellobiose metabolism in the organism. In addition, developing strategy for inserting *lacS* into the genome of *S. acidocaldarius* would enable us to improve the genetic tool for studying these species, which is highly under-developed (3). Development genetic tools for *S. acidocaldarius* would be of great of importance for engineering novel metabolic capabilities in the organism.

Although genetic tools for *Sulfolobus* species have great improvement in the past few years (2, 4, 12), inter-laboratory reproducibility of these tools remains inadequate. Besides, application of *Sulfolobus* species for metabolic engineering is grossly undermine the existence of active mobile elements the result in unstable genomes (30, 37). The genome of *S. acidocaldarius* is free of these active mobile elements, which is why is it essential to establish cellobiose metabolism in the organism (8). However, *S. acidocaldarius* contain a restriction endonuclease (*SuaI*) that digests unprotected foreign DNA; hence, protection of any vector or DNA fragment that is introduced into the organism is very essential. Transformation strategy has improved over the years for the species; heterologous gene expression gene in these organisms is still limited by availability of autonomously replicating vector system and effective selection strategy (3, 27, 36). So far the most effective selection strategies are based on auxotrophies such uracil and lactose auxotrophies (4, 12, 38).

Selection based on auxotrophy requires the creation of an auxotrophic strain, in the case of uracil auxotroph, it usually involves the creation of spontaneous mutants (19). Spontaneous uracil auxotrophs have been successfully selected using 5'-fluoroorotic acid (5-FOA) an analog of orotic acid that inhibits growth of uracil prototrophic wild type microbial strains (5, 25). The toxicity of 5-FOA is exerted when the molecule is metabolize and the process is initiated by orotate phosphoribosyltransferase (OPRT) encoded by *pyrE* and orotidine-5'-phosphate decarboxylase (OPD) encoded by *pyrF* (20). Mechanism of FOA involves the inhibition of thymidylate synthase by 5-fluoro-deoxy uridine monophosphate, an intermediate of FOA metabolism (9, 35). Strains lacking OPRT or OPD can grow in medium supplemented with uracil because they are unable to metabolize 5-FOA; hence are insensitive to the analog.

In this study, we will be isolating spontaneous uracil auxotrophic strains of *S. acidocaldarius* via FOA selection. The isolated uracil auxotroph(s) will serve as hosts for evaluating the impact of LacS on cellobiose metabolism in *S. acidocaldarius* organism. This study will provide an avenue for us to further improve the genetic tool for studying *Sulfolobus* species.

MATERIALS AND METHODS

Strains and growth media.

Sulfolobus acidocaldarius 639 used in this study was obtained from DSMZ (Braunschweig, Germany). The organism was cultivated aerobically at 75ºC and 200 rpm in ATCC #1723 (YT) medium (pH = 2.5). The organism was cultivated on solid medium that consist of YT or T (YT without yeast extract) at 70°C and each medium was solidified with 0.8% gellan gum (Spectrum or Sigma). The optical density of the liquid culture was measured at 600 nm (OD_{600}) using a Beckman-Coulter DU-800 UV/Visible spectrophotometer. Each growth medium was filter sterilized through filters with 0.2-µM pore size and membrane filters. Reagents and Chemicals

used in this study were obtained from Sigma, Spectrum, Fisher scientific, New England Biolabs and Fermentas.

Analysis of cellobiose metabolism

S. acidocaldarius was grown sugars in minimal medium (NYT) that consist of YT without yeast extract and tryptone. The NYT medium was supplemented with either 1.0 g/L cellobiose, 1.0 g/L xylose or a combination of 1.0 g/L of both sugars in 50 ml cultures. Prior to inoculation, the organism was acclimatized to each sugar by growing it in YT supplemented with the sugar(s) of interest. Growth of the organism was monitored by measuring the optical density at 600 nm. To monitor the residual sugar in the growth medium, the organism was grown in NYT medium supplemented with 2.0 g/L of xylose and a combination of 1.0 g/L each of the following sugars: xylose and cellobiose (XCb); xylose and lactose (XL); xylose and glucose (XG); xylose, glucose and cellobiose (XGCb); and xylose, glucose and lactose (XGL). 1.0 ml of each culture was withdrawn after 48 h of growth and pelleted at 17,900 × g for 1 min. The supernatant from each sample was filtered through 0.2-µm syringe filters and analyzed for residual sugar. The samples were analyzed using an Agilent 1100 Series Binary LC System equipped with an Agilent 1200 series refractive index detector. The samples were passed isocratically through Bio-Rad Aminex HPX-87H 300x7.8 column (with Micro-Guard Cation H Refill Cartridges, 30 x 4.6 mm) at a flow-rate of 0.6 ml. The mobile phase consists of 4 mM H_2SO_4 in HPLC grade purity water (Honeywell Burdick & Jackson, Morristown, NJ). The temperature of the column was maintained at 50°C, respectively.

Selection of Spontaneous uracil auxotrophs.

Spontaneous uracil auxotroph used in the study was selected by growing *S. acidocaldarius* in liquid YTFU medium which consist of YT medium supplemented with 50 µg/ml 5-Fluoroorotic acid (5-FOA) and 5 µg/ml uracil at 75°C (200 rpm). After 6 days of incubation, the cells were sub-cultured into the following media: YT, YTFU (20 µg/ml uracil + 50 µg/ml 5-FOA) and YTF (with 50 µg/ml 5-FOA; no uracil) to verify the presence of spontaneous uracil auxotrophs. Uracil auxotrophs were further verified and characterized by growing the mutants on: YT, YTFU, YTF and YTU (YT with 20 µg/ml uracil; no 5-FOA). After several passages on these plates, genomic DNA was extracted from selected strains (uracil auxotrophs) on YTFU plates for sequence analysis of the *pyrEF* genes. The mutations on the *pyrEF* genes were characterized.

Plasmid construction.

To insert *S. solfataricus lacS* (Sso3019) into the genome of *S. acidocaldarius*, we constructed suicide vector pESF-S. The vector was constructed by inserting *lacS* between *S. acidocaldarius pyrE* and *pyrF* (Figure 1) as part of the operon driven by the promoter of *pyrE*. The translation of *lacS* and *pyrF* was placed under the control of the RBS of *pyrF*; *lacS* gene was kept *in-frame* by the addition of an adenine "A" nucleotide between its start codon and the stop codon of *pyrE*. Each fragment used in the construction of pESF-S was amplified with primers with overhanging sequence that overlap the adjacent fragment by at least 30 bp (Supplementary Table 1). The *pyrE* fragment contains 210 bp upstream of the gene's start codon that includes the native promoter of *pyrEF* operon. The *pyrF* fragment contains 52 bp upstream of its start codon (RBS element) and

250 bp downstream of its stop codon. The additional sequence at the 5' and 3' ends *pyrE* and *pyrF* respectively, were added to in to increase chances of homologous recombination with the genome of *S. acidocaldarius*. The complete plasmid was assembled by inserting the *pyrE::lacS::pyrF* fragment into the KpnI and SphI sites of pUC19 at (Figure 1).

The complete plasmid was assembled using the Sequence and Ligation Independent Cloning (SLIC) strategy (29). The fragments (inserts and linear pUC19 backbone) used for the construction of pESF-S were treated independently with 0.5 µl of T4 polymerase at room temperature for 30 min to chew-back the 3' ends of the fragments. Each reaction was stopped by the addition of approximately 1 mM dGTP. The T4 polymerase treated fragments were combined in molar ratio of 1.5:1 (inserts: pUC19) and incubated at 37°C for 30 min in 10 µl reaction; the control experiment consisted of the pUC19 backbone alone. 5 µl of each cloning reactions was transformed into chemically competent *E. coli* DH10B and the transformants were analyzed.

Methylation and transformation of *S. acidocaldarius*

The suicide vector was methylated by transforming into *E. coli* 1821 (New England Biolabs) harboring pM.EsaBC4I which was kindly given to us by Dr. D. Grogan. Methylation of the plasmid protects it against *SuaI* endonuclease, a native restriction enzyme in *S. acidocaldarius* (15). The methylated plasmid was transformed into electro-competent *S. acidocaldarius* E4-39 1% sucrose solution; the competent cells were prepared following procedure "G" in Kurosawa and Grogan (26). Electroporation (1.25 KV, 50 µF and 750 Ω) was carried out in 1 mm cuvette and the cells were recovered with equal volume of a 2X recovery buffer (1% sucrose, 20 mM β-alanine/1.5 mM malate buffer [pH = 4.5] and 10 mM $MgSO_4$) as described by Kurosawa and Grogan (26). The recovered cells were incubated at 70°C for 30 min, cooled on ice for 2 – 5 min, spread on YT or T plates and incubated at 70°C in humidified Sterilite™ plastic containers. The competent cells were prepared.

X-gal staining and characterization of transformants

Transformants were sprayed with 5 mg/ml X-gal (5-bromo-4-chloro-3-indolyl-b-D-galactoside) dissolved in dimethylformamide (DMF) and incubated for at 70°C for at least 60 min. Blue colonies on transformation plates were scored as positive transformants and selected for further characterization. Genomic DNA from selected transformants was analyzed by PCR to determine the presence *lacS* in the genome of the organism and the mode integration. The *pyrE::lacS::pyrF* fragment was amplified form genome DNA samples using forward primer (P/N 35) for *pyrE* and reverse primer (P/N 38) for *pyrF*. We also amplified pUC19 origin of replication from the genomic DNA samples to determine the mode of insertion of the *pyrE::lacS::pyrF* into the genome. Confirmed strain was grown in YT supplemented with cellobiose or lactose to acclimatized the strains to the respectively sugar. The ability of strain harboring *lacS* to utilized cellobiose or lactose as sole carbon and energy source was analyzed as described earlier.

RESULT AND DISCUSSION

Cellobiose metabolism in *S. acidocaldarius*

We evaluated cellobiose metabolism in *S. acidocaldarius* by growing the organism on sugar or in combination with other sugars as co-carbon sources after a period of acclimatization. The organism failed to grow on 2.0 g/L cellobiose as sole carbon source confirming previous reports ((18, 22), but grew on 2.0 g/L xylose and a mixture of 1.0 g/L each of xylose and cellobiose with specific growth rate of 0.083 and 0.087 h^{-1} respectively (Figure 2A). To verify that the organism only utilized xylose during growth on a mixture of xylose and cellobiose, we analyzed the residual in the growth medium found that the organism did not utilize cellobiose (Figure 3). Inability of *S. acidocaldarius* to consume cellobiose suggested that metabolism of the sugar is impaired in the organism. Impairment of cellobiose metabolism could result from lack of cellobiose transporter or absence of functional β-glucosidase (or β-glycosidase). Cellobiose metabolism in *S. solfataricus* is believed to be catalyzed by β-glycosidase encoded by *lacS* because inactivation of this gene in *S. solfataricus* terminated the ability of the organism to utilize cellobiose and lactose (21). Cellobiose is hydrolyzed into two glucose moieties which are subsequently catabolized via glycolytic pathways. We postulated that expression of β-glycosidase (LacS) from *S. solfataricus* would restore cellobiose metabolism in *S. acidocaldarius* LacS in *S. acidocaldarius*; provided the organism possesses a transporter for the sugar. We based our assumption on the fact that glucose metabolism in present in *S. acidocaldarius* (18, 22); and sugar transporters are poorly annotated in the organism (8). To heterologously express LacS in *S. acidocaldarius*, we have to develop genetic tools which are poorly developed for the organism. Development of tool for studying this organism is important, considering that the organism is being evaluated as a potential candidate for cellulosic biofuel production.

Isolation of *S. acidocaldarius E4-39*

Auxotrophic recipient hosts deficient in key metabolic pathways are commonly used in the study of *Sulfolobus* species, because the species are usually not sensitive antibiotics (7). A number of uracil auxotrophs have been successfully used to study different *Sulfolobus* species (3). We sought to obtain spontaneous uracil auxotrophic strains of *S. acidocaldarius* by exposing the organism to 5-FOA in liquid and analyzed resulting culture for mutation in *pyrEF*. We exposed wild-type strain of *S. acidocaldarius* to 5-FOA for six days to enrich the population of spontaneous uracil auxotrophs. Prior to enrichment, spontaneous auxotrophs were not observed on YTFU plate; but the population of apparent uracil auxotrophs increased to approximately 1.1×10^7 after exposing the organism to 5-FOA (Table 1). We also observed colonies on YT plates after exposing the organism to 5-FOA indicating that some of the apparent mutants reverted to wild-type phenotype. The mutations in the spontaneous auxotrophs were most present in *pyrE* (Saci_1597); and the mutations ranged from base-pair substitutions (transition and transversion) to frame-shift mutations, in accordance with previous report (19). Most of the spontaneous mutants have a frame-shift mutation that resulted from the insertion of a 'T' at position 55. Most of the mutants were fairly stable, but they do occasionally revert to wild-type. We selected strain 'E4-39' with 17-bp duplication within *pyrE* (Figure 4) to serve as recipient for studying *in vivo* activity of LacS in the organism.

Evaluation of LacS activity in *S. acidocaldarius*

We established β-glycosidase activity in *S. acidocaldarius E4-39* by heterologously expressing LacS from *S. solfataricus* (Figure 5A). The gene encoding LacS was inserted in the chromosome of *S. acidocaldarius E4-39* via suicide vector pESF-S (Figure 1). Prior to transformation, the plasmid was methylated to protect it against *SuaI*, a restriction endonuclease present in *S. acidocaldarius* (16, 34). After 5 – 7 days of incubation on YT plates, colonies that strained blue when exposed to X-gal were selected and characterized for presence of *lacS* in chromosome of the organism (Figure 5A). The ability of these transformants to cleave X-gal presumptively confirms that LacS is active in the host. X-gal was not cleaves by cells transformed with pEbgaSF (data not shown), which was constructed by substituting *lacS* in pESF-S with *bgaS* the homologue of the gene in *S. acidocaldarius*.

Our results also revealed that each blue colony that was transformed with pESF-S contained *lacS* in its genome (Figure 5B). However, most of the transformants (>80%) contained the entire suicide vector as confirmed by the presence of origin of replication of pUC19 and two bands approximately 1.2 and 2.7 kb corresponding to the *pyrEF* and *pyrE::lacS::pyrF* (Figure 5B). The presence of pESF-S in the genome of the transformants indicates the occurrence of single recombination event (Figure 5). In strains mutated by a double recombination event, the *pyrE::lacS::pyrF* cassette replaced the native *pyrEF* resulting in the amplification of a single band (≈ 2.7 kb) with primers that anneals at the 5' and 3' ends of *pyrE* and *pyrF* respectively (Figure 5B). Regardless of mode of insertion, the ability of the transformants to cleave X-gal was stably maintained, indicating the presence of β-glycosidase activity (4). Heterologous expression of *lacS* in *S. acidocaldarius* via pJ-lacS, a shuttle vector resulted in strains that are heterogeneous and unstable in terms of their ability to cleave X-gal (2). We verified unstable expression *lacS* in *S. acidocaldarius* using pC-lacSb, a shuttle vector that is very similar to pJ-lacS .

***S. acidocaldarius ESF* does not growth on cellobiose**

We selected and grew one of the transformants, *S. acidocaldarius ESF*, harboring the *pyrE::lacS::pyrF* cassette on YT medium supplemented with either cellobiose or lactose to acclimatize it to the sugars. The *pyrE::lacS::pyrF* cassette was inserted into the strain's chromosome via double recombination. After several rounds of acclimatization, we evaluated the correlation between X-gal cleavage and cellobiose metabolism, and found that the strain does grow on either sugar as a sole source of carbon (Figure 2B). The inability of the *S. acidocaldarius ESF* to grow cellobiose despite its ability to hydrolyze X-gal suggests that the organism could not import the sugar as previously observed with the wild-type (Figure 3). We anticipated that cellobiose would be hydrolyzed into two glucose moieties by LacS, if the sugar was imported in the cytoplasm of the organism. LacS has been shown to be important for cellobiose metabolism in *S. solfataricus* (21). Unlike *S. solfataricus*, a transporter for cellobiose has not been described for *S. acidocaldarius* (14, 37); but absence of a transporter for a sugar in the organism does not imply absence of that sugar metabolism. *S. acidocaldarius* metabolizes xylose and arabinose even though no transporter has been confirmed for these sugars (8, 22-23). It will be important to confirm that the organism does not uptake cellobiose via uptake assay before making a definite conclusion on the transport of the sugar in the organism. However, our

results suggest that absence of cellobiose in the organism is most likely due to its inability to import the sugar.

CONCLUSION

In this study, we functionally expressed β-glycosidase encoded by *S. solfataricus lacS* in *S. acidocaldarius* with the aim to establish cellobiose metabolism; but the resulting strain did not grow on the sugar. Our results suggested that absence of cellobiose metabolism in *S. acidocaldarius* could be due to absence of transporter for the sugar. Absence of cellobiose metabolism undermines the potential application of the organism as a platform for cellulosic biofuel engineering. Therefore, to establish we would first verify that the organism does import the sugar; and subsequent introduce a thermophilic cellobiose transporter. This study provided us with additional genetic tools such auxotrophic strains and strategy for homologous recombination in the organism.

Acknowledgments
We thank Dr. Dennis Grogan at the University of Cincinnati for providing the *E. coli* strain ER1821-pM.EsaBC4I and for providing us with the know-how on *S. acidocaldarius* transformation. This work was part of the DOE Joint BioEnergy Institute (http://www.jbei.org) supported by the U. S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U. S. Department of Energy.

REFERENCES

1. **Bartolucci, S., M. Rossi, and R. Cannio.** 2003. Characterization and Functional Complementation of a Nonlethal Deletion in the Chromosome of a {beta}-Glycosidase Mutant of Sulfolobus solfataricus. J. Bacteriol. **185:**3948-3957.
2. **Berkner, S., D. Grogan, S.-V. Albers, and G. Lipps.** 2007. Small multicopy, non-integrative shuttle vectors based on the plasmid pRN1 for *Sulfolobus acidocaldarius* and *Sulfolobus solfataricus*, model organisms of the (cren-)archaea. Nucleic Acids Research. **35:**e88-.
3. **Berkner, S., and G. Lipps.** 2008. Genetic tools for Sulfolobus spp.: vectors and first applications. Archives of Microbiology **190:**217-230.
4. **Berkner, S., A. Wlodkowski, S.-V. Albers, and G. Lipps.** 2010. Inducible and constitutive promoters for genetic systems in <i>Sulfolobus acidocaldarius</i&gt. Extremophiles **14:**249-259.
5. **Boeke, J. D., F. Croute, and G. R. Fink.** 1984. A positive selection for mutants lacking orotidine-5′-phosphate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance. Molecular and General Genetics MGG **197:**345-346.

6. **Brock, T. D., K. M. Brock, R. T. Belly, and R. L. Weiss.** 1972. Sulfolobus: A new genus of sulfur-oxidizing bacteria living at low pH and high temperature. Archives of Microbiology **84:**54-68.

7. **Cammarano, P., A. Teichner, P. Londei, M. Acca, B. Nicolaus, J. L. Sanz, and R. Amils.** 1985. Insensitivity of archaebacterial ribosomes to protein synthesis inhibitors. Evolutionary implications. The EMBO journal **4:**811-6.

8. **Chen, L., K. Brugger, M. Skovgaard, P. Redder, Q. She, E. Torarinsson, B. Greve, M. Awayez, A. Zibat, H.-P. Klenk, and R. A. Garrett.** 2005. The Genome of *Sulfolobus acidocaldarius*, a Model Organism of the Crenarchaeota. Journal of Bacteriology **187:**4992-4999.

9. **Chouini-Lalanne, N., M. C. Malet-Martino, R. Martino, and G. Michel.** 1989. Study of the metabolism of flucytosine in Aspergillus species by 19F nuclear magnetic resonance spectroscopy. Antimicrobial Agents and Chemotherapy **33:**1939-1945.

10. **Cubellis, M. V., C. Rozzo, P. Montecucchi, and M. Rossi.** 1990. Isolation and sequencing of a new [beta]-galactosidase-encoding archaebacterial gene. Gene **94:**89-94.

11. **Dehal, P. S., M. P. Joachimiak, M. N. Price, J. T. Bates, J. K. Baumohl, D. Chivian, G. D. Friedland, K. H. Huang, K. Keller, P. S. Novichkov, I. L. Dubchak, E. J. Alm, and A. P. Arkin.** 2009. MicrobesOnline: an integrated portal for comparative and functional genomics. Nucleic Acids Research.

12. **Deng, L., H. Zhu, Z. Chen, Y. Liang, and Q. She.** 2009. Unmarked gene deletion and host–vector system for the hyperthermophilic crenarchaeon <i>Sulfolobus islandicus</i&gt. Extremophiles **13:**735-746.

13. **Dien, B. S., M. A. Cotta, and T. W. Jeffries.** 2003. Bacteria engineered for fuel ethanol production: current status. Applied Microbiology and Biotechnology **63:**258-266.

14. **Elferink, M. G. L., S.-V. Albers, W. N. Konings, and A. J. M. Driessen.** 2001. Sugar transport in *Sulfolobus solfataricus* is mediated by two families of binding protein-dependent ABC transporters. Molecular Microbiology **39:**1494-1503.

15. **Grogan, D.** 2003. Cytosine methylation by the SuaI restriction-modification system: implications for genetic fidelity in a hyperthermophilic archaeon. Journal of Bacteriology **185:**4657-4661.

16. **Grogan, D. W.** 2003. Cytosine Methylation by the SuaI Restriction-Modification System: Implications for Genetic Fidelity in a Hyperthermophilic Archaeon. J. Bacteriol. **185:**4657-4661.

17. **Grogan, D. W.** 1991. Evidence that β-Galactosidase of Sulfolobus solfataricus Is Only One of Several Activities of a Thermostable β-d-Glycosidase. Applied and Environmental Microbiology **57:**1644-1649.

18. **Grogan, D. W.** 1989. Phenotypic characterization of the archaebacterial genus *Sulfolobus* : comparison of five wild-type strains. Journal of Bacteriology **171:**6710 - 6719.

19. **Grogan, D. W., G. T. Carver, and J. W. Drake.** 2001. Genetic fidelity under harsh conditions: Analysis of spontaneous mutation in the thermoacidophilic archaeon

Sulfolobus acidocaldarius. Proceedings of the National Academy of Sciences **98:**7928-7933.

20. **Grogan, D. W., and R. P. Gunsalus.** 1993. Sulfolobus acidocaldarius synthesizes UMP via a standard de novo pathway: results of biochemical-genetic study. Journal of Bacteriology **175:**1500-1507.

21. **Haseltine, C., R. Montalvo-Rodriguez, A. Carl, E. Bini, and P. Blum.** 1999. Extragenic Pleiotropic Mutations That Repress Glycosyl Hydrolase Expression in the Hyperthermophilic Archaeon *Sulfolobus solfataricus*. Genetics **152:**1353-1361.

22. **Hatzinikolaou, D. G., E. Kalogeris, P. Christakopoulos, D. Kekos, and B. J. Macris.** 2001. Comparative growth studies of the extreme thermophile *Sulfolobus acidocaldarius* in submerged and solidified substrate cultures. World Journal of Microbiology and Biotechnology **17:**229-234.

23. **Joshua, C. J., R. Dahl, P. I. Benke, and J. D. Keasling.** 2011. Absence of Diauxie during Simultaneous Utilization of Glucose and Xylose by Sulfolobus acidocaldarius. J. Bacteriol. **193:**1293-1301.

24. **Kengen, S. W. M., E. J. Luesink, A. J. M. Stams, and A. J. B. Zehnder.** 1993. Purification and characterization of an extremely thermostable β-glucosidase from the hyperthermophilic archaeon Pyrococcus furiosus. European Journal of Biochemistry **213:**305-312.

25. **Kondo, S., A. Yamagishi, and T. Oshima.** 1991. Positive selection for uracil auxotrophs of the sulfur-dependent thermophilic archaebacterium Sulfolobus acidocaldarius by use of 5-fluoroorotic acid. J. Bacteriol. **173:**7698-7700.

26. **Kumar, R., S. Singh, and O. Singh.** 2008. Bioconversion of lignocellulosic biomass: biochemical and molecular perspectives. Journal of Industrial Microbiology and Biotechnology **35:**377-391.

27. **Kurosawa, N., and D. W. Grogan.** 2005. Homologous recombination of exogenous DNA with the *Sulfolobus acidocaldarius* genome: Properties and uses. FEMS Microbiology Letters **253:**141-149.

28. **Magasanik, B.** 1961. Catabolite repression. Cold Spring Harbor Symposia on Quantitative Biology **26:**249-56.

29. **Mamie Z, L., and E. Stephen J.** 2007. SLIC sub-cloning using T4 DNA polymerase treated inserts with RecA.

30. **Martusewitsch, E., C. W. Sensen, and C. Schleper.** 2000. High Spontaneous Mutation Rate in the Hyperthermophilic Archaeon Sulfolobus solfataricus Is Mediated by Transposable Elements. J. Bacteriol. **182:**2574-2581.

31. **Monod, J.** 1942. Recherches sur la croissance des cultures bactâeriennes. Hermann, Paris,.

32. **Mosier, N., C. Wyman, B. Dale, R. Elander, Y. Y. Lee, M. Holtzapple, and M. Ladisch.** 2005. Features of promising technologies for pretreatment of lignocellulosic biomass. Bioresource Technology **96:**673-686.

33. **Park, A.-R., H.-J. Kim, J.-K. Lee, and D.-K. Oh.** 2010. Hydrolysis and Transglycosylation Activity of a Thermostable Recombinant β-Glycosidase from <i>Sulfolobus acidocaldarius</i&gt. Applied Biochemistry and Biotechnology **160:**2236-2247.

34. **Prangishvili, D. A., R. P. Vashakidze, M. G. Chelidze, and I. Y. Gabriadze.** 1985. A restriction endonuclease SuaI from the thermoacidophilic archaebacterium Sulfolobus acidocaldarius. FEBS Letters **192:**57-60.

35. **Rathod, P. K., N. P. Leffers, and R. D. Young.** 1992. Molecular targets of 5-fluoroorotate in the human malaria parasite, Plasmodium falciparum. Antimicrobial Agents and Chemotherapy **36:**704-711.

36. **Schleper, C., K. Kubo, and W. Zillig.** 1992. The particle SSV1 from the extremely thermophilic archaeon Sulfolobus is a virus: demonstration of infectivity and of transfection with viral DNA. Proceedings of the National Academy of Sciences of the United States of America **89:**7645-7649.

37. **She, Q., R. K. Singh, F. Confalonieri, Y. Zivanovic, G. Allard, M. J. Awayez, C. C. Chan-Weiher, I. G. Clausen, B. A. Curtis, and A. De Moors.** 2001. The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2. Proceedings of the National Academy of Sciences of the United States of America **98:**7835 - 7840.

38. **Worthington, P., V. Hoang, F. Perez-Pomares, and P. Blum.** 2003. Targeted Disruption of the {alpha}-Amylase Gene in the Hyperthermophilic Archaeon Sulfolobus solfataricus. J. Bacteriol. **185:**482-488.

39. **Wyman, C. E., B. E. Dale, R. T. Elander, M. Holtzapple, M. R. Ladisch, and Y. Y. Lee.** 2005. Comparative sugar recovery data from laboratory scale application of leading pretreatment technologies to corn stover. Bioresource Technology **96:**2026-2032.

TABLES

Table 1. Characterization of spontaneous uracil auxotrophs

5-FOA Exposure	Uracil auxotrophs (cfu/ml/OD)		
	YT	**YTF**	**YTFU**
Before	2.5×10^8	Nil	Nil
After	3.4×10^2	Nil	1.1×10^7

FIGURES

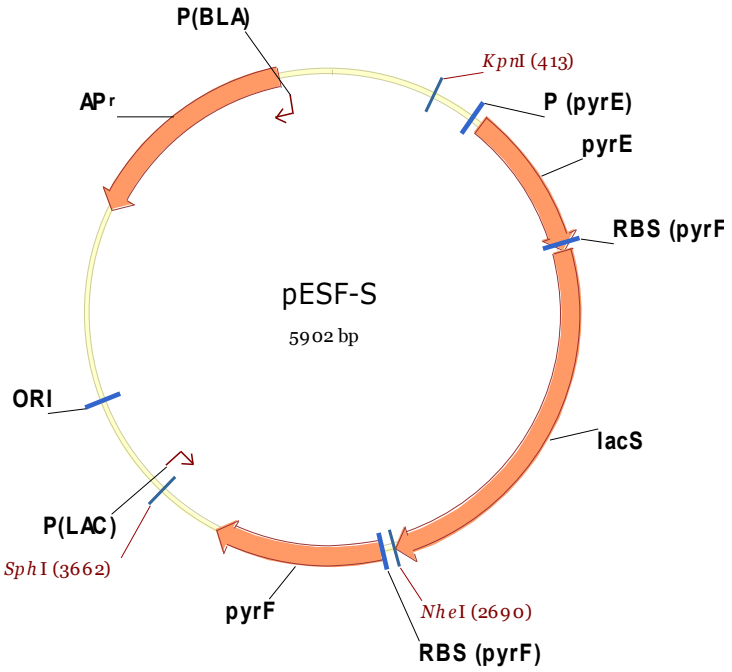


Figure 1. Map of pESF-S. Suicide vector consists of an artificial operon constructed by inserting *S. solfataricus lacS* between *pyrE* and *pyrF*. The artificial operon was inserted into the KpnI and SphI site of pUC19. The vector was designed to replace *pyrEF* with the new operon in *S. acidocaldarius E4-39* and translation of *lacS* was placed under the control of the RBS of *pyrF*.

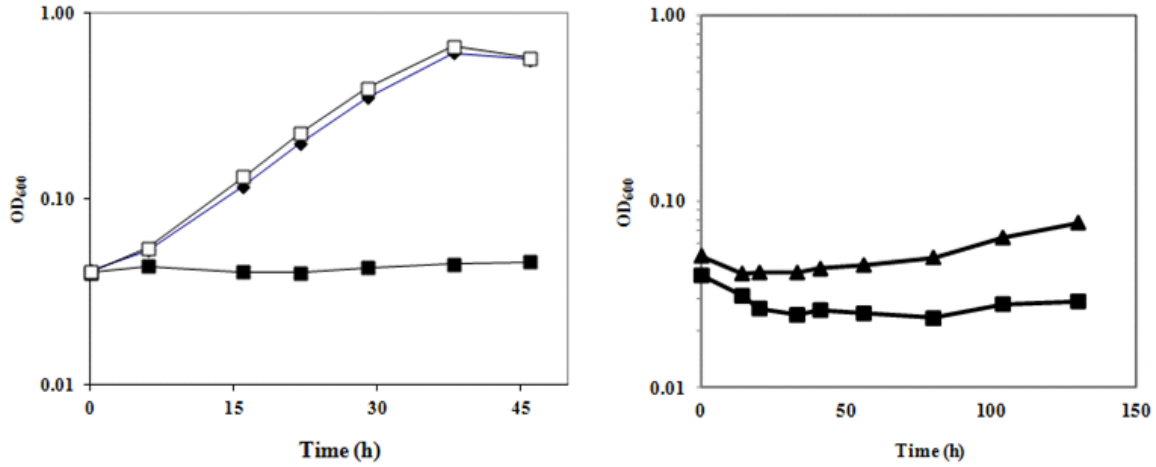


Figure 2. Growth of WT and ESF strains of *S. acidocaldarius* on cellobiose. (A) WT type strain was grown on xylose (solid diamond) and cellobiose (solid square) as sole carbon sources or on both sugars as co-carbon sources (open square). (B) *S. acidocaldarius ESF* harboring *S. solfataricus* in its genome was grown on cellobiose and lactose (solid triangle). Each strain was exposed to each in a complex medium before growth on the growth experiment in a minimal medium.

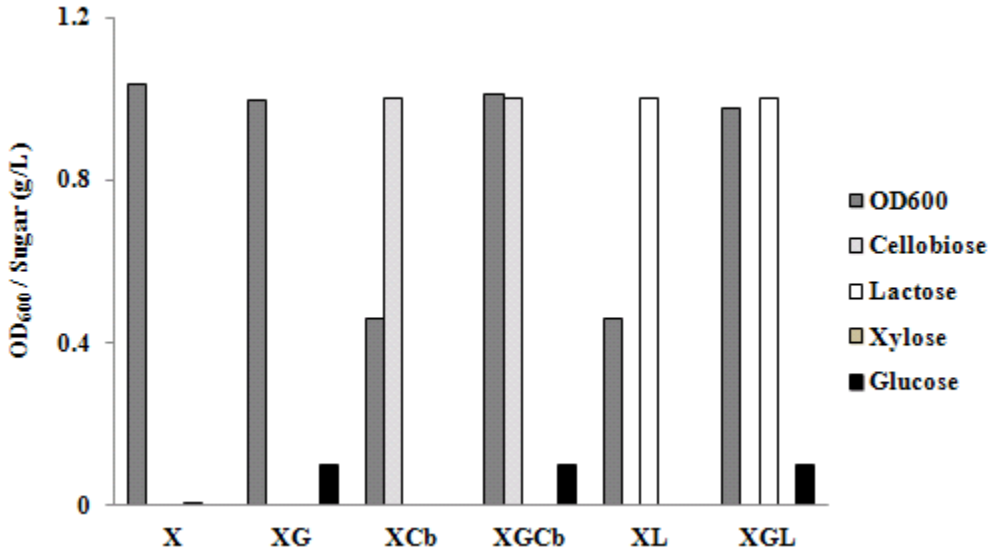


Figure 3. Measurement of residual sugars. *S. acidocaldarius* (WT) was grown on different combination of xylose (X), glucose (G), cellobiose (Cb) and lactose (L) in minimal media. Growth (OD_{600}) and residual sugars was measured after 48 h

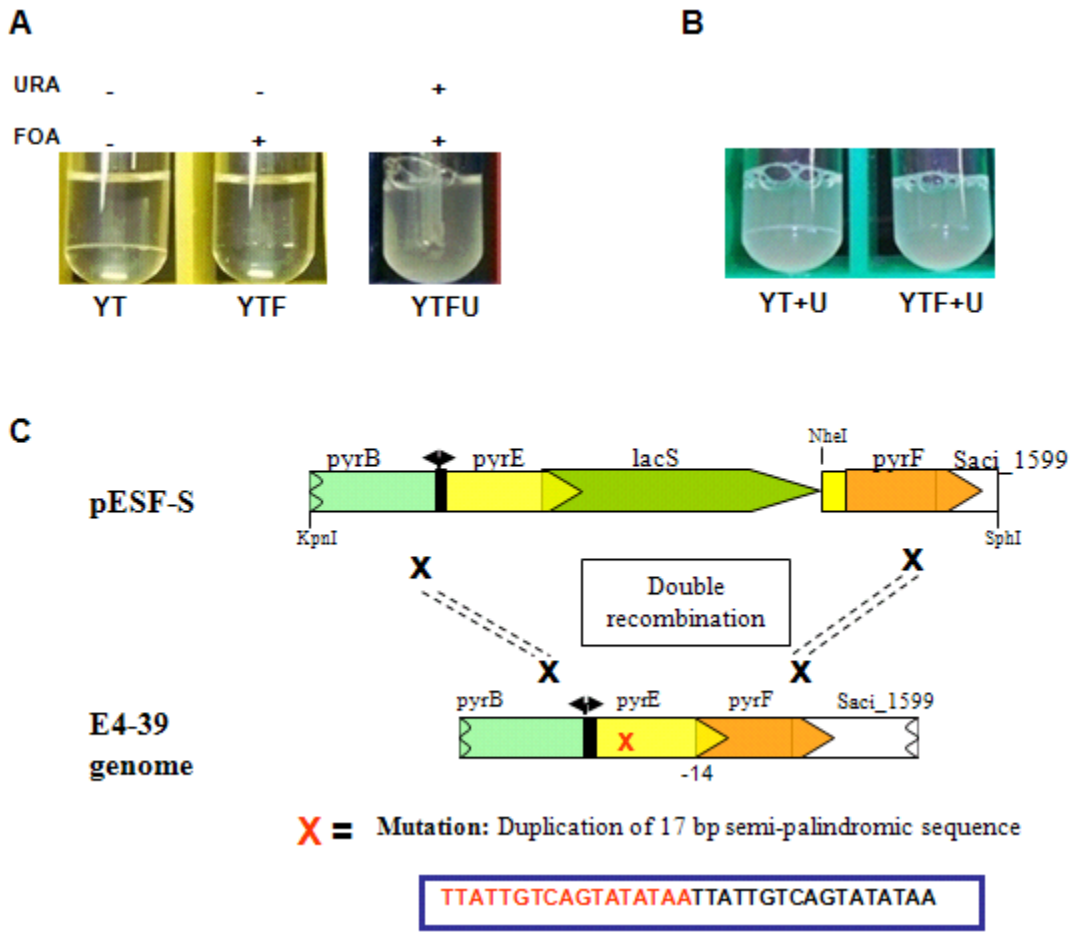


Figure 4. Selection of spontaneous uracil auxotrophs and insertion of *lacS* into *S. acidocaldarius E4-39*. (A) Cells exposed to 50 µg/ml 5-fluoorotic acid (5-FOA) and 5 µg/ml uracil in complex medium (YT) for 6 days were grown in YT, YTF (YT +50 µg/ml 5-FOA) and YTFU (YTF + 20 µg/ml uracil) for 48 h. (B) After 48 h, 20 µg/ml uracil was added to tubes YT and YTF and incubated for additional 48 h to verify the presence of uracil auxotrophs. (C) Double recombination strategy for replacing mutated *pyrE* in strain E4-39 with the *pyrE::lacS::pyrF* in suicide vector pESF-S. Strain E4-39 harbors a 17-bp duplication (X) within its *pyrE*.

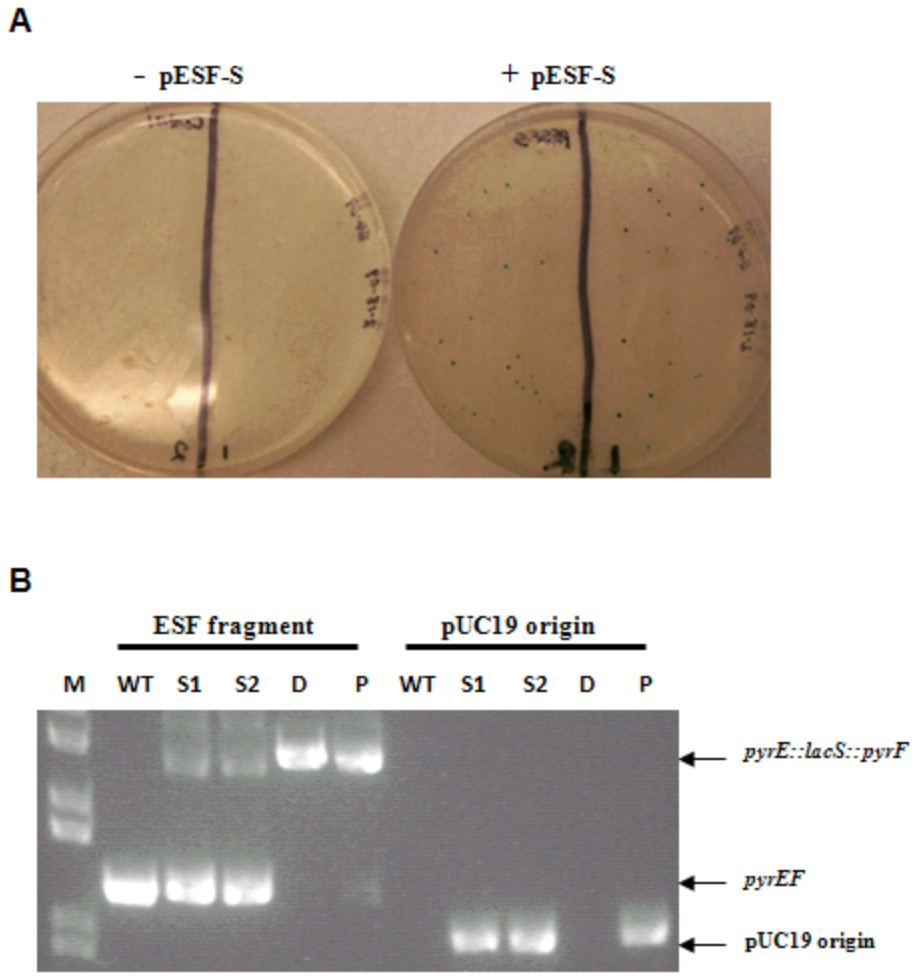


Figure 5. *S. acidocaldarius* transformants. (A) Plate showing blue colonies expressing β-glycosidase encoded by *lacS* from *S. solfataricus*. Each transformant was generated using pESF-S harboring *lacS*. (B) Amplification of *pyrEF, pyrE:;lacS;;pyrF* and pUC19 origin region from WT and *S. acidocaldarius E4-39* transformed with pESF-S.

SUPPLEMENTARY MATERIALS

Table 1. Primers

P/N	Name	Length (bp)	Amplicon	Sequence
45	pESF-SFE	50	*pyrE**	ACGACGGCCAGTGAATTCGAGCTCGGTACC CAAAGAAAGCTATGGATATT
46	pESF-SRE	50	*pyrE**	AAACCTAAAGCTATTTGGAAATGAGTACAT TCTAGCTTTTTCCAATATTT
47	pESF-SFS	50	*lacS*[a]	ATATGGTGAAAAATATTGGAAAAAGCTAGA ATGTACTCATTTCCAAATAG
48	pESF-SRS	50	*lacS*[a]	CTTTTTTTCATTCTCAGCGACGAG**GCTAGC**T TAGTGCCTTAATGGCTTTA
49	pESF-SFF	50	*pyrF**	CCAGTAAAGCCATTAAGGCACTAA**GCTAGC** CTCGTCGCTGAGAATGAAAA
50	pESF-SRF	50	*pyrF**	AAGTTTATCGAAAAGTAATGTTAAGCGTTTG AATGACGTTAGATAAATTC

* *pyrE* and *pyrF* were amplified from *S. acidocaldarius*

[a]*lacS* was amplified from *S. solfataricus*

Chapter Five: Discussion, Conclusion and Future Directions

DISCUSSION

The goal of this study was develop *Sulfolobus acidocaldarius* into a platform for cellulosic biofuel production and metabolic engineering in general. The development of cellulosic biofuel as viable renewable energy alternative to fossil fuels is challenged by the high cost of production. Most of the cost incurred from cellulosic biofuel production comes from the use of multiple and expensive pre-treatment steps to release fermentable sugars and the inefficient conversion of these sugars into fuels by fermenting organism (13-14). Consolidation of the steps involved in the cellulosic biofuel production and improvement of the microbial strains could reduce the cost of producing these biofuels (4). Currently available microbial strains are very limited in their ability to utilize most cellulosic sugars; some strains catabolize the hexose sugars, usually glucose but not pentoses (6). Similarly, most model microbes are not capable of utilizing most sugars in the presence of glucose due to carbon catabolite repression imposed by glucose (12).

This study explored the opportunity to develop *S. acidocaldarius*, a hyperthermo-acidophilic crenarchaea with optimal growth at 75-80°C and pH 2-3 (5) into a consolidated microbial system for the cellulosic sugars into biofuels. The ability of the organism to utilize most abundant cellulosic sugars was examined in Chapter 2. Our result revealed that the organism was capable of simultaneous metabolism of sugars such as xylose, arabinose and galactose in the presence of glucose with catabolite repression. Glucose-induced catabolite repression was previously suggested to be present in the organism (3, 8, 11). A novel mechanism of simultaneous metabolism was also uncovered in which the organism consumed each sugars in a mixture at a rate that is roughly proportional to its starting concentrations in the growth medium (Chapter 2). Broad sugar metabolism and absence of diauxie during growth on multiple sugars, including glucose enhances the profile of *S. acidocaldarius* as a potential candidate for cellulosic biofuel production. However, genetic tools for engineering these capabilities in the organism are not readily available.

The development of *S. acidocaldarius* as a platform for metabolic engineering is challenged by the absence of reproducible genetic tools such as plasmid and host strains with selectable traits. A number of shuttle vectors have been used for heterologous gene expression in *Sulfolobus* species, but the most successful and reproducible plasmids so far, are those based on pRN1 (1-2, 10). This cryptic plasmid (pRN1) from *S. islandicus REN1H1* replicates via rolling circle mechanism, but its origin of replication. Chapter 3 described effort to identify the origin of replication of pRN1 and the development of pRN1-based shuttle vectors for heterologous gene expression in *S. acidocaldarius*. A 241-bp region downstream of the replication protein (Orf904) of pRN1 was identified and evaluated for its role as the putative origin of replication of the plasmid. Our result revealed that this putative origin was very critical for *in vivo* replication of pRN1. The putative origin and *orf56/orf904* were identified as the minimum replication unit of the pRN1; this result agrees with previous report that Orf56 and Orf904 are critical for pRN1 replication (1, 10). A 100-bp stem-loop structure within the putative origin was proposed to be the critical *cis*-acting double-strand origin involved in the formation of the cruciform structure that initiates plasmid replication (Chapter 3). Shuttle vectors develop in this study were also used as vehicle to evaluate heterologous gene expression in *S. acidocaldarius*; and a β-glycosidase from *S. solfataricus* that was encoded by *lacS* (Sso3019) was successfully expressed in the organism (Chapter 3). In addition to the shuttle vectors, we also isolated and characterized spontaneous uracil auxotrophic strains of *S. acidocaldarius* to serve as recipient strains.

Development of these genetic tools and heterologous expression of LacS enhances the metabolic engineering potential of *S. acidocaldarius*.

S. acidocaldarius does not metabolize cellobiose, a key intermediates in cellulose metabolism that is metabolized by other *Sulfolobus* species (7, 9). Chapter 4 focused on understanding the reason for the absence of cellobiose metabolism in the *S. acidocaldarius*. The gene encoding LacS from *S. solfataricus* was inserted the genome of *S. acidocaldarius E4-39* via homologous recombination. We inserted *lacS* into *pyrEF* operon between the two genes and placed its translation under the control of *pyrF* RBS. The resulting strains were evaluated for β-glycosidase activity and cellobiose metabolism. Strains harboring *lacS* exhibited β-glycosidase activity as indicated by their ability to cleave X-gal; but these strains did not grow on cellobiose. We also inserted *bgaS*, a homologue of *lacS* in *S. acidocaldarius* into the *pyrEF* operon, but the resulting strains neither exhibited glycosidase activity nor grow on cellobiose (Chapter 4). This study provided evidence that absence of cellobiose metabolism in *S. acidocaldarius* could be due to its inability to import the sugar. However, the study provided with the opportunity to further develop more genetic tools for studying this organism.

CONCLUSION

The research described in this dissertation demonstrates the potential of *S. acidocaldarius* as a platform for metabolic engineering. The organism was able to simultaneously utilize multiple cellulose sugars as carbon and energy sources without catabolite repression that is commonly observed in bacteria and eukaryotes. Absence of catabolite repression in the organism enhances its cellulosic biofuel application by eliminating the need to de-regulating catabolite repression, an effort that has been shown to be very challenging in *E. coli* and yeast. This study uncovered a novel and well coordinated mode multiple sugar consumption in the organism. The dissertation also described studies that were aimed at developing genetic tools for studying the organism. A number of potent tools such as recipient strains with selectable traits; plasmid and strategies for genome manipulations were described were described. Some of these tools are currently being used to heterologous engineer new capabilities in the organism. The study evaluated the effect of thermophilic β-glycosidase from *S. solfataricus* on the ability of *S. acidocaldarius* to utilize cellobiose. The result revealed no correlation between absence of a functional β-glycosidase and cellobiose metabolism in the organism; but suggested that the organism lacks the transporter for this sugar. Overall, this study indicates that *S. acidocaldarius* has the potential to be a good platform for metabolic engineering; but lots of work has to be done in the area of developing readily available and reproducible genetic tools.

FUTURE DIRECTIONS

To achieve to potential of *S. acidocaldarius* as a platform for metabolic engineering, especially production of biofuels from cellulose biomass, further works need to be carried out. The dissertation reported the ability of the organism to coordinated metabolism multiple sugars, details of the mechanism or its regulation is not fully understood. Preliminary metabolite analysis suggested that rate of sugar consumption during simultaneous metabolism is determined by the cytosolic concentration of individual sugars rather than their concentrations in the growth

medium. To fully understand the mechanism and regulation simultaneous sugar catabolism in *S. acidocaldarius*, functional transcriptomics and proteomics studies need to be carried. It would important to identify the factor responsible for correlating rate of sugar consumption to its concentration in the growth medium. Further metabolite analysis and sugar uptake assay need to be carried out to verify the impact of intra- and extra-cellular sugar concentration on simultaneous catabolism in the organism. Efforts should also be directed towards improving the available genetic tools to ensure that functions in the organism are analyzed *in vivo* rather than *in vitro*. Further analysis need be performed to elucidate the entire components of pRN1 origin of replication and detailed mechanism of replication. More effective and efficient selection strategy needs to be developed for this organism to improve its appeal for metabolic engineering. So far, only few inducible promoter systems are available for this organism; more inducible promoters need to be developed to enable over-expression of thermophilic proteins which is usually challenging in *E. coli* or yeast. It will be important to confirm that this organism does not import cellobiose by carrying out sugar uptake assay. Currently effort is being made to express cellobiose ABC-transporters from *S. solfataricus* in the organism because the sugar is very critical for cellulosic biofuel engineering.

REFERENCES

1. **Berkner, S., D. Grogan, S.-V. Albers, and G. Lipps.** 2007. Small multicopy, non-integrative shuttle vectors based on the plasmid pRN1 for *Sulfolobus acidocaldarius* and *Sulfolobus solfataricus*, model organisms of the (cren-)archaea. Nucleic Acids Research. **35:**e88-.
2. **Berkner, S., A. Wlodkowski, S.-V. Albers, and G. Lipps.** 2010. Inducible and constitutive promoters for genetic systems in <i>Sulfolobus acidocaldarius</i&gt. Extremophiles **14:**249-259.
3. **Bini, E., and P. Blum.** 2001. Archaeal catabolite repression: a gene regulatory paradigm. Advances in Applied Microbiology **50:**339-66.
4. **Bokinsky, G., P. P. Peralta-Yahya, A. George, B. M. Holmes, E. J. Steen, J. Dietrich, T. Soon Lee, D. Tullman-Ercek, C. A. Voigt, B. A. Simmons, and J. D. Keasling.** 2011. Synthesis of three advanced biofuels from ionic liquid-pretreated switchgrass using engineered Escherichia coli. Proceedings of the National Academy of Sciences.
5. **Brock, T. D., K. M. Brock, R. T. Belly, and R. L. Weiss.** 1972. Sulfolobus: A new genus of sulfur-oxidizing bacteria living at low pH and high temperature. Archives of Microbiology **84:**54-68.
6. **Dien, B. S., M. A. Cotta, and T. W. Jeffries.** 2003. Bacteria engineered for fuel ethanol production: current status. Applied Microbiology and Biotechnology **63:**258-266.
7. **Grogan, D. W.** 1989. Phenotypic characterization of the archaebacterial genus *Sulfolobus* : comparison of five wild-type strains. Journal of Bacteriology **171:**6710 - 6719.
8. **Haseltine, C., M. Rolfsmeier, and P. Blum.** 1996. The glucose effect and regulation of alpha-amylase synthesis in the hyperthermophilic archaeon *Sulfolobus solfataricus*. Journal of Bacteriology **178:**945-50.
9. **Hatzinikolaou, D. G., E. Kalogeris, P. Christakopoulos, D. Kekos, and B. J. Macris.** 2001. Comparative growth studies of the extreme thermophile *Sulfolobus acidocaldarius*

in submerged and solidified substrate cultures. World Journal of Microbiology and Biotechnology **17:**229-234.

10. **Lipps, G.** 2009. Molecular biology of the pRN1 plasmid from Sulfolobus islandicus. Bichemical Society Transactions **37:**42 - 45.
11. **Lubelska, J., M. Jonuscheit, C. Schleper, S.-V. Albers, and A. Driessen.** 2006. Regulation of expression of the arabinose and glucose transporter genes in the thermophilic archaeon *Sulfolobus solfataricus*. Extremophiles **10:**383-391.
12. **Magasanik, B.** 1961. Catabolite repression. Cold Spring Harbor Symposia on Quantitative Biology **26:**249-56.
13. **Mosier, N., C. Wyman, B. Dale, R. Elander, Y. Y. Lee, M. Holtzapple, and M. Ladisch.** 2005. Features of promising technologies for pretreatment of lignocellulosic biomass. Bioresource Technology **96:**673-686.
14. **Wyman, C. E., B. E. Dale, R. T. Elander, M. Holtzapple, M. R. Ladisch, and Y. Y. Lee.** 2005. Comparative sugar recovery data from laboratory scale application of leading pretreatment technologies to corn stover. Bioresource Technology 96:2026-2032.