ABSTRACT

LIAN, HONG. Transcriptomic and Physiological Analysis of a Recombinant *Pyrococcus furiosus* Strain Metabolically Engineered to Produce 3-Hydroxypropionate from CO_2 and Maltose. (Under the direction of Dr. Robert M. Kelly).

As a new paradigm for renewable fuels, Electrofuels expand the boundaries of traditional biofuels that rely on photosynthetic carbon fixation. Electrofuels could overcome the drawback inherent to biofuels approaches via direct utilization of CO_2 to bypass energy intermediate sugar, more efficient non-photosynthetic CO_2 fixation pathways and durable low potential energy sources. Here, we aim to metabolically engineer a hyperthermophilic archeaon, *P. furiosus* strain COM1, as a host for production of desired liquid fuels or valuable chemicals (e.g., 3-hydroxypropionate) by incorporation of the 3HP/4HB cycle from the thermoacidophile *M. sedula*. Sub-pathway 1 (SP1) of the 3-hydroxypropionate/4-hydroxybutyrate (3HP/4HB) cycle is composed of the heteromultimeric acetyl-CoA/propionyl-CoA carboxylase (E1) (encoded by Msed_0147, Msed_0148, and Msed_1375), malonyl-CoA reductase (E2) (encoded by Msed_0709), and malonic semialdehyde reductase (E3) (encoded by Msed_1993). These three enzymes sequentially convert HCO_3^- and acetyl-CoA into 3-hydroxypropionate (3-HP). To achieve our goal, the *in vitro* study of SP1 was first performed to confirm all enzyme functions individually, and that if assembled could convert acetyl-CoA to 3-HP. The genes encoding these enzymes were produced by heterologous expression in *E. coli* or *P. furiosus*. In this process, the rate-limiting carboxylase in carbon fixation and its accessory enzyme biotin protein ligase (BPL) captured our specific attention, as did carbonic anhydrase (CA), the enzyme catalyzing conversion of CO_2 to bicarbonate as substrate for carboxylase. Therefore, Msed_0390, and Msed_2010 were identified in the *M. sedula* genome to encode a functional CA and BPL.

The *in vitro* work provided the basis for incorporating CA and BPL by metabolic engineering into recombinant *P. furiosus* to facilitate 3-HP/4-HB cycle function for production of fuels and chemicals. All three enzymes (E1, E2 and E3) were produced in recombinant form, with E2 and E3 in *E. coli,* and the three subunits encoding E1 in *P. furiosus*. Recombinant forms of *M. sedula* BPL and CA were also produced in *E. coli*. Formation of 3-HP was demonstrated *in vitro* by the assembly of recombinant E1, E2 and E3, using either acetyl-CoA or malonyl-CoA as substrate. The five *M. sedula* enzymes mentioned above were then metabolically engineered into *P. furiosus* strain COM1 to enable production of 3-hydroxypropionate (3-HP) from maltose and CO_2. *P. furiosus* was grown in a 2-liter fermentor at 95°C until late exponential phase, at which time the temperature was reduced to 72°C to initiate 3-HP production. The addition of genes encoding *M. sedula* CA and biotin BPL, led to two-fold higher 3-HP titers than in their absence. Furthermore, CO_2 mass transfer was found to be rate-limiting, as 10-fold higher 3-HP concentrations (from 0.3 to 3 mM) were obtained when sparging and agitation rates were increased from 15 mL/min to 50 mL/min, and 250 rpm to 400 rpm, respectively. Transcriptomic analysis using a hybrid oligonucleotide microarray revealed that the *M. sedula* genes were among the most highly transcribed in the recombinant *P. furiosus* genome, both at 2.5 h and 40 h after the temperature shift. Monovalent cation /H+ antiporter subunits were found up-regulated in MW76, presumably involved in the export of 3-HP to maintain cytosolic homeostasis to prevent intracellular acidification. From 2.5 h to 40 h, the *P. furiosus* COM1 transcriptome showed significant down-regulation of genes involved in energy metabolism, amino acid biosynthesis, DNA metabolism and protein synthesis, while MW76 maintained relatively high cellular activity. Up-regulation of biotin synthesis genes in MW76 was also noted at

40h, presumably to provide sufficient biotin for biotinylation of the *M. sedula* carboxylase subunit. The results of this study will be critical in designing further genetic refinements to the MW76 strain and to coordinate these changes to bioprocessing strategies.

Transcriptomic and Physiological Analysis of a Recombinant *Pyrococcus furiosus* Strain Metabolically Engineered to Produce 3-Hydroxypropionate from CO_2 and Maltose

by
Hong Lian

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

Chemical Engineering

Raleigh, North Carolina

2014

APPROVED BY:

Robert M. Kelly Committee Chair	Balaji M. Rao
David F. Ollis	Amy M. Grunden

BIOGRAPHY

Hong Lian was born in Zhangzhou, a small city in the South of China. In 2004, she was enrolled into East China University of Science and Technology (ECUST) and started her journey of being a bio-engineer. ECUST is where people in China produced antibiotics via fermentation for the first time. In the college, she had a chance to contact respected professors and outstanding fellows. In Shanghai, she experienced the viability of this city for its internationality. She wanted to learn more advanced technology and knowledge in the United States. After graduation and earning her Bachelors of Science in Engineering in Biochemical Engineering, she headed to North Carolina State University where she was happy to become a member of the Electrofuels team, working on saving the world from the energy crisis. In the lab, she has become friends with bugs, such as *E. coli*, *M. sedula*, and *P. furiosus*---though these bugs had given her some tough times, finally they show their characteristics as loyal and excellent friends.

ACKNOWLEDGMENTS

I would like to thank Dr. Kelly for allowing me the opportunity to study under his guidance. I would also like to thank fellow members of the Kelly group who have offered me so much. I would like to thank US DOE ARPA-E projects to fund the research we are doing in our lab. Especially thank our ARPA-E team, Dr. Yejun Han, Dr. Aaron Hawkins, and Andrew Loder. Also, I want to thank Dr. Arpan Mukherjee for the help of microarray experiments.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 4

Chapter 1

Biological Conversion of Carbon Dioxide and Hydrogen into Liquid Fuels and Industrial Chemicals

The rise of Electrofuels

What is an Electrofuel and why it is advantageous over traditional Biofuels?

As a new paradigm for renewable energy, Electrofuels expand the boundaries of traditional biofuels that rely on photosynthetic carbon fixation. To provide energy-dense, infrastructure-compatible liquid fuels at dramatically higher conversion efficiencies, the Electrofuels program was launched by the US Department of Energy (DOE) Advanced Research Projects Agency-Energy (ARPA-E) in 2009. Various projects were instituted to develop metabolically-engineered microbial strains capable of assimilating energy, fixing inorganic carbon, and producing fuel molecules, all of which are independent of photosynthesis (1). Such foundational efforts in microbiology, genomics, metabolic engineering, and synthetic biology are anticipated to revolutionize the biofuels sector.

Biofuel production relies primarily on photosynthesis, in which solar radiation is assimilated through Photosystems I and II to drive the Calvin-Benson-Bassham (CBB) cycle for CO_2 fixation (1). Via CBB cycle, inorganic carbon is converted to glucose, and further directed to biosynthesis of various cell materials, including polysaccharide-rich structures (lignocellulosic biomass). The energy stored in biomass can be converted to liquid fuels fermentatively, however, with relatively low efficiency. One energetically unfavorable part lies in the conversion of biomass into fermentable sugars. There is a high cost of pretreatment to liberate cellulose from the lignin seal and its crystalline structure (2, 3), typically involving enzymes for cellulose hydrolysis. Another limiting factor is the low efficiency of photosynthesis. In reality, the photosynthetic efficiency of terrestrial plants is around 1%, and up to 3-4% in the best cases, leading to typical biomass yields of below 10 $g/m^2/day$ (4).

Under limited land resources, biofuels feedstocks and food crops have significant competition in open markets. Though aquatic species, such as microalgae, have become attractive because of their significantly higher average photosynthetic efficiency than typical land crops (biomass yield averages around 20 $g/m^2/day$, with peaks of 60 $g/m^2/day$), acquiring high quantities of algal biomass from artificial culture for downstream processing is problematic. The design of enclosed photobioreactors to prevent culture collapse or takeover by low oil strains is challenging and expensive (4,5). To summarize, the current cellulosic biofuel mode relies on photosynthesis in green plants or algae, both of which constrain the further improvement of energy efficiency in biofuel production.

Electrofuels could overcome the drawback inherent to biofuels approaches via direct utilization of CO_2 to bypass energy intermediate sugar, more efficient non-photosynthetic CO_2 fixation pathways and durable low potential energy sources. At the core of the program are the chemolithoautotrophic organisms, capable of fixing and reducing inorganic carbon, and thus dependent on non-photosynthetic pathways for deriving energy from a variety of inorganic substrates (1).

Electrofuels: State of the art

(i) Carbon fixation pathways:

To date, in addition to the CBB cycle, five more carbon fixation pathways have been discovered in nature to assimilate CO2 into cell material (6,7): Reductive citric acid cycle (RCC), Reductive acetyl-CoA pathway (RAP), 3-hydroxypropionate Bicycle (3HP), 3-hydroxypropionate/4-hydroxybutyrate cycle (3HP/4HB), and Dicarboxylate/4-hydroxybutyrate cycle (DC/4HB). These have been identified in various organisms, in some cases involving enzymes that are oxygen-tolerant or oxygen-sensitive. The direct net products of these pathways are pyruvate (3HP bicycle) and acetyl-CoA (all other four pathways), both of which are energy-rich intermediates that directly feed liquid fuel synthesis (See Figure 1) (8). Electrofuels research is evaluating the characteristics of each of these pathways, as well as synthetic pathways not found in the natural world (1).

RAP is the only non-cyclic CO_2 fixation pathway that directly utilizes reduced C1 compounds (CO, formic acid, formaldehyde, methanol, etc.) as free metabolites or as moieties attached to specific C1 carrier compounds, while in other pathways, CO_2 is assimilated into the carbon backbone of metabolites via carboxylation reactions (6,9). The RAP is also unique due to its additional physiological roles for energy conservation and as an electron sink (9). Except for the CBB cycle that may sometimes serve solely for recycling of electron carriers (10,11), other pathways support autotrophic growth (9). The RCC, 3HP bicycle, 3HP/4HB and DC/4HB cycles show significant structural similarity and all involve two conserved metabolites: acetyl-CoA and succinyl-CoA. The RCC and the 3-HP bicycle operate solely in bacterial lineages, while the DC/4HB and the 3HP/4HB cycle are used only by archaea. Despite the considerably overlapping parts within the 3HP bicycle, 3HP/4HB cycle and DC/4HB cycle, they seem to have evolved independently (9).

RAP is generally considered as the most energy efficient (12). Its extremely low energy requirement is consistent with the energetic constraints of methanogens and acetogens that live close to the thermodynamic limit (13). However, this pathway has a high demand for metals (Mo or W, Co, Ni, and Fe), which are more water-soluble in the reduced oxidation state (i.e., under anoxic conditions). It also requires bioenergetically expensive coenzymes (tetrahydropterin and cobalamin) and operates under strict anaerobiosis. That is why the distribution of RAP in nature is restricted (6). RCC is also an energy efficient route (12), for example, used by the anaerobic photosynthetic green sulphur bacteria that live at a depth of over 100 m below the surface of anoxic water-bodies, where extremely low light energy is available (14). The energy costs of the 3HP bicycle are higher, and it has rather limited distribution in the green non-sulfur phototrophs Chloroflexaceae family (15). Both 3HP/4HB and DC/HB cycles are not energetically efficient (i.e., higher ATP requirement than the other pathways) (12). However, their thermotolerance is an important feature for (hyper)thermophilic crenarchaea. 3HP/4HB can be used strategically in mesophiles as well, due to its postulated existence in mesophilic "marine group I" archaea representing ~20% of all picoplankton cells in the world oceans (16). Additionally, archaea do not need large amount of sugars and fatty acids for cell wall and cell membrane construction (6,17). Therefore, the product from 3HP/4HB and DC/HB, acetyl-CoA and pyruvate, can be utilized primarily as principal biosynthetic precursors (6).

(ii) Energy assimilation and host development:

Chemolithoautotrophic organisms derive energy from the proton motive force generated by electron flow from a reductant (e.g., H_2, H_2S, Fe^{2+}) to an oxidant (e.g. O_2, S, CO_2). This energy flow can provide the driving force for synthesis of ATP and reducing equivalents (NAD(P)H) required for carbon fixation (1). Based on this, Electrofuels platform organisms could include microorganisms, such as *Ralstonia eutropha* (H_2) (18), *Acidithiobacillus thiooxidans* (H_2S) (19), *Nitrosomonas europaea* (NH_3) (20), and *Acidithiobacillus ferrooxidans* (Fe^{2+}) (21). Heterotrophs, such as *Escherichia coli* and *Pyrococcus furiosus*, can also be genetically engineered into chemolithoautotrophs for growth and fuel production on formate or H_2 (1).

With hydrogen gas as source of reducing power for CO_2 fixation, either a natural H_2-utilizing autotroph or a suitable heterotroph can be engineered to produce the desired Electrofuel (22). *Ralstonia eutropha* represents the ‘autotrophic’ approach, and its genetic tools include plasmid-based expression systems (23) and chromosomal modifications via homologous recombination (24,25,26). Some species in the genus *Clostridium* are also natural H_2-utilizing autotrophs, whose genetic tools await further development. Beyond the optimization of endogenous CO_2-fixation machinery, further concern is that the carbon flow should be directed to Electrofuel production rather than to native cellular metabolism (22). The second avenue for Electrofuel development using heterotrophs will be addressed in this study, focusing on *P. furiosus*. This can be accomplished by using a highly competent strain (COM1) that enables chromosomal modification (27,28). Overexpression of its native cytoplasmic hydrogenase I (SHI) has been achieved (29), thereby potentially facilitating H_2 uptake for regeneration of reducing power.

(iii) Liquid fuel synthesis:

Electrofuels leverage recent advances in metabolic engineering and synthetic biology to direct carbon and energy fluxes from pyruvate and acetyl-CoA to final fuel molecules dependent on a variety of pathways (1). (see **Figure 1**).

3HP/4HB cycle and *Metallosphaera sedula*

The archaeal order *Sulfolobales* contains a number of thermophilic microorganisms whose genomes encode enzymes of the 3HP/4HB pathway (30). In particular, Fuchs and co-workers established that this pathway occurs in *Metallosphaera sedula* (Topt 73°C) (31). Also, the presence of high acetyl-CoA/propionyl-CoA carboxylase activities indicates the operation of this cycle in *Sulfolobus metallicus* (Topt 75°C), *Acidianus brierleyi* (Topt 70°C) and *Acidianus ambivalens* (Topt 80°C) (32,33,34). *M. sedula* can grow heterotrophically on peptides, autotrophically on CO2 (sole carbon source) and H2, and mixotrophically, including on casamino acids and FeSO4 or metal sulfides (35). The 3HP/4HB cycle in *M. sedula*, is summarized in (**Table 1** and **Figure 2**). The cycle starts with one acetyl-CoA and assimilates two CO2 molecules (in the form of bicarbonate), forming 3-HP and 4-HB as key intermediates, and finally regenerates two acetyl-CoA molecules. A key feature of 3HP/4HB pathway is that all of the high-energy reductant is supplied as NADPH. In this cycle, Enzymes 1 (E1) and 2 (E2) are both bi-functional, catalyzing two different reactions. So far, all the enzymes in this cycle have been confirmed in *M. sedula* through biochemical characterization (22).

The 3HP/4HB cycle is linked to central carbon metabolism by siphoning off succinyl-CoA, which is oxidatively converted to oxaloacetate, pyruvate, and PEP (36). By contrast, DC/HB (Dicarboxylate/4-hydroxybutyrate cycle) uses oxygen-sensitive pyruvate synthase to convert acetyl-CoA to a C3 compound for gluconeogenesis. To further support the aerobic operation of 3HP/4HB cycle, succinyl-CoA reductase uses NADPH as the cofactor rather than ferredoxin, and the latter is always used by anaerobic enzymes; 4-hydroxybutyryl-CoA dehydratase is exceptionally oxygen tolerant (37). This cycle utilizes bicarbonate as substrate (not CO_2), which can be advantageous for organisms living in slightly alkaline marine water where bicarbonate is the main inorganic carbon species (6).

Though gene disruption by cross-species complementation has been reported in *M. sedula* (38), the potential for metabolic engineering in this organism requires continued development of its genetic system (22). Therefore, an alternative host should be selected for implementing the 3HP/4HB pathway for biological conversion of carbon dioxide into a variety of chemical and fuel molecules.

***Pyrococcus furiosus* as electrofuel host**

To select an optimal host for the 3-HP/4-HB CO_2 fixation pathway, one would consider a thermophilic or hyperthermophilic organism that can metabolize hydrogen and also has a tractable genetic system. The hyperthermophilic archaeon *P. furiosus* fulfills these criteria. A genetic system has been developed in a naturally competent variant of *P. furiosus* by a deletion of the *pyrF* gene, required for uracil biosynthesis. This *pyrF* deletion strain (COM1) is a naturally competent uracial auxotroph that has been successfully used to generate markerless deletion mutants (27). *P. furiosus* grows between 70 to 103^{o}C (T_{opt} 100^{o}C) (39), so enzymes from 3HP/4HB cycle can be expressed, folded and function at their optimal temperature (~70 to 75^{o}C). As mentioned previously, a key feature of the 3HP/HB cycle is using NADPH as reductant. *P. furiosus* can produce H2 as an end product of sugar fermentation (40). It contains three [NiFe]-hydrogenases. One is membrane-bound (Pfu MBH) that accepts electrons from reduced ferredoxin (Fdred) and catalyses the H2 evolution (40,41), in an energy-conserving manner (42). The other two are cytoplasmic enzymes, termed soluble hydrogenase I (Pfu SHI) and hydrogenase II (Pfu SHII), utilizing NADP(H) as physiological electron carrier (40,43). Pfu SHI and Pfu SHII are thought to function similarly in H2 evolution, but SHII is less active (44). Kinetic analysis indicates that Pfu SHI has an order of magnitude higher H_2-uptake activity than H_2-evolution activity at 80°C. As such, it contributes to the generation of NADPH for biosynthesis, by recycling the 'waste' product of respiration, H2, produced by Pfu MBH (45). The recently reported ability to over-express SHI in engineered *P. furiosus* (29) enhances this organism's capability to metabolize H_2 to regenerate NADPH for continued driving of 3HP/4HB cycle. Therefore, by integration into the chromosome of *P. furiosus* under the control of constitutive native promoters, production of acetyl-CoA from the 3HP/4HB cycle, and its conversion to desired liquid fuels from other fuel synthesis pathways, can be achieved and then optimized.

Carboxylases

Classification and functions of carboxylases. Carboxylases are among the most important enzymes in the biosphere, catalyzing the incorporation of a CO_2 molecule into an organic substrate. They play critical roles in the fixation of inorganic carbon (CO_2 or HCO_3^-), carbon assimilation and anaplerosis, all important biosynthetic processes, as well as contribute to redox metabolism (46,49). Such diverse functions and wide distribution in nature benefit from using a single C1 unit, CO_2, which is far less toxic than formaldehyde, the alternative C1 unit used in biology (47,48). CO_2 is readily available to the cell, either as atmospheric gas or in its hydrated form (HCO_3^-), the substrate for all biotin-dependent carboxylases.

Depending on different functionality, carboxylases can be classified into five groups (46) (see **Table 2**):

1) The <u>autotrophic carboxylases</u> operate in autotrophic pathways to form central precursor molecules (PEP, pyruvate, acetyl-CoA, and TCA cycle intermediates) solely from inorganic

carbon, while central precursor molecules are converted from organic growth substrate via assimilatory carboxylases serving in heterotrophic pathways.

2) Anaplerotic carboxylases contribute to the refilling of intermediates in TCA cycle when they are utilized for biosynthesis.

3) Unlike the above three group carboxylases that ultimately provide central precursors, biosynthetic carboxylases are involved in the synthesis of cellular building blocks starting from central precursor molecules.

4) The last one, redox-balancing carboxylases transfer reducing equivalents [NAD(P)H] onto CO_2, the terminal electron acceptors, to maintain redox balance.

***Comparison of autotrophic carboxylases and their connection to pathway physiology*.** All the natural autotrophic CO_2 fixation pathways have carboxylase as key enzymes, except for the reductive acetyl-CoA pathway, in which two molecules of CO_2 are initially converted into acetyl-CoA by reduction via formate dehydrogenase and CO dehydrogenase/acetyl-CoA synthase (46,49). Carboxylases exert significant pressure on the total energy cost of an autotrophic pathway (46). The most energetically efficient pathway, reductive citric acid cycle (rTCA), contains three reducing carboxylases: isocitrate dehydrogenase (NADPH-dependent), α-ketoglutarate:ferredoxin oxidoreductase (ferredoxin-dependent), and pyruvate: ferredoxin oxidoreductase (ferredoxin-dependent). The CBB cycle is more energy-intensive than the rTCA cycle (6), because its carboxylase, RubisCO, catalyzes the carboxylation of ribulose-1,5-bisphosphate in a non-reductive mechanism. A separate reduction of the initial fixation product should be performed during the next steps, but following the activation of the fixed CO_2 group through phosphorylation with extra consumption of ATP. This extra consumed energy comprises nearly 60% of the total energetic difference between the CBB cycle and rTCA cycle (46). Moreover, RubisCO exhibits a wasteful oxygenase side-reaction in which ribulose-1,5-bisphosphate is oxidized by molecular oxygen, thereby losing the captured carbon (50). The 3-hydroxypropionate (3HP) bicycle is equally energetically expensive as the CBB cycle (6) and the extra energy cost over rTCA cycle can mostly be ascribed to the property of the central carboxylase, acetyl-CoA/propionyl-CoA carboxylase, catalyzing a non-reductive carboxylation. Although a further reduction of fixed CO_2 is required here, as in the CBB cycle, the activation of the fixed CO_2 group for reduction occurs within this carboxylation step, requiring hydrolysis of ATP. In contrast, in the CBB cycle, activation is catalyzed by 3-phosphoglyceratekinase (46). Similar to the 3HP bicycle, the 3-hydroxypropionate/4-hydroxybutyrate (3HP/4HB) cycle uses biotin-dependent acetyl-CoA/propionyl-CoA carboxylase (51,52) to catalyze energy-intensive carboxylation reactions. The dicarboxylate/hydroxybutyrate (DC/HB) cycle employs pyruvate:ferredoxin oxidoreductase and phosphoenolpyruvate (PEP) carboxylase (53), leading to a more energetically favorable reductive carboxylation.

It has been proposed that there is apparent tradeoff between oxygen sensitivity and activity of carboxylases (46). The catalytically efficient reducing carboxylases of the ferredoxin type in rTCA are highly oxygen-sensitive, because they have easily oxidized iron-sulfur clusters (54). *Hydrogenobacter thermophilus* encodes two different isoforms of α-ketoglutarate: ferredoxin oxidoreductase: one is preferentially expressed under anoxic conditions, while the other is synthesized under oxic conditions. The catalytic efficiency of the oxygen-tolerant isoform is lower than that of oxygen-sensitive one by an order of magnitude (55).The oxygen sensitivity of carboxylases limit the rTCA cycle to certain anoxic environments, and, as such, it is mainly found in anaerobic and microaerobic organisms. The oxygen tolerance of RubisCO determines the wide distribution of the CBB cycle in strictly aerobic and facultatively anaerobic autotrophic organisms (6). Due to the side function as oxygenase, there are multiple RubisCO isoforms existing in many facultatively anaerobic proteobacteria, differentially expressed as a function of the ratio of CO_2 to O_2 that it encounters (56). The 3HP bicycle operates in the facultatively aerobic green non-sulfur bacterium *Chloroflexus aurantiacus* (57). Corresponding to the oxygen sensitivity of its ferredoxin-dependent carboxylase, the DC/HB cycle only exists in anaerobic autotrophic species of the Orders Thermoproteales and Desulfurococcales (12,58). In contrast, the 3HP/4HB cycle is distributed in (micro) aerobic, members of the Sulfolobales, and aerobic marine group I archaea (12,52).

In general, the carboxylases that catalyze carboxylation via a reductive mechanism are more efficient, and contribute significantly to the low energy cost of the rTCA cycle and DC/HP cycle, in exchange of the oxygen tolerance of the organisms. The CBB cycle, the 3-HP bicycle, and the 3HP/4HB cycle all use non-reducing carboxylases, and though more energy intensive, they are adapted to (micro)aerobic environment inhabited by their host microorganism (46).

Carboxylases are believed to catalyze the rate-limiting steps of carbon fixation pathways, based on detailed bioenergetics analysis (59). Some carboxylases may show higher specific activity, but less affinity, toward CO_2 or HCO_3^-. In general, carboxylases with slow kinetics need to be expressed at a high level for autotrophic growth, and hence substantially consume the carbon and nitrogen resources of the cell (9). However, pyruvate synthase and 2-ketoglutarate synthase, which are known to have very low specific activity, can still facilitate the fast growth of *Hydrogenobacter thermophilus* strain TK-6, when it is oxidizing H_2 under aerobic conditions (55).

Application of carboxylases in synthetic biology. Synthetic CO_2 fixation pathways can be applied to production of Electrofuels, biopharmaceuticals, or fine chemicals from CO_2 and an appropriate energy source. In addition to the traditional focus on optimization of existing natural pathways and their corresponding carboxylases determining carbon flux (46, 60), synthetic biology can be used to create novel "customized" CO_2 fixation pathways via combination of different enzymatic reactions from different sources or organisms. Bar-Even *et al*. computationally assembled the synthetic carbon fixation pathways, some of which showed significant quantitative advantages over their natural counterparts on kinetics, energetic and topology (59). This can overcome the historical and evolutionary burdens intrinsic to the six natural autotrophic pathways.

Carboxylases can be the limiting (bottleneck) enzyme in synthetic pathways (46). For example, pyruvate carboxylase of *Corynebacterium glutamicum* redirects carbon flux to the synthesis of fermentation products that are derived from the TCA cycle. Overexpression of this enzyme resulted in a 50% increase in lysine production (61). Likewise, acetyl-CoA carboxylase overproduction resulted in a 6-fold increase in the rate of fatty acid synthesis in *E. coli* (62). The succinate production by *E. coli* was also increased upon overproduction of phosphoenolphosphate (PEP) carboxylase (63). Likewise, succinic acid production from glucose in *E. coli* was increased 3.5-fold by overexpression of PEP-carboxylase (62). To drive the operation of synthetic CO_2 fixation pathways to the highest rates, systematic search tool was used by Bar-Even *et al.* (59) to investigate the pathways employing known favorable carboxylating enzymes, characterized by high specific activities and affinities toward CO_2 or HCO_3^-. The most promising synthetic pathways revealed from their analysis featured the most efficient carboxylating enzyme, PEP carboxylase. Alternatively, novel robust carboxylases can be designed, since the capacity of known enzymes to catalyze the reactions of non-native substrates can be evaluated by computational screening (64).

Biotin protein ligase, an accessory enzyme to acetyl-CoA carboxylase

As described previously, the 3HP/4HB cycle and 3HP bicycle have central bifunctional carboxylase, namely acetyl-CoA/propionyl-CoA carboxylase. Acetyl-CoA carboxylases (ACC) are multiple subunit complexes, composed of biotin carboxylase (BC), biotin carboxyl carrier protein (BCCP) and carboxyl transferase (CT). A functional carboxylase should have an activated BCCP, with a biotin molecule covalently bound in a specific lysine residue. This biotin is carboxylated by BC through accepting a carboxyl group from bicarbonate and forming carboxybiotin. BCCP then swings carboxybiotin to the CT component where the carboxyl group is transferred to acetyl-CoA (propionyl-CoA) and produces malonyl-CoA (methylmalonyl-CoA) (65). Biotinylation of BCCP is catalyzed by biotin protein ligase (BPL), which promotes an amide linkage between the carboxyl group of biotin molecule and the ε-amino group of a specific lysine residue of BCCP (66). This process is Mg^{2+}-and ATP-dependent, and occurs in two steps. The first step generates the reactive intermediate biotinyl-5′-AMP. In the second step, the ε-amino group of the target lysine of BCCP reacts with the activated carboxyl group to form biotinylated-BCCP with release of AMP (67) (see **Figure 3**). *Corynebacterium glutamicum* is a biotin auxotrophic Gram-positive bacterium engineered for large-scale production of amino acids (68). The biotin-dependent pyruvate carboxylase is key enzyme for anaplerotic oxaloacetate synthesis in *C. glutamicum* under biotin-sufficient conditions, applied to L-lysine production (61). L-lysine production by *C. glutamicum* is positively affected by biotin (69). Overproduction of the pyruvate carboxylase improved supply of the L-lysine precursor oxaloacetate during anaplerosis (61). On the other hand, the activity of pyruvate carboxylase can be further enhanced. Overexpression of the BPL gene (*birA*) from *C. glutamicum* increased the level of the pyruvate carboxylase holoenzyme, resulting in a significant growth advantage in glucose minimal medium, as well as a two-fold higher L-lysine yield on glucose compared with the control strain (68).

Carbon dioxide capture

CO_2 concentrating mechanisms and carbonic anhydrase in RubisCO-based carbon fixation. Since biological carbon fixation depends on the concentration of CO_2 in the environment, mechanisms to concentrate CO_2 have evolved to compensate for the imbalance between the high demand of inorganic carbon and low ambient CO_2 concentration (22,70). It is well known that some chemoautotrophs, various photoautotrophic microorganisms, green algae, and plants have developed distinct mechanisms to increase the concentration of CO_2 in close proximity to ribulose-1,5-bisphosphate carboxylase/oxygenase (RubisCO), the CO_2-fixing enzyme of the CBB pathway (70,71,72). As described previously, RubisCO has a low affinity for CO_2 and does not discriminate well between CO_2 and the competing substrate, O_2. Elevation of CO_2 concentration near the active site of RubisCO can reduce its O_2-

dependent oxygenase activity, so the wasteful photorespiration and the loss of carbon can be inhibited (73).

Carbonic anhydrases (CAs) are essential components of CO_2 concentrating mechanisms (CMMs). Their variety of functions accord with the type of CMMs in which they are involved, and can be summarized mainly as follows (70,71,72,74):

(a) Conversion of bicarbonate to CO_2, for fixation by RubisCO. This is always the final step required in all known CMMs.

(b) Conversion of CO_2 to bicarbonate, for pre-fixation by phosphoenolpyruvate (PEP) carboxylase. This mostly occurs during four carbon (C4) mechanism and crassulacean acid metabolism (CAM) in terrestrial higher plants. The pre-fixation product will be decarboxylated and release CO_2 again for RubisCO.

(c) Co-localization with RubisCO in specialized microcompartments (i.e. 'carboxysomes' in cyanobacteria or 'pyrenoid' in eukaryotic green algae such Chlamydomonas spp., where carbon fixation occurs.)

(d) Facilitation of diffusion of CO_2 via forcing rapid equilibration between CO_2 and bicarbonate.

In a review (74), Zabaleta *et al*. summarized several already well described CCMs, including the C4 mechanism, CAM, bicarbonate accumulation systems and capsular structures (carboxysomes) around RubisCO. Furthermore, they proposed a basal CCM (bCCM) universally present in all plants, functioning to recycle the large amounts of CO_2 liberated from TCA cycle and photorespiration in mitochondria for carbon fixation in chloroplasts. The bCCM consists of at least two types of mitochondrial CA: a γ-type CA domain of the mitochondrial NADH dehydrogenase complex and a matrix localized β-type CA that converts CO_2 to HCO_3^- and/or transfers HCO_3^- from mitochondria into the cytosol. HCO_3^- is then transported from the cytosol into the chloroplast by a putative bicarbonate translocase or by a pathway resembling partial C4 metabolism. In chloroplast, β-type CA dehydrates bicarbonate into CO_2 as substrate for RubisCO.

CMMs in cyanobacteria are summarized by Badger and Price (72). At least four different inorganic carbon uptake systems (including two bicarbonate transporters and two CO_2 uptake systems) are present to accumulate HCO_3^- in the cytosol of the cell. HCO_3^- subsequently diffuses into the interior of the microcompartment ('carboxysome'), where it is dehydrated by a specialized CA co-localized with RubisCO. The carboxysome is believed to be a CO_2 diffusion barrier (75), so RubisCO can be CO_2-saturated. The availability of multiple cyanobacterial genomes enables the comparative genomics of the CCM in cyanobacteria and reveals different carboxysome types, HCO_3^- and CO_2 transport systems and carbonic anhydrase types (72).

Novel CO_2-concentrating mechanisms and carbonic anhydrase for Electrofuels. Carbon concentrating mechanisms have only been elucidated for organisms living by the CBB cycle (RubisCO-based). Utilizing the CBB cycle, *Ralstonia eutropha* has been genetically engineered as an important 'Electrofuel-producer'. Kusian *et al*. (76) found that the deletion of a β-type CA in this facultative chemoautotroph leads to its death at the CO_2 concentrations found in ambient air. Furthermore, this *R. eutropha* mutant recovers its ability to grow under air via complementation of heterologous CA genes from *E. coli* and human, indicating sufficient CA activity is required and the function of some CAs can be replaced by foreign CAs beloning to different families.

CAs are ubiquitous enzymes in nature and their roles have been noted widely in non-CBB carbon fixation pathways. For example, a β-type CA was discovered in *Methanobacterium thermoautotrophicum*, a thermophilic methanoarchaeon synthesizing cell carbon from CO_2 through incomplete reductive tricarboxylic acid cycle (TCA) (77,78). This CA was presumed to concentrate HCO_3^- in the vicinity of pyruvate carboxylase and phosphoenolpyruvate (PEP) carboxylase for the synthesis of oxaloacetate (79). Carbon fixation pathways are generally energy-consuming processes. However, acetogens can reduce CO_2 into acetate using the reductive acetyl-CoA or Wood/Ljungdahl pathway during which ATP synthesis results (80). CA activity has been detected in acetogens, *Acetobacterium woodii* and *Clostridium thermoaceticum*, which may facilitate their energy conservation (81). Just as different carbon fixation pathways feature different first committed-step CO_2 fixing enzymes, the detection of novel CA in organisms living by non-RubisCO-based carbon fixation probably represents a starting point for exploring novel CMMs.

The CO_2-fixing enzymes of 3HP, 3HP/4HB and DC/4HB cycles use bicarbonate rather than CO_2 as substrate (9), hence a functional CA is possibly a key requirement for CO_2 and bicarbonate interconversion, and thus for an efficient electrofuel production pathway. It would be useful to find out more about the molecular details of how CO_2 is fed to the carbon-fixing machinery via specific CA and/or CCMs, since the growth environment, physiological conditions and CO_2-utilizing enzyme reservoir of these organisms are quite different from those of photo- and chemoautotrophs living by RubisCO-based carbon fixation.

As mentioned previously, the heterotrophic archaeon *P. furiosus*, the host for an engineered 3HP/4HB pathway, appears to contain only a γ-type CA (82). The extent to which this CA can facilitate CO_2-utilization by the 3HP/4HB cycle in *P. furiosus* is unclear at present. Genome sequencing indicates the existence of genes encoding CAs in *M. sedula* (35), and the potential interaction between CA and the 3HP/4HB cycle raises a question: is it necessary to engineer CA in *P. furiosus*? In any case, the means for concentrating and/or hydrating CO_2 is obviously an area that merits attention for Electrofuels development (22).

3-hydroxypropionic acid: A valuable intermediate in the 3HP/4HB cycle

3-hydroxypropionic acid, a valuable industrial chemical. 3-Hydroxypropionic acid (abbreviated as 3-HP; MF $C_3H_6O_3$; CAS Registry Number 503-66-2; pK4.51; MW 90.07884; density 1.08) is a three-carbon, non-chiral organic molecule, containing one carboxyl group and one β-hydroxyl group. US DOE ranks it at the top third among twelve value added chemicals derived from biomass (83,84). 3-HP can serve as a build- ing block for multiple commercial commodity and specialty chemicals, such as 1,3-propanediol ($C_3H_8O_2$), acrylic acid ($C_3H_4O_2$), methyl acrylate ($C_4H_6O_2$), acrylamide (C_3H_5NO), ethyl 3-HP ($C_5H_{10}O_3$), malonic acid ($C_3H_4O_4$), propiolactone ($C_3H_4O_2$), and acrylonitrile (C_3H_3N) (83). It can also be used for polymer synthesis. In particular, poly-(3-hydroxypropionic acid) [poly (3-HP)] is used in manufacture of surgical and drug-releasing materials, because of its good biocompatibility and biodegradability. [Poly (3-HP)] exhibits an unusual glass transition temperature and heat stability (85). Combining the properties of poly-(3-hydroxybutyrate) [poly (3HB)] and poly-(2-hydroxypropionic acid) [poly (2HP)], the resulting [Poly (3-HP)] is more resistant to hydrolytic cleavage than [Poly (2-HP)], but more enzymatically cleavable than [poly-(3-HB)] without carrying methyl groups at the polymer backbone. These advantages allow [poly-(3-HP)] to substitute traditional petroleum-based polymers. As estimated, 3.63 million tons of 3-HP is sold in global markets annually (83) (see **Figure 4**).

Chemical method for 3-HP production. 3-HP can be synthesized from several chemical routes, including hydration from acrylic acid in the presence of liquid acid or solid acid catalyst (86), or the oxidation of 1,3-propanediol (1,3-PD) via a palladium-containing supported catalyst (87). However, both acrylic acid and 1,3-propanediol are expensive and make the process economically unfavorable (88). Some other synthesis options such as hydrolysis of 3-hydroxypropionitrile or β-propiolactone are also not commercially feasible, since the starting materials have toxicity and carcinogenicity (89).

Biological routes for 3-HP production. The commercial market for 3-HP and the disadvantage of chemical synthesis methods for 3-HP production has motivated studies of biological production since the early 2000s. 3-HP can be made naturally as end product through a range of metabolic pathways in some microorganisms (88). *Lactobacillus* species produce 3-HP via metabolizing glycerol (90,91); *Alcaligenes faecalis*, *Byssochlamys* sp., *Geotrichum* sp. and *Trichoderma* sp. have metabolic pathways that convert acrylic acid to 3-HP (92,93,94). A uracil catabolism pathway was discovered in *Saccharomyces kluyveri* that can degrade uracil to form 3-HP (95). On the other hand, 3-HP is an intermediate of two autotrophic CO_2 fixation pathways: 3-hydroxypropionate bicycle and 3-hydroxypropionate/4-hydroxybutyrate cycle. The former exists in *Chloroflexus aurantiacus*, a thermophilic photosynthetic green non-sulfur bacterium (57), while the latter is functional in members of the Sulfolobales, such as *Metallosphaera sedula* (31,96).

Generally, a high product concentration (> 100g/L), productivity (> 2g/L/h) and substrate-to-product yield (> 50%) are important for the production of a commercially viable chemical. None of the aforementioned natural microorganisms can fulfill those criteria for 3-HP production, lacking cheap and abundant substrates and well-developed genetics systems and fermentation processes (88). Therefore, the development of genetically engineered microorganisms and optimization of the production process are essential. As summarized by Kumar *et al.* (88), various biochemical pathways have been suggested for 3-HP production from glucose or glycerol as substrate (see **Figure 5-6**). Seven pathways utilizing glucose were proposed and patented by US-based agricultural company Cargill, all of which have been investigated experimentally; only Pathway II and III exhibit promising prospects of application (84,97,98, 99,100,101,102).

A recombinant *E. coli* strain was constructed (103) according to the Pathway II scheme, utilizing malonyl-CoA reductase from *C. aurantiacus* and nicotinamide nucleotide transhydrogenase and acetyl-CoA carboxylase from *E. coli*. Since one intermediate of this pathway, malonyl-CoA, is also an intermediate for fatty acid synthesis, Lynch *et al.* reduced the level of enoyl-ACP reductase expression in *E. coli*, and thus the level of fatty acid synthesis. Furthermore, they re-directed the carbon flux towards acetyl-CoA by deleting the competing pathways, leading to lactate, ethanol, acetate and methylglyoxal. This yielded the best result reported to date - a high titer of 49 g/L at 69 h with a 0.46 mol/mol yield on glucose (104).

β-alanine is an important precursor for 3-HP in pathways III, IV, VI, and VII (see **Figure 5**). A recombinant *E. coli* with constructed pathway III (105) can produce 9.2 g/L 3-HP. Pathways VI and VII have been suggested to address the lack of ATP generation in pathways III and IV. This is because the energy consuming steps can be bypassed through the use of alanine 2,3-aminomutase, which transfers an amino group from α- to β-carbon. Its isoenzyme, lysine 2,3-aminomutase, has been engineered by Liao *et al.* (106) to convert α-alanine to β-alanine. However, little detailed information has been provided for 3-HP microbial production through pathway VI and VII (88).

Glycerol has become a better substrate than glucose for 3-HP production (88), since the rapid growth of biodiesel production generates large quantity of cheap by-product glycerol ($120–$180/MT glycerol compared with $525/MT glucose - www.icispricing.com; www.indexmundi.com). With glycerol as a substrate, there are two different pathways leading to 3-HP formation: CoA-dependent and CoA-independent (s.ee Figure 6). It is known that both pathways naturally exist in *K. pneumoniae* (88,107). The CoA-dependent pathway requires more details on catalytic mechanisms underlying the critical propionaldehyde dehydrogenase (PduP) (88). The enzymes and the two-step CoA-independent pathway have been studied extensively. Both CoA-dependent and CoA-independent pathways share the first-step conversion of glycerol to 3-hydroxypropanealdehyde (3-HPA), catalyzed by glycerol dehydratase (GDHt). For a CoA-independent pathway, 3-HPA is then oxidized directly to 3-HP by aldehyde dehydrogenase (ALDH). This pathway can be assembled by the GDHt (DhaB) from *K. pneumonia*, as well as ALDHs from different microorganisms. As concluded by *Kumar et al*. (88), several ALDHs have been considered for this purpose

(108,109,110,111,112,113), especially three ALDHs isolated from *E. coli*, *K. pneumoniae* and *Azospirillium brasilense* that have high activities for 3-HPA. Both NAD^+ and $NADP^+$ an be used by these ALDHs, but a higher activity was observed with NAD^+ for most aldehydes.

3-HP production via a CoA-independent pathway has been demonstrated in *E. coli*, accompanied by several challenges, such as the intracellular accumulation of toxic 3-HPA caused by imbalance between the activities of the GDHt (DhaB) and ALDH (AldH). To address the problem, HPA should be rapidly utilized by efficient ALDHs. In recombinant *E. coli* strain SH-BGK1, the native AldH was replaced with α-ketoglutaric semialdehyde dehydrogenase (KGSADH) of *A. brasilense*, considered as the most suitable ALDH for 3-HP production (113). Another strategy adopted by Raj *et al.* was to control the enzyme ratio in *E. coli* SH-BGA1, in which *aldH* is cloned in a high-copy plasmid and *dhaB* in a low-copy plasmid (114).

Moreover, *E. coli* fails to synthesize coenzyme B12 , a necessary coenzyme of GDHt (DhaB) (115,116), so the external addition is required in the culture medium (88). Coenzyme B12 can be damaged by a reaction between the activated Co-C bond of the coenzyme and oxygen, so GDHt loses activity completely after one hour of oxygen exposure (117). In view of this, Mori *et al.* introduced glycerol dehydratase reactivase (GdrAB) into recombinant *E. coli*, which improves the stability of DhaB by replacing a damaged coenzyme B12 molecule with a new intact molecule (118). Likewise, Kwak *et al*. (119) developed another recombinant *E. coli* by cloning DhaB and DhaR (reactivation factor of glycerol dehydratase) from *Lactobacillus brevis*. Despite the fact that 3-HP can be produced at higher titer and productivity by the aforementioned *E. coli* strains, the expensive supplementation of coenzyme B12 negatively impacts the economic feasibility, and a coenzyme B12 natural producer which can work as host strain, such as *K. pneumonia*, might be used to remedy this problem.

K. pneumonia has been studied most extensively for 3-HP production from glycerol. Gene manipulation methods can be adapted from that of *E. coli* due to the evolutionary and biochemical proximity of the two species (120). Moreover, *K. pneumoniae* shows a high glycerol assimilation rate, and the glycerol metabolic pathways have been well characterized (88).

In *K. pneumoniae*, glycerol is assimilated through either 'oxidative' or 'reductive' pathways. The production of 3-HP occurs in the reductive pathway, in which 3-HPA can branch off to either 3-HP or 1,3-propanediol (PDO), depending on the activity of the relevant enzymes (oxidoreductase and 3-HPA-specific ALDH) and the availability of redox cofactors (NAD+ and NADH) (88) (See Figure 6). One strategy for 3-HP accumulation is to prevent carbon flow toward PDO by deletion of oxidoreductase and overexpression of ALDH, while an alternative strategy of co-production of PDO and 3-HP has been adopted by many researchers (108,121,122) .

As mentioned previously, the conversion of 3-HPA to 3-HP by ALDH reduces NAD^+ to NADH. NAD^+ could be regenerated by the oxidative metabolism of glycerol and/or an electron transport chain in the presence of oxygen. However, oxygen significantly suppresses the biosynthesis of coenzyme B12, and the expression of *dhaB*, which thereby down-regulates the whole reductive glycerol metabolism (117). This leads to a dilemma as to whether the aerobic requirement should be fulfilled for continuous NAD^+ regeneration, or anaerobic conditions should be maintained for effective coenzyme B synthesis and adequate carbon flux toward 3-HP. Maintaining proper aeration conditions becomes critically important for the successful production of 3-HP with *K. pneumonia* (88). Seeking an alternative solution, PDO can be co-produced with 3-HP simultaneously, since NADH or NAD^+ can be recycled within the reductive pathway (See Figure 6). In this case, 3-HP production can be performed under anaerobic or microaerobic conditions. The highest 3-HP titer achieved so far in *K. pneumoniae* is 48.9 g/L, along with 25.3 g/L PDO and 24.9 g/L lactic acid in 28 h (122). This titer is almost as high as the best case from a recombinant *E. coli* fed by glucose (104). Furthermore, the 3-HP produced in this *K. pneumoniae* strain has higher yield (mol/mol) and productivity (g/L/h) (see **Table 3**).

Summary

As a new paradigm for renewable fuels, Electrofuels expand the boundaries of traditional biofuels that rely on photosynthetic carbon fixation. Electrofuels could overcome the drawback inherent to biofuels approaches via direct utilization of CO_2 to bypass energy intermediate sugar, more efficient non-photosynthetic CO_2 fixation pathways and durable low potential energy sources. At the core of the program are the chemolithoautotrophic organisms, capable of fixing and reducing inorganic carbon, and thus dependent on non-photosynthetic pathways for deriving energy from a variety of inorganic substrates (1). Here, we aim to metabolically engineer a hyperthermophilic archeaon *P. furiosus* as a host for production of desired liquid fuels or valuable chemicals 3-hydroxypropionate by incorporation of 3HP/4HB cycle from thermoacidophile *M. sedula*. Additionally, the rate-limiting carboxylase in carbon fixation and its accessory enzyme biotin protein ligase draw our specific attention, as well as carbonic anhydrase, and the potential carbon concentration mechanisms involved in non-photosynthetic assimilation of inorganic carbon from thermoacidophilic environments.

REFERENCES

1. **Conrado RJ, Haynes CA, Haendler BE, Toone EJ.** 2013. Electrofuels: A new paradigm for renewable fuels, p. 1037-64. *In* Lee JW(ed), Advanced Biofuels and Bioproducts. Springer, New York, N.Y.

2. **Sticklen MB.** 2008. Plant genetic engineering for biofuel production: towards affordable cellulosic ethanol. Nat. Rev. Genet. **9**:433-443.

3. **Mosier N, Wyman C, Dale B, Elander R, Lee YY, Holtzapple M, Ladisch M.** 2005. Features of promising technologies for pretreatment of lignocellulosic biomass. Bioresour. Technol. **96**:673-686.

4. **Brown A.** 2010. Algae - The future for bioenergy? Summary and conclusions from the IEA Bioenergy ExCo64 Workshop.

5. **Vasudevan PT, Briggs M.** 2008. Biodiesel production--current state of the art and challenges. J. Ind. Microbiol. Biotechnol. **35**:421-430.

6. **Berg IA.** 2011. Ecological aspects of the distribution of different autotrophic CO_2 fixation pathways. Appl. Environ. Microbiol. **77:**1925-36.

7. **Thauer RK.** 2007. Microbiology. A fifth pathway of carbon fixation. Science **318:**1732-1733.

8. **Connor MR, Liao JC.** 2009. Microbial production of advanced transportation fuels in non-natural hosts. Curr. Opin. Biotechnol. **20:**307-15.

9. **Bar-Even A, Noor E, Milo R.** 2012. A survey of carbon fixation pathways through a quantitative lens. J. Exp. Bot. **63:**2325-42.

10. **Wang X, Falcone DL, Tabita FR.** 1993. Reductive pentose phosphate-independent CO_2 fixation in *Rhodobacter sphaeroides* and evidence that ribulose bisphosphate carboxylase/oxygenase activity serves to maintain the redox balance of the cell. J. Bacteriol. **175:**3372-3379.

11. **Tichi MA, Tabita FR.** 2000. Maintenance and control of redox poise in *Rhodobacter capsulatus* strains deficient in the Calvin-Benson-Bassham pathway. Arch. Microbiol. **174:**322-333.

12. **Berg IA, Kockelkorn D, Ramos-Vera WH, Say RF, Zarzycki J, Hügler M, Alber**

BE, Fuchs G. 2010. Autotrophic carbon fixation in archaea. Nat. Rev. Microbiol. **8:**447-60.

13. **Deppenmeier U, Müller V.** 2008. Life close to the thermodynamic limit: how methanogenic archaea conserve energy. Results Probl. Cell Differ. **45:**123-152.

14. **Manske AK, Glaeser J, Kuypers MMM, Overmann J.** 2005. Physiology and phylogeny of green sulfur bacteria forming a monospecific phototrophic assemblage at a depth of 100 meters in the Black Sea. Appl. Environ. Microbiol. **71:**8049-8060.

15. **Strauss G, Fuchs G.** 1993. Enzymes of a novel autotrophic CO_2 fixation pathway in the phototrophic bacterium *Chloroflexus aurantiacus*, the 3-hydroxypropionate cycle. Eur. J. Biochem. **215:**633-43.

16. **Karner MB, DeLong EF, Karl DM.** 2001. Archaeal dominance in the mesopelagic zone of the Pacific Ocean. Nature **409:**507-510.

17. **Gonzalez O, Gronau S, Falb M, Pfeiffer F, Mendoza E, Zimmer R, Oesterhelt D.** 2008. Reconstruction, modeling & analysis of *Halobacterium salinarum* R-1 metabolism. Mol. Biosyst. **4:**148-159.

18. **Brigham CJ, Gai CS, Lu J, Speth DR, Worden RM, Sinskey AJ.** Engineering *Ralstonia eutropha* for production of isobutanol from CO_2, H_2, and O_2, p. 1065-90. *In* Lee, J. W. (ed), Advanced Biofuels and Bioproducts. Springer, New York, NY.

19. **Valdes J, Ossandon F, Quatrini R, Dopson M, Holmes DS.** 2011. Draft genome sequence of the extremely acidophilic biomining bacterium *Acidithiobacillus thiooxidans* ATCC 19377 provides insights into the evolution of the Acidithiobacillus genus. J. Bacteriol. **193:**7003-4.

20. **Hommes NG, Sayavedra-soto LA, Daniel J, Arp DJ.** 2003. Chemolithoorganotrophic growth of *Nitrosomonas europaea* on fructose. J. Bacteriol. **185:** 6809-14.

21. **Valdés J, Pedroso I, Quatrini R, Dodson RJ, Tettelin H, Blake R, Eisen J a, Holmes DS.** 2008. *Acidithiobacillus ferrooxidans* metabolism: from genome sequence to industrial applications. BMC Genomics **9:**597.

22. **Hawkins AS, McTernan PM, Lian H, Kelly RM, Adams MW.** 2013. Biological conversion of carbon dioxide and hydrogen into liquid fuels and industrial chemicals. Curr. Opin. Biotechnol. **24:**1-9.

23. **Sichwart S, Hetzler S, Bröker D, Steinbüchel A.** 2011. Extension of the substrate

utilization range of *Ralstonia eutropha* strain H16 by metabolic engineering to include mannose and glucose. Appl. Environ. Microbiol. **77:**1325-1334.

24. **Orita I, Iwazawa R, Nakamura S, Fukui T.** 2012. Identification of mutation points in *Cupriavidus necator* NCIMB 11599 and genetic reconstitution of glucose-utilization ability in wild strain H16 for polyhydroxyalkanoate production. J. Biosci. Bioeng. **113:**63-69.

25. **Kawashima Y, Cheng W, Mifune J, Orita I, Nakamura S, Fukui T.** 2012. Characterization and functional analyses of R-Specific enoyl coenzyme A hydratases in polyhydroxyalkanoate-producing *Ralstonia eutropha*. Appl. Environ. Microbiol. **78:**493-502.

26. **Lindenkamp N, Volodina E, Steinbuchel A.** Genetically modified strains of *Ralstonia eutropha* H16 with beta-ketothiolase gene deletions for production of copolyesters with defined 3-Hydroxyvaleric acid contents. Appl. Environ. Microbiol. **78:**5375-5383.

27. **Lipscomb GL, Stirrett K, Schut GJ, Yang F, Jenney FE, Scott RA, Adams MWW, Westpheling J.** 2011. Natural competence in the hyperthermophilic archaeon *Pyrococcus furiosus* facilitates genetic manipulation: Construction of markerless deletions of genes encoding the two cytoplasmic hydrogenases. Appl. Environ. Microbiol. **77:**2232-2238.

28. **Bridger SL, Lancaster WA, Poole FL, Schut GJ, Adams MWW.** 2012. Genome sequencing of a genetically tractable *Pyrococcus furiosus* strain reveals a highly dynamic genome. J. Bacteriol. **194:** 4097-4106.

29. **Chandrayan SK, McTernan PM, Hopkins RC, Sun J, Jenney FE, Adams MWW.** 2012. Engineering hyperthermophilic Archaeon *Pyrococcus furiosus* to overproduce its cytoplasmic [NiFe]-hydrogenase. J. Biol. Chem. **287:**3257-3264.

30. **Auernik KS, Cooper CR, Kelly RM.** 2008. Life in hot acid: pathway analyses in extremely thermoacidophilic archaea. Curr. Opin. Biotechnol. **19:**445-453.

31. **Hügler M, Huber H, Stetter KO, Fuchs G.** 2003. Autotrophic CO_2 fixation pathways in archaea (Crenarchaeota). Arch. Microbiol. **179:**160-73.

32. **Burton NP, Williams TD, Norris PR.** 1999. Carboxylase genes of *Sulfolobus metallicus*. Arch. Microbiol. **172:**349-353.

33. **Ishii M, Miyake T, Satoh T, Sugiyama H, Oshima Y, Kodama T, Igarashi Y.** 1997.

Autotrophic carbon dioxide fixation in *Acidianus brierleyi*. Arch. Microbiol. **166:**368-371.

34. **Menendez C, Bauer Z, Huber H, Gad'on N, Stetter KO, Fuchs G.** 1999. Presence of acetyl coenzyme A (CoA) carboxylase and propionyl-CoA carboxylase in autotrophic Crenarchaeota and indication for operation of a 3-hydroxypropionate cycle in autotrophic carbon fixation. J. Bacteriol. **181:**1088-98.

35. **Auernik KS, Kelly RM.** 2009. Physiological versatility of the extremely thermoacidophilic archaeon *Metallosphaera sedula* supported by heterotrophy, autotrophy and mixotrophy transcriptomes. Appl. Environ. Microbiol. **76:**931-5.

36. **Estelmann S, Hügler M, Eisenreich W, Werner K, Berg IA, Ramos-Vera WH, Say RF, Kockelkorn D, Gad'on N, Fuchs G.** 2011. Labeling and enzyme studies of the central carbon metabolism in Metallosphaera sedula. J. Bacteriol. **193:**1191-200.

37. **Scott R, Naser U, Friedrich P, Selmer T, Buckel W, Golding BT.** 2004. Stereochemistry of hydrogen removal from the "unactivated" C-3 position of 4-hydroxybutyryl-CoA catalysed by 4-hydroxybutyryl-CoA dehydratase. Chem .Commun. **10:**1210-1211.

38. **Maezato Y, Johnson T, McCarthy S, Dana K, Blum P.** 2012. Metal resistance and lithoautotrophy in the extreme thermoacidophile *Metalosphaera sedula*. J. Bacteriol. **194:** 6856-63.

39. **Fiala G, Stetter KO.** 1986. *Pyrococcus furiosus* sp. nov. represents a novel genus of marine heterotrophic archaebacteria growing optimally at 100°C. Arch. Microbiol. **145:**56-61.

40. **Sun J, Hopkins RC, Jenney FE, McTernan PM, Adams MWW, Rodrigues-Lima F.** 2010. Heterologous expression and maturation of an NADP-dependent [NiFe]-hydrogenase: A key enzyme in biofuel production. PLoS One **5:**e10526.

41. **Schut GJ, Bridger SL, Adams MWW.** 2007. Insights into the metabolism of elemental sulfur by the hyperthermophilic archaeon *Pyrococcus furiosus*: characterization of a coenzyme A-dependent NAD(P)H sulfur oxidoreductase. J. Bacteriol. **189:**4431-41.

42. **Sapra R, Bagramyan K, Adams MWW.** 2003. A simple energy-conserving system: proton reduction coupled to proton translocation. Proc. Natl. Acad. Sci. U. S. A. **100:**7545.

43. **Ma K, Weiss R, Adams MWW.** 2000. Characterization of hydrogenase II from the

hyperthermophilic archaeon *Pyrococcus furiosus* and assessment of its role in sulfur reduction. J. Bacteriol. **182:**1864.

44. **Ma K, Adams MWW.** 2001. Hydrogenases I and II from *Pyrococcus furiosus*. Methods Enzymol. **331:**208.

45. **Silva PJ, Van Den Ban ECD, Wassink H, Haaker H, de Castro B, Robb FT, Hagen WR.** 2000. Enzymes of hydrogen metabolism in *Pyrococcus furiosus*. Eur. J. Biochem. **267:**6541-6551.

46. **Erb TJ.** 2011. Carboxylases in natural and synthetic microbial pathways. Appl. Environ. Microbiol. **77:**8466-77.

47. **Aresta M, Dibenedetto A.** 2007. Utilisation of CO_2 as a chemical feedstock: opportunities and challenges. Dalton Trans. **28:**2975-2992.

48. **Hopkinson RJ, Barlow PS, Schofield CJ, Claridge TDW.** 2010. Studies on the reaction of glutathione and formaldehyde using NMR. Org. Biomol. Chem. **8:**4915-4920.

49. **Ragsdale SW.** 2008. Enzymology of the Wood-Ljungdahl pathway of acetogenesis. Ann. N. Y. Acad. Sci. **1125:**129-136.

50. **Badger MR, Bek EJ.** 2008. Multiple Rubisco forms in proteobacteria: their functional significance in relation to CO_2 acquisition by the CBB cycle. J. Exp. Bot. **59:**1525-1541.

51. **Hügler M, Krieger RS, Jahn M, Fuchs G, Hügler M.** 2003. Characterization of acetyl-CoA/propionyl-CoA carboxylase in *Metallosphaera sedula*. Eur. J. Biochem. **270:**736-744.

52. **Berg IA, Kockelkorn D, Buckel W, Fuchs G.** 2007. A 3-hydroxypropionate/4-hydroxybutyrate autotrophic carbon dioxide assimilation pathway in Archaea. Science. **318:**1782.

53. **Huber H, Gallenberger M, Jahn U, Eylert E, Berg IA, Kockelkorn D, Eisenreich W, Fuchs G.** 2008. A dicarboxylate/4-hydroxybutyrate autotrophic carbon assimilation cycle in the hyperthermophilic Archaeum *Ignicoccus hospitalis*. Proc. Natl. Acad. Sci. U. S. A. **105:**7851-6.

54. **Kletzin A, Adams MW.** 1996. Molecular and phylogenetic characterization of pyruvate and 2-ketoisovalerate ferredoxin oxidoreductases from *Pyrococcus furiosus*

and pyruvate ferredoxin oxidoreductase from *Thermotoga maritima*. J. Bacteriol. **178:**248-257.

55. **Yamamoto M, Arai H, Ishii M, Igarashi Y.** 2006. Role of two 2-oxoglutarate:ferredoxin oxidoreductases in *Hydrogenobacter thermophilus* under aerobic and anaerobic conditions. FEMS Microbiol. Lett. **263:**189-193.

56. **Yoshizawa Y, Toyoda K, Arai H, Ishii M, Igarashi Y.** 2004. CO_2-responsive expression and gene organization of three ribulose-1,5-bis-phosphate carboxylase/oxygenase enzymes and carboxysomes in *Hydrogenovibrio marinus* strain MH-110. J. Bacteriol. **186:**5685-5691.

57. **Holo H.** 1989. *Chloroflexus aurantiacus* secretes 3-hydroxypropionate, a possible intermediate in the assimilation of CO_2 and acetate. Arch. Microbiol. **151:** 252-6.

58. **Ramos-Vera WH, Berg IA, Fuchs G.** 2009. Autotrophic carbon dioxide assimilation in *Thermoproteales revisited*. J. Bacteriol. **191:**4286-97.

59. **Bar-Even A, Noor E, Lewis NE, Milo R.** 2010. Design and analysis of synthetic carbon fixation pathways. Proc. Natl. Acad. Sci. U. S. A. **107:**8889-94.

60. **Raines CA.** 2006. Transgenic approaches to manipulate the environmental responses of the C3 carbon fixation cycle. Plant. Cell Environ. **29:**331-339.

61. **Peters-Wendisch PG, Schiel B, Wendisch VF, Katsoulidis E, Möckel B, Sahm H, Eikmanns BJ.** 2001. Pyruvate carboxylase is a major bottleneck for glutamate and lysine production by *Corynebacterium glutamicum*. J. Mol. Microbiol. Biotechnol. **3:**295-300.

62. **Millard CS, Chao YP, Liao JC, Donnelly MI.** 1996. Enhanced production of succinic acid by overexpression of phosphoenolpyruvate carboxylase in *Escherichia coli*. Appl. Environ. Microbiol. **62:**1808-1810.

63. **Erb TJ, Brecht V, Fuchs G, Müller M, Alber BE.** 2009. Carboxylation mechanism and stereochemistry of crotonyl-CoA carboxylase/reductase, a carboxylating enoyl-thioester reductase. Proc. Natl. Acad. Sci. U. S. A. **106:**8871-8876.

64. **Assary RS, Broadbelt LJ.** 2011. Computational screening of novel thiamine-catalyzed decarboxylation reactions of 2-keto acids. Bioprocess Biosyst. Eng. **34:**375-388.

65. **Tong L.** 2005. Acetyl-coenzyme A carboxylase: crucial metabolic enzyme and attractive target for drug discovery. Cell. Mol. Life Sci. **62:**1784-1803.

66. **Chapman-Smith A, Jr. JEC, Cronan JE.** 1999. *In vivo* enzymatic protein biotinylation. Biomol. Eng. **16:**119-25.

67. **Streaker ED, Beckett D.** 2006. Nonenzymatic biotinylation of a biotin carboxyl carrier protein: unusual reactivity of the physiological target lysine. Protein Sci. **15:**1928-1935.

68. **Peters-Wendisch P, Stansen KC, Götker S, Wendisch VF.** 2012. Biotin protein ligase from *Corynebacterium glutamicum*: role for growth and L-lysine production. Appl. Microbiol. Biotechnol. **93:**2493-502.

69. **Ko YT, Chipley JR.** 1984. Role of biotin in the production of lysine by *Brevibacterium lactofermentum*. Microbios. **40:**161-171.

70. **Cannon GC, Heinhorst S, Kerfeld CA.** 2010. Carboxysomal carbonic anhydrases: Structure and role in microbial CO_2 fixation. Biochim. Biophys. Acta **1804:**382-92.

71. **Badger MR, Price GD.** 1994. The role of carbonic anhydrase in photosynthesis. Annu. Rev. Plant Physiol. Plant Mol. Biol. **45:**369-392.

72. **Badger MR, Price GD.** 2003. CO_2 concentrating mechanisms in cyanobacteria: molecular components, their diversity and evolution. J. Exp. Bot. **54:**609-622.

73. **Tabita FR.** 1999. Microbial ribulose 1,5-bisphosphate carboxylase/oxygenase: A different perspective. Photosynth. Res. **60:**1-28.

74. **Zabaleta E, Martin MV, Braun H-P.** 2012. A basal carbon concentrating mechanism in plants? Plant Sci. **187:**97-104.

75. **Dou Z, Heinhorst S, Williams EB, Murin CD, Shively JM, Cannon GC.** 2008. CO_2 fixation kinetics of *Halothiobacillus neapolitanus* mutant carboxysomes lacking carbonic anhydrase suggest the shell acts as a diffusional barrier for CO_2. J. Biol. Chem. **283:**10377-84.

76. **Kusian B, Sültemeyer D, Bowien B, Su D.** 2002. Carbonic anhydrase is essential for growth of *Ralstonia eutropha* at ambient $CO_{(2)}$ concentrations. J. Bacteriol. **184:**5018-5026.

77. **Smith KS, Ferry JG.** 1999. A plant-type (beta-class) carbonic anhydrase in the thermophilic methanoarchaeon *Methanobacterium thermoautotrophicum*. J. Bacteriol. **181:**6247-53.

78. **Simpson PG, Withman WB.** 1993. Anabolic pathways in methanogens, p. 445-472.

In Ferry JG(ed), Methanogenesis: ecology, physiology, biochemistry, and genetics. Chapman and Hall, New York, N.Y.

79. **Smith KS, Cosper NJ, Stalhandske C, Scott R a, Ferry JG.** 2000. Structural and kinetic characterization of an archaeal beta-class carbonic anhydrase. J. Bacteriol. **182:**6605-13.

80. **Stephen WR, Pierce E.** 2008. Acetogenesis and the Wood-Ljungdahl pathway of CO_2 fixation. Biochim. Biophys. Acta. **1784:** 1873-1898.

81. **Schnappauf G, Braus GH, Gössner AS, Drake HL, Braus-Stromeyer SA.** 1997. Carbonic anhydrase in *Acetobacterium woodii* and other acetogenic bacteria. J. Bacteriol. **179:**7197-7200.

82. **Smith KS, Jakubzick C, Whittam TS, Ferry JG.** 1999. Carbonic anhydrase is an ancient enzyme widespread in prokaryotes. Proc. Natl. Acad. Sci. U. S. A. **96:**15184-9.

83. **Aden A, Bozell J, Holladay J, J.White, Manheim A, Elliot D, Lasure L, Jones S, Gerber M, Ibsen K, Lumberg L, Kelley S.** 2004. Top value-added chemicals from biomass: Results of screening for potential candidates from sugars and synthesis gas. U.S. Dep. Energy **1**:13-19.

84. **Jiang X, Meng X, Xian M.** 2009. Biosynthetic pathways for 3-hydroxypropionic acid production. Appl. Microbiol. Biotechnol. **82:**995-1003.

85. **Mochizuki M, Hirami M.** 1997. Structural effects on the biodegradation of aliphatic polyesters. Polym. Adv. Technol. **8:**203-209.

86. **Meng XS, Abraham T, Tsobanakis P.** 2007. Process for preparing 3-hydroxycarboxylic acids. Patent US20070015936A1.

87. **Behr A, Botulinski A, Carduck F-J, Schneider M.** 1996. process for preparing 3-hydroxypropionic acid. Patent EP0579617A1.

88. **Kumar V, Ashok S, Park S.** 2013. Recent advances in biological production of 3-hydroxypropionic acid. Biotechnol. Adv. **31:** 945-961.

89. **Pina CD, Falletta E RM.** 2011. A green approach to chemical building blocks. The case of 3-hydroxypropanoic acid. Green Chem. **13:**1624-32.

90. **Talarico TL, Casas IA, Chung TC, Dobrogosz WJ.** 1988. Production and isolation of reuterin, a growth inhibitor produced by *Lactobacillus reuteri*. Antimicrob. Agents Chemother. **32:**1854-1858.

91. **Garai-Ibabe G, Ibarburu I, Berregi I, Claisse O, Lonvaud-Funel A, Irastorza A, Dueñas MT.** 2008. Glycerol metabolism and bitterness producing lactic acid bacteria in cidermaking. Int. J. Food Microbiol. **121:**253-261.

92. **Ansede JH, Pellechia PJ, Yoch DC.** 1999. Metabolism of acrylate to beta-hydroxypropionate and its role in dimethylsulfoniopropionate lyase induction by a salt marsh sediment bacterium, *Alcaligenes faecalis* M3A. Appl. Environ. Microbiol. **65:**5075-5081.

93. **Takamizawa K, Horitsu H, Ichikawa T, Kawai K, Suzuki T.** 1993. β-hydroxypropionic acid production by *Byssochlamys* sp. grown on acrylic acid. Appl. Microbiol Biotechnol. **40:**196-200.

94. **Dave H, Ramakrishna C, Desai JD.** 1996. Degradation of acrylic acid by fungi from petrochemical activated sludge. Biotechnol. Lett. **18:**963-4.

95. **Andersen G, Björnberg O, Polakova S, Pynyaha Y, Rasmussen A, Møller K, Hofer A, Moritz T, Sandrini MPB, Merico A-M, Compagno C, Akerlund H-E, Gojković Z, Piskur J.** 2008. A second pathway to degrade pyrimidine nucleic acid precursors in eukaryotes. J. Mol. Biol. **380:**656-66.

96. **Saini R, Kapoor R, Kumar R, Siddiqi TO, Kumar A.** 2011. CO_2 utilizing microbes - A comprehensive review. Biotechnol. Adv. **29:**949-60.

97. **Gokarn RR, Selifonova OV, Jessen HJ, Steven JG, Selmer T, BuckelW.** 2001. 3-hydroxypropionic acid and other organic compounds. Patent US7186541B2.

98. **Henry CS, Broadbelt LJ, Hatzimanikatis V.** 2010. Discovery and analysis of novel metabolic pathways for the biosynthesis of industrial chemicals: 3-hydroxypropanoate. Biotechnol. Bioeng. **106:**462-473.

99. **Liao HH, Gokarn RR, Gort SJ, Jessen HJ, Selifonova OV.** 2005. Alanine 2,3-aminomutase. Patent EP1575881B1.

100. **LiaoHH,Gokarn RR,Gort SJ, JessenHJ, SelifonovaOV.** 2007. Production of 3-hydropropionic acid using beta-alanine/pyruvate aminotransferase. Patent US20070107080A1.

101. **Marx A, Wendisch VF, Rittmann D, Buchholz S.** 2007. Microbiological production of 3-hydroxypropionic acid. Patent WO2007042494A3.

102. **Van Maris AJA, Konings WN, van Dijken JP, Pronk JT.** 2004. Microbial export of

lactic and 3-hydroxypropanoic acid: implications for industrial fermentation processes. Metab. Eng. **6:**245-255.

103. **Rathnasingh C, Raj SM, Lee Y, Catherine C, Ashok S, Park S, Mohan S.** 2012. Production of 3-hydroxypropionic acid via malonyl-CoA pathway using recombinant *Escherichia coli* strains. J. Biotechnol. **157:**633-40.

104. **Lynch MD, Gill RT, Warnecke LT.** 2011. Method for producing 3-hydroxypropionic acid and other products. Patent WO2011038364A1.

105. **Jessen HJ, Liao HH, Steven J, Selifonova OV.** 2011. Beta-alanine/alpha-ketoglutarate aminotransferase for 3-hydroxypropionic acid production. Patent US8030045B2.

106. **Liao HH,Gokarn RR, Gort SJ, Jessen HJ, Selifonova O.** 2003. Nucleic acid and polypeptide sequences for alanine 2,3-aminomutase and their use for production of β-alanine, pantothenate, 3-hydroxypropionic acid, and 1,3-propanediol. Patent WO03/062173A2.

107. **Luo LH, Kim CH, Heo S-Y, Oh B-R, Hong W-K, Kim S, Kim D-H, Seo J-W.** 2012. Production of 3-hydroxypropionic acid through propionaldehyde dehydrogenase PduP mediated biosynthetic pathway in *Klebsiella pneumoniae*. Bioresour. Technol. **103:**1-6.

108. **Huang Y, Li Z, Shimizu K, Ye Q.** 2012. Simultaneous production of 3-hydroxypropionic acid and 1,3-propanediol from glycerol by a recombinant strain of *Klebsiella pneumoniae*. Bioresour. Technol. **103:**351-9.

109. **Jo J-E, Mohan Raj S, Rathnasingh C, Selvakumar E, Jung W-C, Park S.** 2008. Cloning, expression, and characterization of an aldehyde dehydrogenase from *Escherichia coli* K-12 that utilizes 3-Hydroxypropionaldehyde as a substrate. Appl. Microbiol. Biotechnol. **81:**51–60.

110. **Ko Y, Ashok S, Zhou S, Kumar V, Park S.** 2012. Aldehyde dehydrogenase activity is important to the production of 3-hydroxypropionic acid from glycerol by recombinant *Klebsiella pneumoniae*. Process. Biochem. **47:**1135-43.

111. **Luo LH, Seo J-W, Baek J-O, Oh B-R, Heo S-Y, Hong W-K, Kim D-H, Kim CH.** 2011. Identification and characterization of the propanediol utilization protein PduP of *Lactobacillus reuteri* for 3-hydroxypropionic acid production from glycerol. Appl. Microbiol. Biotechnol. **89:**697-703.

112. **Raj SM, Rathnasingh C, Jung WC, Selvakumar E, Park S.** 2010. A novel NAD+-

dependent aldehyde dehydrogenase encoded by the puuC gene of *Klebsiella pneumoniae* DSM 2026 that utilizes 3-hydroxypropionaldehyde as a substrate. Biotechnol. Bioprocess. Eng. **15:**131-8.

113. **Rathnasingh C, Raj SM, Jo J-E, Park S.** 2009. Development and evaluation of efficient recombinant *Escherichia coli* strains for the production of 3-hydroxypropionic acid from glycerol. Biotechnol. Bioeng. **104:**729-739.

114. **Mohan Raj S, Rathnasingh C, Jung W-C, Park S.** 2009. Effect of process parameters on 3-hydroxypropionic acid production from glycerol using a recombinant *Escherichia coli*. Appl. Microbiol. Biotechnol. **84:**649-657.

115. **Daniel R, Bobik TA, Gottschalk G.** 1998. Biochemistry of coenzyme B12-dependent glycerol and diol dehydratases and organization of the encoding genes. FEMS Microbiol. Rev. **22:**553-566.

116. **Toraya T.** 2002. Enzymatic radical catalysis: coenzyme B12-dependent diol dehydratase. Chem. Rec. **2:**352-66.

117. **Xu XL, Zhang GL, Wang LW, Ma BB, Li C.** Quantitative analysis on inactivation and reactivation of recombinant glycerol dehydratase from *Klebsiella pneumoniae* XJPD-Li. J. Mol. Catal. B: Enzym. **56:**108-14.

118. **Mori K, Tobimatsu T, Hara T, Toraya T.** 1997. Characterization, sequencing, and expression of the genes encoding a reactivating factor for glycerol-inactivated adenosylcobalamin-dependent diol dehydratase. J. Biol. Chem. **272:**32034-32041.

119. **Kwak S, Park Y-C, Seo J-H.** 2013. Biosynthesis of 3-hydroxypropionic acid from glycerol in recombinant *Escherichia coli* expressing *Lactobacillus brevis* dhaB and dhaR gene clusters and *E. coli* K-12 aldH. Bioresour. Technol. **135:**432-9.

120. **Celińska E.** 2012. As a 1, 3-propanediol producer-the metabolic engineering approach. Crit. Rev. Biotechnol. **32:**274-288.

121. **Ashok S, Raj SM, Rathnasingh C, Park S.** 2011. Development of recombinant *Klebsiella pneumoniae* ΔdhaT strain for the co-production of 3-hydroxypropionic acid and 1,3-propanediol from glycerol. Appl. Microbiol. Biotechnol. **90:**1253-1265.

122. **Huang Y, Li Z, Shimizu K, Ye Q.** 2013. Co-production of 3-hydroxypropionic acid and 1,3-propanediol by *Klebseilla pneumoniae* expressing aldH under microaerobic conditions. Bioresour. Technol. **128:**505-12.

TABLES

Table 1. Enzymes of 3-hydroxypropionate/4-hydroxybutyrate cycle

Enzyme #	Enzymes from 3-HP/4-HB Pathway	Msed ORF
E1 (α, β, γ)	Acetyl/propionyl-CoA carboxylase	Msed_0147(a)
		Msed_0148(b)
		Msed_1375(g)
E2	Malonyl-CoA/Succinyl-CoA reductase	Msed_0709
E3	Malonate semialdehyde reducatse	Msed_1993
E4	3-Hydroxypropionyl-CoA synthetase	Msed_1456
E5	3-Hydroxypropionyl-CoA dehydratase	Msed_2001
E6	Acryloyl-CoA reductase	Msed_1426
E7	Methylmalonyl-CoA epimerase	Msed_0639
E8 (α)	Methylmalonyl-CoA mutase-catalytic subunit	Msed_0638
E8 (β)	Methylmalonyl-CoA mutase-cobalamin B12-binding subunit	Msed_2055
E9	Succinate semialdehyde reductase	Msed_1424
E10	4-Hydroxybutyrate-CoA ligase	Msed_0394
		Msed_0406
		Msed_1353
E11	4-Hydroxybutyrl-CoA dehydratase	Msed_1321
E12	Crotonyl-CoA hydratase/ (S)-3-hydroxybutyrl-CoA dehydrogenase	Msed_0399
E13	Acetoacetyl-CoA b-ketothiolase	Msed_0656

<table>
<tr><th colspan="3">Table 2. Classification of carboxylases based on different functions.</th></tr>
<tr><th>Group</th><th colspan="2">Function</th></tr>
<tr><td>Autotrophic carboxylases</td><td rowspan="2">Synthesis of central precursors (PEP, pyruvate, acetyl-CoA, TCA cycle intermediates)</td><td>Start from inorganic carbon</td></tr>
<tr><td>Assimilatory carboxylases</td><td>Start from organic carbon</td></tr>
<tr><td>Anaplerotic carboxylases</td><td colspan="2">Refilling of intermediates in TCA cycle</td></tr>
<tr><td>Biosynthetic carboxylases</td><td colspan="2">Synthesis of cellular building blocks starting from central precursor molecules</td></tr>
<tr><td>Redox-balancing carboxylases</td><td colspan="2">Maintain redox balance</td></tr>
</table>

Table 3. Maximal production of 3-HP in recombinant *E. coli* and *K. pneumonia* (Adapted from (89)**).**

Organism	Carbon source	Aeration condition	Titer (g/L)	Yield (mol/mol)	Productivity (g/L/h)	Ref
E. coli	glucose	Aerobic	49.0	0.46	0.71	(104)
K. pneumoniae	glycerol	Microaerobic	48.9	0.41	1.75	(122)

FIGURES

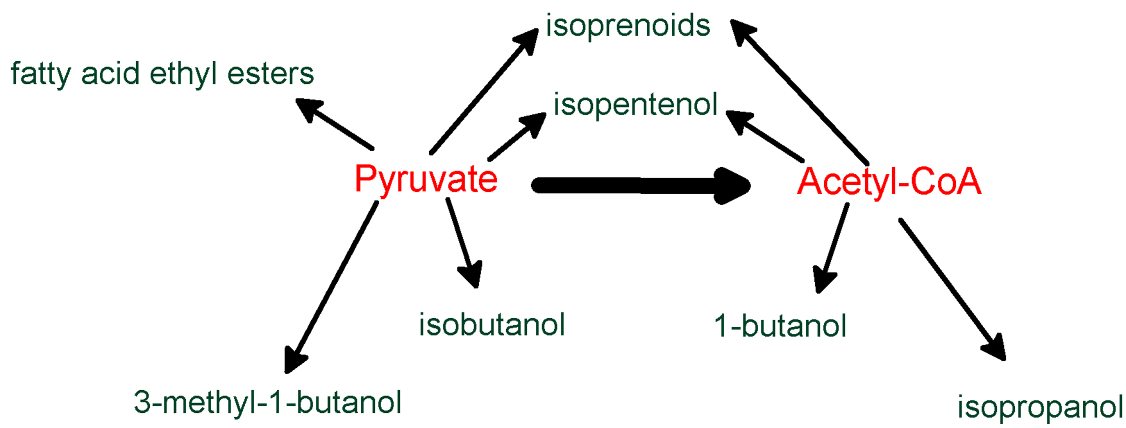


Figure 1. Pyruvate and acetyl-CoA as precursors for liquid fuel molecules.

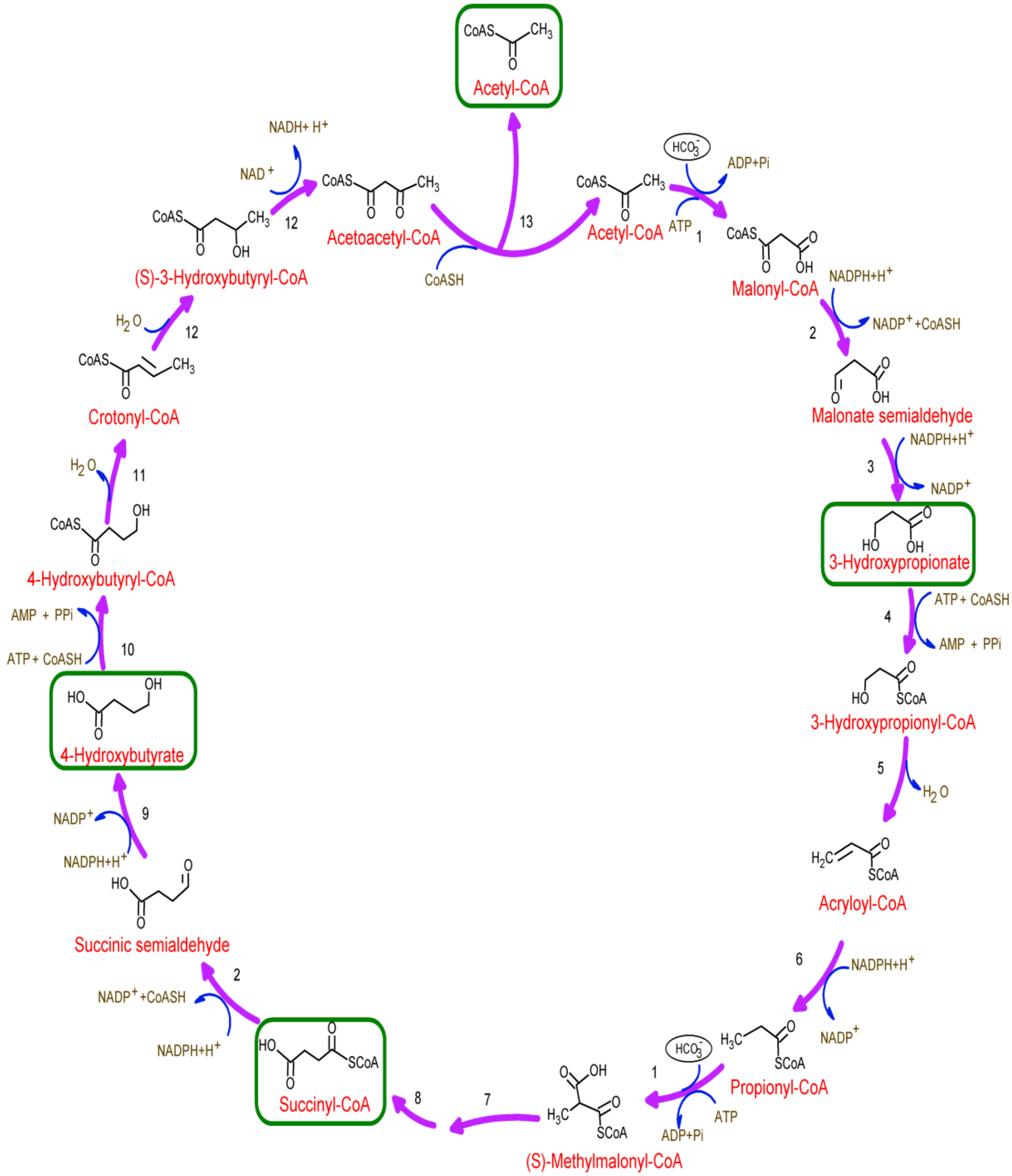


Figure 2. 3-hydroxypropionate/4-hydroxybutyrate cycle. The numbers for each reaction step refer to enzymes in Table 1.

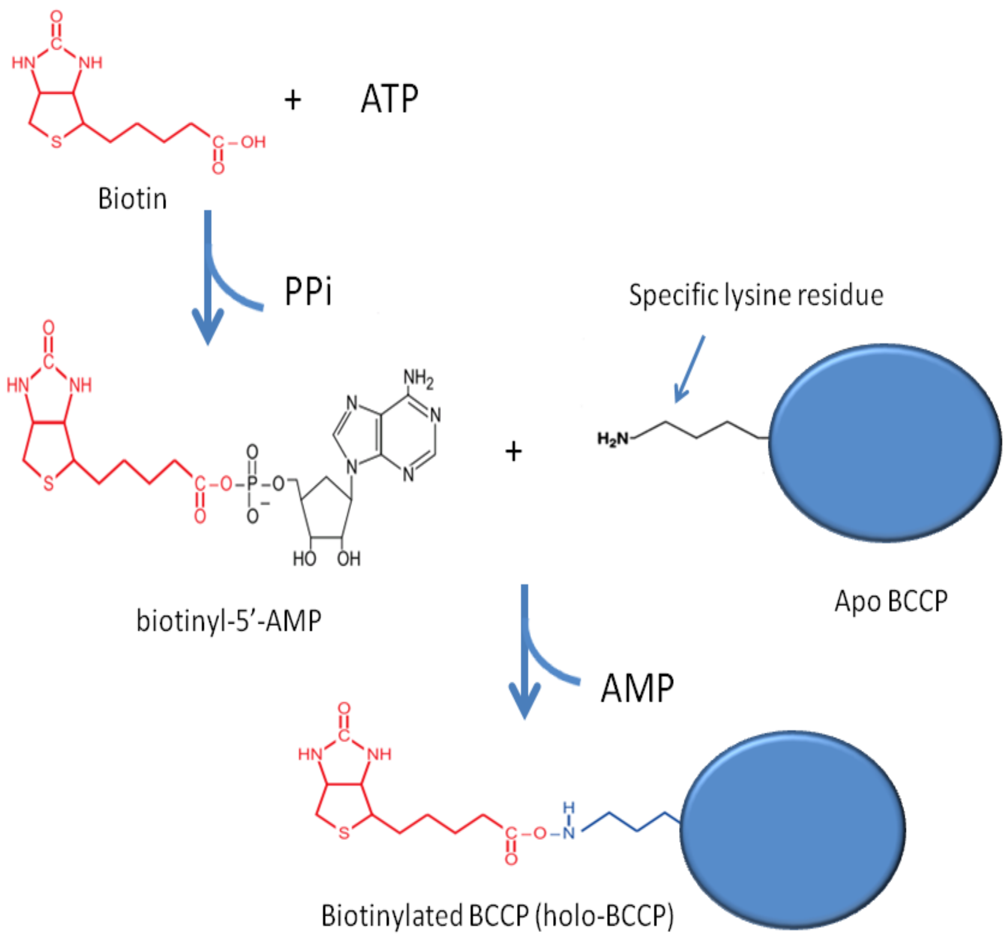


Figure 3. The biotin protein ligase reaction mechanism. (Adapted from (66)**)**

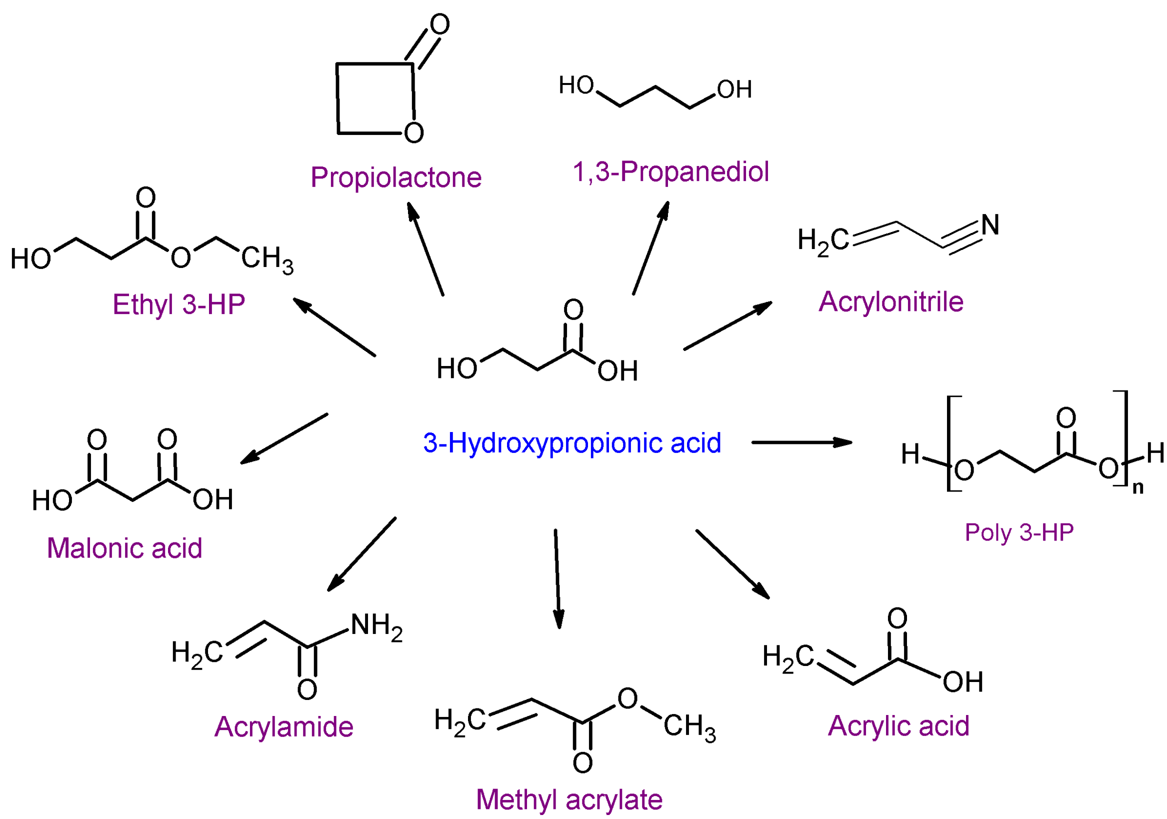


Figure 4. 3-hydroxypropionate (3-HP) as precursor for valuable chemicals (Adapted from (83)**).**

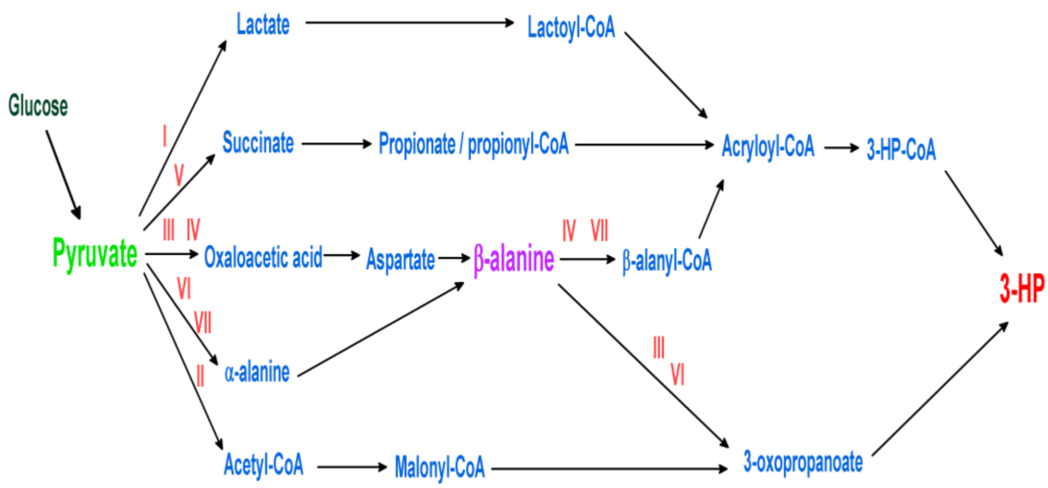


Figure 5. 3-HP synthesis from glucose via seven metabolic pathways (as labeled in I to VII) (Adapted from (88)**).**

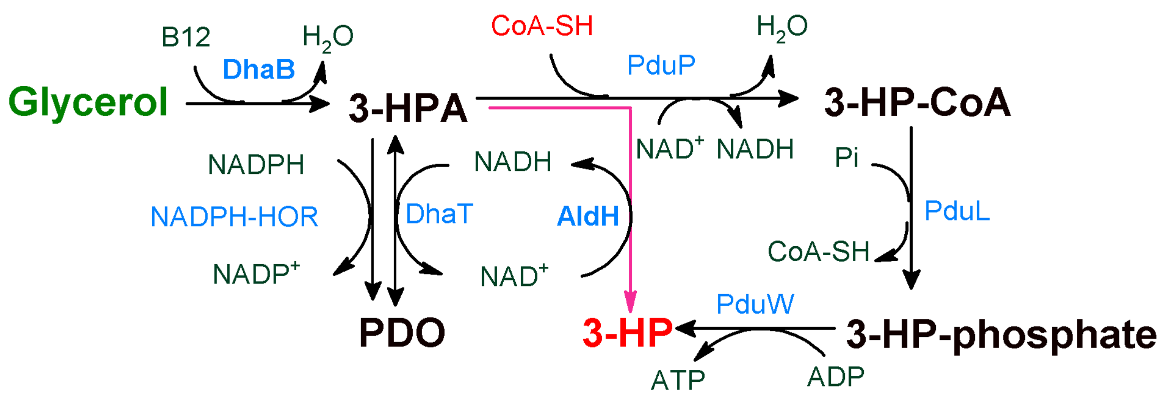


Figure 6. 3-HP synthesis from glycerol via CoA-dependent (reductive pathway) or CoA-independent pathway. DhaB, glycerol dehydratase, **PduP,** propionaldehyde dehydrogenase, **AldH,** aldehyde dehydrogenase, **PDO** 1,3-propanediol **(Adapted from** (88) **).**

Chapter 2

Carbonic Anhydrase (CA) from the Extremely Thermoacidophilic Archaeon *Metallosphaera sedula* (Msed_0390, CAMS) Represents a New sub-class of β-Carbonic Anhydrases

ABSTRACT

Msed_0390, encoded in the genome of the extremely thermoacidophilic archaeon *Metallosphaera sedula*, was shown to encode a functional carbonic anhydrase (CA), designated as CAMS. This microorganism utilizes the 3-hydroxypropionate/4-hydroxybutyrate (3-HP/4-HB) cycle for carbon fixation under autotrophic conditions. Comparison of this enzyme with other known classes of CAs based on primary amino acid sequence and subunit composition suggests that it belongs to a new subclass of β-CAs. In view of its origin from thermophilic archaea, this enzyme may exhibit promising applications in carbon dioxide capture and sequestration technology for processing CO_2 emissions from coal combustion and other waste gases with high acid gas content. The possible involvement of CAMS in non-photosynthetic assimilation of inorganic carbon from thermoacidophilic environments and participation in carbon concentrating mechanisms (CMMs) is considered. This work provides the basis for incorporating this enzyme by metabolic engineering into the hyperthermophile *Pyrococcus furiosus* to facilitate 3-HP/4-HB cycles functioned for production of fuels and chemicals.

Introduction

Current Classification of Carbonic Anhydrases (CA)

Carbonic anhydrases (CAs, EC 4.2.1.1) catalyze the reversible hydration of carbon dioxide: $CO_2 + H_2O \leftrightarrow HCO_3^- + H^+$ (1). CAs are metalloenzymes distributed among all domains of life: Archaea, Eukarya and Bacteria. Based on primary amino acid sequence and three-dimensional structure, CAs are currently categorized into five classes (α, β, γ, δ and ζ) (2, 3) (see **Table 1**). The α and γ class CAs both contain a catalytically essential metal ion (zinc for α class, zinc/iron for γ class) coordinated by three histidine residues (and one additional water molecule for α class, two additional water molecules for γ class). The β-class CAs have an active site zinc coordinated by two cysteines and one histidine residue (4), and a variant fourth ligand (5). The α, β, and γ CAs seem to share a common zinc hydroxide mechanism for catalysis. This is indicated by kinetic analyses of a γ-class CA (CAM) from *Methanosarcina thermophila* (6), the 'plant' type β-class CA from *Pisum sativum* (7), the 'cab' type β-class CA (CAB) from *Methanobacterium thermoautotrophicum* (4), and the human α-class isozymes CA I, II and III (8,9). The first step of this mechanism is the interconversion between CO_2 and HCO_3^-, triggered by the nucleophilic attack of a zinc-bound hydroxide ion on CO_2, in which the rate is described by the steady-state parameter k_{cat}/K_m. The second step involves regeneration of the active site by the loss of a hydrogen ion from zinc-bound water, reflected in the steady-state parameter k_{cat} (8,9).

The five CA classes have different evolutionary roots (10). The α class CAs are monomeric enzymes, found mostly in mammals with a few in bacteria and algae, but are missing in the Archaea (8). β-CAs are the most ubiquitous class, and are described in multiple lineages, ranging from plants to algae to bacteria to archaea (2). CAM is the archetype of γ-class CA, isolated and characterized from the methanoarchaeon *Methanosarcina thermophila* (11). Phylogenetic analysis of CAM homologs derived from the databases shows wide distribution of γ-class in all three domains of life.

Two novel CAs have been identified in the marine diatom *Thalassiosira weissflogii*, TWCA1 and CDCA1, which were classified as δ and ζ classes, respectively (12,13). The monomeric TWCA1 shares no sequence homology to α or γ classes. Its catalytic zinc is coordinated by three histidines, resembling α and γ CAs, along with an active site geometry similar to that of α class. This zinc ion can be substituted by Co(II) *in vivo* under Zn-limited conditions (14). The similar active site dimension between α and δ CAs allows for access of bulkier substrates, such as p-nitrophenyl acetate (PNPA), so only these two classes of CAs exhibit esterase activity besides the canonical CO_2 hydration/dehydration activity (15). With a water molecule as a fourth ligand, this enzyme probably follows the zinc hydroxide mechanism, found in α, β, and γ CAs, but this has yet to be confirmed through steady-state kinetics (16). Unlike the δ-CAs widely distributed in marine phytoplankton, ζ-class CA, represented by CDCA1, is found only in marine diatoms. It can readily exchange cadmium and zinc at its active site. This zinc to toxic cadmium replacement is presumably a significant competitive advantage in the metal-poor environment of the oceans (17). The active site conformation in CDCA1 closely resembles that of the β-CA from *P. sativum*, including the conserved two cysteines and one histidine coordinating cadmium and an aspartate and arginine pair. CDCA1 consists of three nearly identical repeats, with each repeat exhibiting robust catalytic activity as a monomer. CDCA1-repeat2 shares little similarity with the β-class CA monomer, but presents as a structural and functional mimic of β-class CA dimer (17).

β-Carbonic Anhydrases

This class of CA can be further divided into three sub-classes: the ‘plant’ type, typically represented by CA from the plant *P. sativum* (18), the ‘cab’ type, represented by CAB, a CA in the thermophilic archaea *Methanobacterium thermoautotrophicum* (19), and the ‘carboxysome’ type, represented by CsoSCA, a carboxysome-associated CA from the chemolithoautotroph *Halothiobacillus neapolitanus* (20). The ‘plant’ type is also found in bacteria, while the ‘Cab’ type forms an exclusively prokaryotic phylogenetic clade composed primarily of sequences from Gram-positive bacteria (21). Actually, CsoSCA was initially annotated as an ε-class CA (20), due to the primary sequence divergence with other known classes. It is localized in the shell of a cellular microcompartment, called the ‘carboxysome’. However, the crystal structure of CsoSCA reveals remarkable structural similarity to known β-CAs in terms of fold and active site region, so it was re-classified as a member of β-CAs (22).

All three types of β-CAs contain the three conserved zinc ligands (one histidine and two cysteines) and two other conserved active site residues (aspartate/arginine pairs), which presumably facilitate substrate binding, proton shuffling, and product release (22). The ‘plant’ and ‘cab’ subclasses are much more similar to each other than they are to the ‘carboxysome’ type (22). The major difference between ‘Cab’ and ‘plant’ types are five active site residue substitutions in their active sites. The ‘Cab’ active site residues His23, Met33, Lys53, Ala58 and Val72 are replaced in the ‘plant’ type β-CA by Gln, Ala, Phe, Val and Tyr, respectively, suggesting a more open and less hydrophobic pocket for CO_2 binding in CAB (5). Besides an almost complete divergent primary sequence from ‘plant’ type and ‘Cab’ type, CsoSCA distinguishes itself by a lack of active site pairing maintained in the former two types, which is organized within a two-fold symmetric homodimer or pair of fused, homologous domains (22).

γ-class Carbonic Anhydrases

CAM is the γ-class archetype CA, isolated and characterized from the anaerobic methanoarchaeon *M. thermophila* (11). The heterologously expressed CAM in the closely related *Methanosarcina acetivorans* had full incorporation of Fe^{2+} as the active site metal (23), instead of Zn^{2+}, observed when initially overproduced and aerobically purified from *E. coli* (24). One explanation is that CA rapidly loses activity upon exposure of Fe^{2+} to air and the aerobic purification leads to the loss of Fe^{3+} (the oxidized iron) and replacement by Zn^{2+} in the buffer not treated with metal chelating agents (25). On the other hand, recombinant production of archaeal proteins in eukaryotic or bacterial hosts may impede biochemical function, associated with problems in inserting active metal centers (25,26,27). Given that iron-containing CAM has several-fold greater activity than the zinc-containing enzyme, CAM likely contains iron when synthesized *in vivo*, taking advantage of Fe^{2+}-rich anaerobic environments of methane-producing species (23).

To date, CAM has been the only γ-class CA biochemically characterized, although homologs have been identified by sequence alignment in databases. CAM contains three active site metal ligands (His 81, His 117, and His 122) and catalytically important residues (Gln 75, Gln73, Glu 62, Glu 84 and Asn 202). Analysis of homolog sequences actually reveal two sub-classes: a minor sub-class ‘Cam’ represented by CAM, and a major sub-class with conservation of CAM residues essential for metal binding and catalysis, except for Glu62 and the acidic loop residues, including Glu84 (25). The latter is designated as ‘CamH’, exemplified by a homolog enzyme CAP, from *Pyrococcus horikoshii*, a fermentatively anaerobic archaeon optimally growing at 98°C, and not a methanogen. Both CAM and CAP have been overproduced in *E. coli* and the crystal structures of their recombinant form have been resolved respectively (28,29), showing similar overall fold and highly conserved active site architecture. They are both homotrimers, and both have three active site metals located at the interface between monomers; one monomer donates one histidine ligand and the adjacent monomer contributes the remaining two histidine ligands. Remarkably, Ca^{2+} ions have been observed bound to a polar valley in the CAP structure, in addition to the catalytic Zn^{2+} ion. This is the first time that Ca^{2+} has been found at the interface between the monomers among the different classes of CA. Perhaps, the absence of an acidic loop structure in CAP creates this second metal binding site, and the Ca^{2+} could play a role in substrate entry and product exit (29).

Functional Roles of Carbonic Anhydrases

As summarized by Kanth *et al.*, human α-CAs are involved in diverse physiological functions, including respiration and transport of CO_2/HCO_3^- in various metabolizing tissues, pH and CO_2 homeostasis, electrolyte secretion, biosynthetic reaction (gluconeogenesis, lipogenesis and ureagenesis), bone resorption and many other physiologic or pathologic processes. They are important therapeutic targets and, as such, their inhibitors and activators are currently used as drugs (30). It was assumed that α-CA is the least evolved phylogenetic group, while β and γ CAs are the most ancient of these enzymes (31), widespread in plants, bacteria and archaea, where they play critical roles in carbon concentration mechanisms (CCMs) to facilitate CO_2 fixation process and CO_2 transport for some important biosynthesis.

Specifically, CA may act as an extracellular or periplasmic enzyme that aids in the conversion of external bicarbonate to CO_2, for subsequent uptake by the cell. In four carbon (C4) and crassulacean acid metabolism (CAM) in higher plants, cytosolic CA conversely mediates the conversion of CO_2 to bicarbonate, which serves as the substrate for PEP carboxylase to synthesize C4 acids. C4 acids can then be transported into the bundle sheath and decarboxylated to release CO_2, finally re-fixed by RubisCO (32,33,34). A cytosolic CA could promote the diffusion of inorganic carbon (Ci) to the chloroplast envelope or convert bicarbonate to CO_2, if bicarbonate was accumulated in the cytosol. In some cases, CA spatially associates with RubisCO in capsular structures, such as the cyanobacterial carboxysome or the pyrenoid, rapidly dehydrating cytosolic bicarbonate pool to raise CO_2 levels for RubisCO to function in a fully saturated state (32,35).

CsoSCA, the 'carboxysome' type of β-CA, was found associated with α type carboxysomes (36). The importance of this association has been studied. Since the carboxysomal shell was considered to be the CO_2 diffusional barrier, this enzyme, as an integral component of the shell, facilitates entry of inorganic carbon (Ci) into the carboxysome interior, where RubisCO is encapsulated, through conversion of cytoplasmic bicarbonate to CO_2. In β-type carboxysomes, a β-CA, CcaA, has been identified as a likely component (37, 38), while another CA candidate, CcmM, was thought to be a homolog to an archaeal γ-CA (24, 39).

Exploration of CA in *Metallosphaera sedula* to facilitate 3-HP/4-HB cycle function (*in vivo* and *in vitro*)

As discussed in Chapter I, there has been almost no research reported for carbon concentration mechanisms (CCMs) and related CAs beyond the range of organisms containing photosynthetic or Rubisco-controlled carbon fixation process. There has been evidence that Rubisco and β-CA are transcriptionally linked in *P. sativum*. Both Rubisco and β-CA expression levels were up-regulated when plants are transferred from a high CO_2 environment to atmospheric (low) CO_2 (40). Previous studies reported an ORF in the extremely thermoacidophilic archaeon *Metallosphaera sedula* (Msed_0390) significantly up-regulated under autotrophic growth compared to heterotrophic growth (41), synchronized with transcription of genes encoding 3HP/4HB cycle enzymes. This ORF was annotated as a β-CA and transcriptional response, and corresponding expression, presumably correlates to inorganic carbon uptake by *M. sedula* for carbon fixation. This CA could be utilized to complement enzymes from the *M. sedula* 3HP/4HB cycle to convert CO_2 directly to fuel molecules or industrial chemicals in metabolic engineering hosts, such as *Pyrococcus furious* (*Pfu*). *Pfu* does not appear to natively express a highly functional CA to aid in the fast conversion of CO_2 to bicarbonate, the substrate of 3HP/4HB cycle. So far, the only evidence for CA in *Pfu* was the detection of cross-reacting proteins in *Pfu* cell extract to antisera raised from typical known CAs in Western blots (21). However, the *Pfu* cell extract exhibited undetectable activity (< 0.01 Wilbur-Anderson unit at 55°C), which could have been related to the upper temperature limit of the assay (55°C) being well below the expected activity optimum (85°C) for a putative *Pfu* CA. Also, the decreased solubility of CO_2 at higher temperatures under atmospheric pressure impedes determination of CA activity. However, the crystal structure of a γ class CA has been reported in *P. horikoshii*, referred to as CAP, but no activity has been noted for this enzyme. Therefore, the capability of wild-type *Pfu* for efficient CO_2 uptake and bicarbonate conversion is questionable. Here, the CA from *M. sedula* was produced recombinantly in *E. coli* and characterized for its potential to work in conjunction with the 3-HP/4-HB cycle from *M. sedula* in metabolically engineered strains of *Pfu*, designed to produce fuels and industrial chemicals from CO_2.

Methods and Materials

Cloning and heterologous expression of *M. sedula* carbonic anhydrase in *E. coli.* The gene encoding the putative carbonic anhydrase in the *M. sedula* genome (Msed_0390) was amplified by using forward primer 5'-ACACAC<u>CATATG</u>AGAATCGTCATATCGTGTATGGAC-3' introducing an NdeI site at the initiation codon, and reverse primer 5'-ACACAC<u>CTCGAG</u> GCTAATCCTCTTTACATCCC-3' introducing an XhoI site after the stop codon. The PCR conditions were as follows: initial denaturation for 1 min at 95°C, 40 cycles of denaturation for 45s at 95°C, primer annealing for 30 sec at 61.2°C, and extension for 35 sec at 72°C, and a final extension for 5 min at 72°C. The PCR product was purified and cloned into pET21b (+) vector, resulting in plasmid pET21b-mse0390. The plasmid was transformed into Competent *E. coli* Rosetta (DE3) cells, capable of encoding rare tRNAs required for translation of *M. sedula* enzymes. Cells were grown in 1L LB media containing 100 μg of ampicillin ml^{-1} and 34 μg of chloramphenicol ml^{-1}. Expression was induced with 20 μM isopropyl-β-D-thiogalactopyranoside (IPTG) (optical density at 600 nm, 0.6); cells were harvested 4 to 5 hr after induction (37°C) by centrifuge and stored at -20 °C until they were used.

Purification of C-terminal 6×His-tagged carbonic anhydrase. *E. coli* cells were re-suspended in 20 mM sodium phosphate, 0.5 M NaCl, pH 7.4 and disrupted by sonication. Cell extract was incubated at 65°C for 20 min to denature *E.coli* native proteins, and the supernatant after centrifuge (16,000 × g, 15 min, 4°C) was collected for purification. The heat-treated cell extract was loaded on a HiTrap IMAC HP column (GE Healthcare Life Sciences; volume, 1 mL) equilibrated with 20 mM sodium phosphate, 0.5 M NaCl, at pH 7.4. The column was washed with ten-bed volume equilibration buffer and developed with a step gradient to 300 mM imidazole at a flow rate of 1 ml min^{-1}. The most active fractions eluted with imidazole were pooled and dialyzed in 50 mM Tris-HCl, pH 8.2 to remove imidazole. The purified protein solution was concentrated via Amicon® Ultra-15 Centrifugal Filter Units (3kDa) (EMD Milipore) and stored at -80°C.

Molecular weight determination. Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was performed, as described by Laemmli (42), using 4-12% gels (Invitrogen). Proteins were visualized by Coomassie blue staining (43). The native molecular mass of the enzyme was estimated by gel filtration chromatography. Protein from the IMAC step was applied to a Superdex 75 HR 16/60 gel filtration column (GE Healthcare Life Sciences; volume, 20 ml), which had been equilibrated with 50 mM sodium phosphate, 100 mM NaCl, pH 7.2. The flow rate was 0.5 ml min^{-1}. The column was calibrated with the following molecular mass standards (Sigma): β-Amylase from sweet potato (200 kDa), alcohol dehydrogenase from yeast (150 kDa), bovine serum albumin (66 kDa), carbonic anhydrase from bovine erythrocytes (29 kDa), cytochrome c from horse heart (12.4 kDa), and aprotinin from bovine lung (6.5 kDa).

CO_2 hydration: Wilbur-Anderson assay. The CO_2 hydration assay followed the Wilbur–Anderson protocol (44) with certain modifications. CO_2-saturated water was made by subliming dry ice in a glass bottle for 30 min, with the bottle sealed with a rubber stopper to prevent leakage. The CO_2-saturated water (3 ml) was immediately added to 2 ml of Tris–Sulfate buffer (100 mM; pH 8.3) containing 0.5 ml of enzyme solution. The enzyme solution was properly diluted to make sure the reaction rate falls within measurable limits. The reaction vessel was maintained at 10°C. The assay was carried out at low temperature to promote solubility of CO_2 in water, and to slow down the rate of reaction within measurable limits. The time (t) required for the pH to drop from 8.0 to 6.3, as catalyzed by the enzyme, was measured. The control for the pH change (8.0–6.3) used enzyme storage buffer (50 mM Tris-HCl, pH 8.2) substituted for the enzyme solution. The Wilbur–Anderson Units were calculated with equation: $1U = (t_c - t)/t$, where t_c is the time for control. The protein content was determined by the method of Bradford assay (45), and the activity was reported in U/ml/mg protein. All measurements were carried out in triplicate.

Esterase assay. Activity for *p*-nitrophenylacetate hydrolysis was determined at 25°C, using a modification of the method of Armstrong *et al*. (46). A 3 mM stock of p-nitrophenylacetate was freshly prepared for each set of assays by dissolving the compound in acetone and then diluting it 25-fold with water. The uncatalyzed rate was measured by adding 1.9 ml of 15 mM Tris-Sulfate, pH 7.6, to 1 ml of this substrate solution and recording the change in an absorbance (348 nm) per min ($\Delta\varepsilon = 5{,}000\ M^{-1}\ cm^{-1}$). After 3 min, 0.1 mL of enzyme solution was added, and the catalyzed reaction was monitored for additional 3 min. A control lacking enzyme was performed by adding 0.1 mL enzyme storage buffer (50 mM Tris-HCl, pH 8.2); carbonic anhydrase from bovine erythrocytes (Sigma) was used as a positive control.

Metal analysis. Metal analysis of CAMS was performed by using Varian 820 ICP-MS (Inductively-coupled plasma-mass spectrometer).

Bioinformatics. Nucleotide and amino acid sequences of CAs were obtained from databases at NCBI (http://www.ncbi.nlm.nih.gov). The alignments of amino acid sequences were performed with Clustal WX verstion 2.0 (47), and displayed with Jalview 2.0 (48). For phylogenetic analysis, nucleotide sequence alignments were performed with Clustal W and used to construct a Neighbor Joining Tree or Maximum likelihood tree with MEGA, version 5 (49).

Results and Discussion

Biochemical and biophysical characteristics of CAMS, the β class carbonic anhydrase encoded in Msed_0390. The heat-treated cell extract of recombinant *E. coli* containing Msed_0390 exhibited CO_2 hydration activity, relative to the control**.** Furthermore, the IMAC purified enzyme (see **Figure 1**) shows CO_2 hydration activity 292 U/mg at 10^{o}C, validating that Msed_0390 encodes a functional CA for CO_2 and HCO_3^- interconversion in *M. sedula*. This enzyme was labeled CAMS (CA from *M. sedula*), and it is the first CA that has been identified from thermoacidophilic archaea. Esterase activity, as determined by hydrolysis of p-nitrophenyl acetate, is a feature of α-CA and δ-CA (15). Similar to CAB, the β-CA from *M.*

thermoautotrophicum, and CAM, a γ-CA from *M. thermophila*, CAMS does not exhibit esterase activity. (See **Table 2** for comparison of typical CA from different classes).

The predicted molecular mass of CAMS is 22 kDa (including a C-terminal 6×His tag), predicted via ExPASy (http://web.expasy.org/compute_pi/). The molecular mass determined by size exclusion chromatagraphy was 20.6 kDa, suggesting this enzyme functions as a monomer. Thus, the molecular assembly of CAMS differs from other β-CAs that typically exist as homodimers, homotetramers and homooctamers, with the fundamental structural unit as a dimer (50). Though CAMS resembles α-CAs that typically occur as monomers, the absence of esterase activity distinguishes it from α-CAs. High amounts of Zn (18 mg/L) were detected in the IMAC purified CAMS (4.0 mg/mL, based on BCA method (51)), compared to zero or trace amount zinc in the protein storage buffer, suggesting this enzyme contains Zn as active site binding metal.

Sequence alignment with previously identified β-CAs. Analysis using Target P1.1 (http://www.cbs.dtu.dk/services/TargetP/) and Signal P 4.0 (http://www.cbs.dtu.dk/services/SignalP/) did not identify any obvious signal sequence or transit peptide involved in targeting of Msed_0390 to a specific cellular location. However, it was predicted to be a trans-membrane protein by PSIPRED (52) (see **Figure 2**). The amino acid sequence for Msed_0390 (CAMS) was used to search homologs by Blastp in the non-redundant sequence database at NCBI. 19 ORFs, mostly derived from other species in the Sulfolobales, were identified with top high sequence similarity to CAMS (> 90% identity) and low Expect value (< 1e-20), indicating a high conservation of this enzyme within this order. The genomes of these species encode genes belonging to the 3-hydroxypropionate/4-

hydroxybutyrate (3HP/4HB) cycle. Neither Msed_0390 nor any of the homologs identified in the Blastp search has been confirmed biochemically as a functional CA. However, some are related to "cab" type β-CAs, such as the hypothetical β-CAs from Methanococcus, Methanobrevibacter, Mycobacterium and Bacillus species. A second Blastp search using these CA candidates as query identified the previously documented CAB from *M. thermoautotrophicum* (19). Phylogenetic analysis by Smith *et al*. (21) grouped the above species into an exclusively prokaryotic clade (D) of β-CAs, distantly related to the other clades, consisting primarily of sequences from Gram-positive bacteria. Furthermore, a multiple sequence alignment was performed here among CAMS and cab type, plant type, and carboxysome type CAs. Like all the β-CAs, CAMS has the conserved motif for active site zinc binding ligands (Cys7, His56, and Cys59), and an aspartate / arginine pair (Asp 9, Arg11), (numbering as in the CAMS sequence), which represents a typical active site architecture of β-CAs. Interestingly, the sequence alignment among CAMS and its homologs from other members of Sulfolobales indicates the absence of Cys7 (numbering as in the CAMS sequence) in some organisms, such as *Metallosphaera cuprina*, *Sulfolobus acidocaldarius*, and *Thermoproteus tena* (see **Figure 3-4**). Though those ORFs have not been biochemically confirmed as functional CAs, a different catalytic mechanism for Sulfolobales CAs that is different from 'cab' type β-CAs might be expected, since that Cys is probably not required for metal binding and coordination in CAMS group enzymes.

Potential applications of CAMS in carbon dioxide capture and sequestration. Although clean energy technologies have been developed and applied, fully half of the electricity in the United States stills relies on the combustion of cheap and abundant coal, which results in 40% of fossil fuel derived CO_2 emissions worldwide (53). Thus, CO_2 capture and sequestration technology can reduce CO_2 emissions significantly. Recently the prospects of carbonic anhydrases (CA) have been demonstrated on acceleration of CO_2 capture from fossil fuels combustion products (53,54), since these enzymes catalyze the reaction:

$$CO_2 + H_2O \leftrightarrow H_2CO_3 \leftrightarrow HCO_3^- + H^+$$

To drive the equilibrium of this reaction towards the right, either an alkaline capture solvent approach (e.g., aqueous monoethanolamine (MEA), potassium carbonate or methyldiethanolamine (MDEA)) should be used to neutralize the proton released, or a biomineralization approach should be used in which CO_2 is precipitated as solid carbonate via divalent calcium or magnesium (53). CAs can facilitate both of the approaches for carbon capture and storage. Especially for the former case, CO_2 is removed from the flue gas stream in the absorber column and then desorbed in a heated stripper column to give relatively pure CO_2 for compression and storage. Solvents such as MEA can bind CO_2 tightly at the first step but this creates large electricity costs at desorption. Therefore, the combined use of CA and aqueous solvents with slower absorption rate of CO_2 (e.g. hindered and tertiary amines) tends to maintain fast and adequate absorption, while decreasing the energy cost of desorption (55). Early work using readily available CAs from mammalian sources exhibits poor stability and activity under the harsh conditions of these processes: temperatures from 50 to over 125°C, high concentrations of organic amine, heavy metals, and sulfur and nitrogen oxides. Hence, CAs from thermophilic organisms subsequently started to draw attention as an enzyme source in solvent-based capture system (53,56). Two CAs have been identified and biochemically characterized from thermophilic organisms: the γ-class CA (CAM) from *M. thermophila* (5), and the β-class CA (CAB) from *M. thermoautotrophicum* (19). Interestingly, though the structure of a γ-class CA from the hyperthermophile *P. horikoshii* has been resolved, no biochemical data have been reported for this enzyme (29). Actually, other thermostable CAs may have been discovered from organisms and applied to carbon sequestration systems. US 2004/0259231 (57) disclosed the use of CAB, as well as the non-thermostable human CA isoform IV, in a CO_2 solubilization and concentration process. Borchert *et al.* (58) described heat-stable α-class CA from *Bacillus clausii* for >99% extraction of CO_2 from gas stream, versus only 33% extraction without CA. CA from the thermophile *Caminibacter mediatlanticus* can also increase the amount of CO_2 extracted with solvents (59). Multiple companies have developed their own carbon capture technology based on systems using CA as catalysts in which they are immobilized on solid support such as particle packed in a bed reactor (Solution Inc. (60)) or a fiber microporous membrane in a "permeator" (Carbozyme Inc. (61)). The immobilization of CA can ensure reusability by increasing their thermal, chemical stabilities and biocatalytic activity (30).

It was pointed out that the ideal CA for use in sequestration systems should have catalytic activity in capture solvents at absorption temperatures (45–60°C) and survive the desorption temperatures (>100°C) (53). However, so far, wild-type thermophilic CAs are still unsatisfactory for a practical CA-catalyzed CO_2 sequestration system. Already known thermostable β-class and γ-class CAs show slower rates than mesophilic α-class CAs (53). Screening of CA should be expanded in archaea (thermophilic or hyperthermophilic sources), and all microbes evolved under harsh conditions (found in vent or high-saline environments)

(30). Based on these CA candidates, protein engineering methods can be employed to obtain CA variants harboring higher activity, thermostability, pH optimum, and toxic impurities/gas tolerance (30). For example, Codexis has acquired a CA from a mesophilic organism with higher activity and stability in MDEA and other amine solvents from high-through-put screening. After directed evolution, this enzyme had its half-life ($t_{1/2}$) increased by five orders of magnitude to 20 hours in 4.2 M MDEA at 75°C (62).

In this work, the carbonic anhydrase discovered from the thermophilic archaeon *M. sedula* is expected to show comparable or even higher thermostability than CAB, since the latter's host, *Methanobacterium thermoautotrophicum*, has an optimal growth temperature of 65°C, which is lower than that of *M. sedula* (70-75°C). Therefore, further study on the thermostability and optimal temperature of CAMS is needed to demonstrate the prospects of this enzyme's application in carbon capture and sequestration technology.

CAMS's physiological role in carbon capture and carbon concentration mechanisms in M. sedula, and its relation to 3HP/4HB cycle. An already proposed model of CA roles in photosynthetic inorganic carbon assimilation by cyanobacteria from soda lakes can shed light on the physiological role and relation of CAMS in CCM and carbon fixation in *M. sedula*. This model was established in an alkaliphilic cyanobacterium *Microcoleus chthonoplastes*, which lives at high alkaline pH (pH 9-10) where inorganic carbon exists only in the carbonate and bicarbonate forms (63). Also, their habitat may have low rates of inorganic carbon diffusion in water, and their thick exopolysaccharide (EPS) sheath may prevent penetration of inorganic carbon into the cells. Therefore, a CCM is probably required and several CAs were considered to be involved. An extracellular CA, CahB1, was assumed to participate in the uptake of exogenous Ci and pH stabilization in the pericellular layers of *M. chthonoplastes* (64). On the other hand, carboxysomal CAs convert accumulated bicarbonate into CO_2 as the acceptable substrate for Rubisco. The extracellular CA can also prevent the leakage of CO_2 from cells by rapid conversion to bicarbonate in periplasmic space and in pericellular layer, so bicarbonate can be transported through the membrane back into the cell by carbon transporters (65).

The case is similar, but opposite, in *M. sedula* growing at low pH (pH 2) where CO_2 has low solubility, but the cytoplasm, where 3HP/4HB cycle utilizes bicarbonate as substrate, is neutral (66). It is still unclear whether CO_2 is transported into cell via CO_2 uptake systems, then converted to bicarbonate inside the cell, or CO_2 is converted to bicarbonate before delivered by HCO_3^- transporters into cell. However, an extracellular or transmembrane CA would be expected for rapid conversion of external CO_2 to bicarbonate and facilitation of its transport into cell. So far, there is no strong evidence to show CAMS as an extracellular protein, but it has a hydrophobic region which makes it a possible transmembrane protein as predicted by PSIPRED. Cellular location of CAMS should be confirmed by experiments, such as Western blots. On the other hand, it would be worthwhile to measure CAMS activity, as a function of pH. Intracellular CAs always function effectively at neutral or acidic pH (64), while the CahB1, as mentioned above, exhibits highest enzymatic activity at alkaline pH, which is consistent with its extracellular localization (65).

Another assumption is the potential spatial relationship between CAMS and 3-HP/4-HB cycle enzymes, especially the key carbon fixing enzyme carboxylase (Msed_0147,0148,1375). As mentioned previously, the arrangement of the CA associated with the carboxysome can facilitate the delivery of CO_2, the substrate for RuBisCo. It would not be surprising if *M. sedula*, harboring the 3HP/4HB cycle, has developed convergent sub-cellular compartment in parallel for tight spatial connection between substrate provider (CA) and substrate consumer (carboxylase). Actually, it was proposed more than 25 years ago that many metabolic pathways are organized as multi-enzyme complexes in highly ordered assemblies. Evidence of this is the synthesis of phenylpropanoids and flavonoids on membrane associated enzyme complexes (67). Perhaps, similar multi-enzyme complexes can be observed for 3-HP/4-HB cycle enzymes and their accessory enzymes, such as CA. As Winkel-Shirley described (68), enzyme complexes can facilitate the direct transfer or channeling of intermediates between active sites. Such organization is compelling for a number of reasons: there could be competition for substrates at the numerous branch points within these pathways, the intermediates are highly reactive and potentially toxic, and the overall concentrations of these compounds are extremely low. Moreover, only in this way can these pathways respond quickly to external and internal signals to adjust the levels and/or types of end products that are synthesized.

In any case, more work is needed to connect the biochemical and biophysical features of CAMS to *M. sedula* physiology. This kind of information is critical to strategic use of the 3-HP/4-HB cycle in metabolic engineering hosts for carbon fixation into fuels and chemicals.

REFERENCES

1. **Meldrum NU, Roughton FJ.** 1933. Carbonic anhydrase. Its preparation and properties. J. Physiol. **80:**113-142.

2. **Syrjänen L, Tolvanen M, Hilvo M, Olatubosun A, Innocenti A, Scozzafava A, Leppiniemi J, Niederhauser B, Hytönen VP, Gorr TA, Parkkila S, Supuran C T.** 2010. Characterization of the first beta-class carbonic anhydrase from an arthropod (*Drosophila melanogaster*) and phylogenetic analysis of beta-class carbonic anhydrases in invertebrates. BMC Biochem. **11:**28.

3. **Zimmerman SA, Ferry JG, Supuran CT.** 2007. Inhibition of the archaeal beta-class (Cab) and gamma-class (Cam) carbonic anhydrases. Curr. Top. Med. Chem. **7:**901-908.

4. **Smith KS, Cosper NJ, Stalhandske C, Scott RA, Ferry JG.** 2000. Structural and kinetic characterization of an archaeal β-Class carbonic anhydrase. J Bacteriol. **182:** 6605-13.

5. **Strop P, Smith KS, Iverson TM, Ferry JG, Rees DC.** 2001. Crystal structure of the "cab"-type beta class carbonic anhydrase from the archaeon *Methanobacterium thermoautotrophicum*. J. Biol. Chem. **276:**10299-305.

6. **Alber BE, Colangelo CM, Dong J, Stålhandske CM, Baird TT, Tu C, Fierke CA, Silverman DN, Scott RA, Ferry JG.** 1999. Kinetic and spectroscopic characterization of the gamma-carbonic anhydrase from the methanoarchaeon *Methanosarcina thermophila*. Biochemistry **38:**13119-28.

7. **Johansson IM, Forsman C.** 1993. Kinetic studies of pea carbonic anhydrase. Eur. J. Biochem. **218:**439-46.

8. **Lindskog S.** 1997. Structure and mechanism of carbonic anhydrase. Pharmacol. Ther. **74:**1-20.

9. **Silverman DN, Lindskog S.** 1988. The catalytic mechanism of carbonic anhydrase: implications of a rate-limiting protolysis of water. Acc. Chem. Res. **21:**30-36.

10. **Supuran CT.** 2008. Carbonic anhydrases - an overview. Curr. Pharm. Des. **14:** 603-614.

11. **Alber B, Ferry JG.** 1994. A carbonic anhydrase from the archaeon *Methanosarcina thermophila*. Proc. Natl. Acad. Sci. U S A. **91:**6909-13.

12. **Roberts SB, Lane TW, Morel FMM.** 1997. Carbonic anhydrase in the marine diatom *Thalassiosira weissflogii* (Bacillariophyceae). J. Phycol. **33:**845-850.

13. **Lane TW, Saito MA, George GN, Pickering IJ, Roger C, Morel FMM.** 2005. Isolation and preliminary characteization of a cadmium carbonic anhydrase from a marine diatom. Nature **435:**42.

14. **Yee D, Morel FMM.** 1996. *In vivo* substitution of zinc by cobalt in carbonic anhydrase of a marine diatom. Limnol. Oceanogr. **41:**e577.

15. **Lee RBY, Smith JAC, Rickaby REM.** 2013. Cloning, expression and characterization of the δ-carbonic anhydrase of *Thalassiosira weissflogii* (Bacillariophyceae). J. Phycol. **49:**170-177.

16. **Cox EH, McLendon GL, Morel FMM, Lane TW, Prince RC, Pickering IJ, George GN.** 2000. The active site structure of *Thalassiosira weissflogii* carbonic anhydrase 1. Biochemistry **39:**12128-30.

17. **Xu Y, Feng L, Jeffrey PD, Shi Y, Morel FMM.** 2008. Structure and metal exchange in the cadmium carbonic anhydrase of marine diatoms. Nature. **452:**56-61.

18. **Kimber MS, Pai EF.** 2000. The active site architecture of *Pisum sativum* beta-carbonic anhydrase is a mirror image of that of alpha-carbonic anhydrases. EMBO J. **19:**1407-18.

19. **Smith KS, Ferry JG.** 1999. A plant-type (beta-class) carbonic anhydrase in the thermophilic methanoarchaeon *Methanobacterium thermoautotrophicum*. J. Bacteriol. **181:**6247-53.

20. **So AK, Espie GS, Williams EB, Shively JM, Heinhorst S, Cannon GC.** 2004. A novel evolutionary lineage of carbonic anhydrase (epsilon class) is a component of the carboxysome shell. J. Bacteriol. **186:**623-30.

21. **Smith KS, Jakubzick C, Whittam TS, Ferry JG.** 1999. Carbonic anhydrase is an ancient enzyme widespread in prokaryotes. Proc. Natl. Acad. Sci. U. S. A. **96:**15184-9.

22. **Sawaya MR, Cannon GC, Heinhorst S, Tanaka S, Williams EB, Yeates TO, Kerfeld CA.** 2006. The structure of beta-carbonic anhydrase from the carboxysomal shell reveals a distinct subclass with one active site for the price of two. J. Biol. Chem. **281:**7546-55.

23. **Macauley SR, Zimmerman A, Apolinario EE, Evilia C, Hou YM, Ferry JG,**

Sowers KR. 2009. The archetype gamma-class carbonic anhydrase (Cam) contains iron when synthesized *in vivo*. Biochemistry **48:**817-9.

24. **Ferry JG.** 1996. Characterization of heterologously produced carbonic anhydrase from *Methanosarcina thermophila*. J. Bacteriol. **178:**3270-74.

25. **Ferry JG.** 2010. The gamma class of carbonic anhydrases. Biochim. Biophys. Acta. **1804:**374-81.

26. **Smith JD, Robinson AS.** 2002. Overexpression of an archaeal protein in yeast: secretion bottleneck at the ER. Biotechnol. Bioeng. **79:**713-23.

27. **Kim R, Sandler SJ, Goldman S, Yokota H, Clark J, Kim S.** 1998. Overexpression of archaeal proteins in *Escherichia coli*. Biotechnology Letters **20:**207-10.

28. **Kisker C, Schindelin H, Alber BE, Ferry JG, Rees DC.** 1996. A left-handed beta-helix revealed by the crystal structure of a carbonic anhydrase from the archaeon *Methanosarcina thermophila*. EMBO J. **15:** 2323-2330.

29. **Jeyakanthan J, Rangarajan S, Mridula P, Kanaujia SP, Shiro Y, Kuramitsu S, Yokoyama S, Sekar K.** 2008. Observation of a calcium-binding site in the gamma-class carbonic anhydrase from *Pyrococcus horikoshii*. Acta Crystallogr. D. Biol. Crystallogr. **64:**1012-9.

30. **Kanth BK, Lee J, Pack SP.** 2013. Carbonic anhydrase: Its biocatalytic mechanisms and functional properties for efficient CO_2 capture process development. Eng. Life Sci. **13:**422-431.

31. **Smith KS, Ferry JG.** 2000. Prokaryotic carbonic anhydrases. FEMS Microbiol. Rev. **24:**335-366.

32. **Badger MR, Price GD.** 1994. The role of carbonic anhydrase in photosynthesis. Annu. Rev. Plant Physiol. Plant Mol. Biol. **45:**369-392.

33. **Hatch MD.** 1987. C4 photosynthesis: a unique blend of modified biochemistry, anatomy and ultrastructure. Biochim.Biophys.Acta. **895:**81-106.

34. **Hatch MD, Burnell J.** 1990. Carbonic anhydrase activity in leaves and its role in the first step of C4 photosynthesis. Plant Physiol. **93:**380-83.

35. **Cannon GC, Heinhorst S, Kerfeld CA.** 2010. Carboxysomal carbonic anhydrases: Structure and role in microbial CO_2 fixation. Biochim. Biophys. Acta. **1804:**382-92.

36. **Baker SH, Williams DS, Aldrich HC, Gambrell AC, Shively JM.** 2000. Identification and localization of the carboxysome peptide CsoS3 and its corresponding gene in *Thiobacillus neapolitanus*. Arch. Microbiol. **173:**278-83.

37. **Fukuzawa H, Suzuki E, Komukai Y, Miyachi S.** 1992. A gene homologous to chloroplast carbonic anhydrase (icfA) is essential to photosynthetic carbon dioxide fixation by Synechococcus PCC7942. Proc. Nati. Acad. Sci. USA. **89:**4437-41.

38. **Jian WY, Price GD, Song L, Badger MR.** 1992. Isolation of a putative carboxysomal carbonic anhydrase gene from the *cyanobacterium Synechococcus* PCC7942. Plant Physiol. **100:**794–800.

39. **Price GD, Coleman JR, Badger MR.** 1992. Association of carbonic anhydrase activity with carboxysomes isolated from the *cyanobacterium Synechococcus* PCC7942. Plant Physiol. **100:**784-793.

40. **Majeau N, Coleman JR.** 1996. Effect of CO_2 concentration on carbonic anhydrase and ribulose-1, 5-bisphosphate carboxylase/oxygenase expression in pea. Plant Physiol. **112:** 569-574.

41. **Auernik KS, Kelly RM.** 2009. Physiological versatility of the extremely thermoacidophilic archaeon *Metallosphaera sedula* supported by transcriptomic analysis of heterotrophic, autotrophic, and mixotrophic growth. Appl. Environ. Microbiol. **76:** 931-935.

42. **Laemmli UK.** 1970. Cleavage of structural proteins during assembly of the head of bacteriophage-T4. Nature **227:**680-685.

43. **Zehr BD, Savin TJ, Hall RE.** 1989. A one-step, low background Coomassie staining procedure for polyacrylamide gels. Anal. Biochem. **182:**157-159.

44. **Wilbur KM, Anderson NG**. 1948. Electrometric and colorimetric determination of carbonic anhydrase. J. Biol. Chem. **176:**147-154.

45. **Bradford MM.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem.**72:**248-254.

46. **Armstrong J, Myers DV, Verporte JA, Edsall JT.** 1966. Purification and properties of human erythrocyte carbonic anhydrase. J. Biol. Chem. **41:** 5137-49.

47. **Larkin MA, Blackshields G, Brown NP, R. Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson**

TJ, Higgins DG. 2007. Clustal W and Clustal X version 2.0. Bioinformatics **23:**2947-2948.

48. **Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ.** 2009. Jalview Version 2-a multiple sequence alignment editor and analysis workbench. Bioinformatics **25:** 1189-1191.

49. **Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S.** 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. **28:**2731-9.

50. **Rowlett RS.** 2010. Structure and catalytic mechanism of the beta-carbonic anhydrases. Biochim. Biophys. Acta. **1804:**362-73.

51. **Brown R.** 1989. Protein measurement using bicinchoninic acid: elimination of interfering substances. Anal. Biochem. **180:**136-9.

52. **Buchan DWA, Minneci F, Nugent TCO, Bryson K, Jones DT.** 2013. Scalable web services for the PSIPRED protein analysis workbench. Nucleic Acids Res. **41:** 340-348.

53. **Savile CK, Lalonde JJ.** 2011. Biotechnology for the acceleration of carbon dioxide capture and sequestration. Curr. Opin. Biotechnol. **22:**818-23.

54. **Chu S.** 2009. Carbon capture and sequestration. Science **325:**1599.

55. **Ciferno J, Fout T, Jones A, Murphy JT.** 2009. Capturing carbon from existing coal-fired plants. Chem. Eng. Prog. **105:**33-41.

56. **Daigle R, Desrochers M.** 2009. Carbonic anhydrase having increased stability under high temperature conditions. Patent US7521217.

57. **Bhattacharya S.** 2004. Enzyme facilitated solubilization of carbon dioxide from emission streams in novel attachable reactors/devices. Patent US20040259231A1.

58. **Borchert M, Saunders P.** 2010. Heat-stable carbonic anhydrases and their use. Patent US7803575.

59. **Borchert M, Saunders P.** 2010. Heat-stable carbonic anhydrases and their use. Patent WO2010151787.

60. **Lalande JM, Tremblay A.** 2005. Process and a plant for the production of Portland cement clinker. Patent US6908507 B2.

61. **Bao L, Trachtenberg M.** 2006. Facilitated transport of CO_2 across a liquid membrane: Comparing enzyme, amine and alkaline. J. Memb. Sci. **280:**330-334.

62. **Savile C, Nguyen L, Alvizo O, Balatskaya S, Choi G, Benoit M, Fusman I, Geilhufe J, Ghose S, Gitin S, Grate J, City R, Broering J.** 2011. Low-cost biocatalyst for acceleration of energy efficient CO_2 capture. From ARPA-E Workshop.

63. **Zavarzin GA, Zhilina TN, Kevbrin VV.** 1999. Alkaliphilic microbial community and its functional diversity, Microbiology. (Moscow) **68:**503-521.

64. **Price GD, Badger MR, Woodger FJ, Long BM.** 2008. Advances in understanding the cyanobacterial CO_2-concentrating-mechanism (CCM): functional components, Ci transporters, diversity, genetic regulation and prospects for engineering into plants. J. Exp. Bot. **59:**1441-61.

65. **Kupriyanova EV, Sinetova MA, Markelova AG, Allakhverdiev SI, Los DA, Pronina NA.** 2011. Extracellular β-class carbonic anhydrase of the alkaliphilic cyanobacterium *Microcoleus chthonoplastes*. J. Photochem. Photobiol. B. **103:**78-86.

66. **Peeples TL, Kelly RM.** 1995. Bioenergetic response of the extreme thermoacidophile *Metallosphaera sedula* to thermal and nutritional stresses. Appl. Environ. Microbiol. **61:**2314-21.

67. **Hrazdina G, Wagner GJ.** 1985. Metabolic pathways as enzyme complexes: Evidence for the synthesis of phenylpropanoids and flavonoids on membrane associated enzyme complexes. Arch. Biochem. Biophys. **237:**88-100.

68. **Winkel-Shirley B.** 2001. Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell Biology, and biotechnology. Plant Physiol. **126:**485-493.

69. **Sly WS, Hu PY.** 1995. Human carbonic anhydrases and carbonic anhydrase deficiencies. Annu. Rev. Biochem. **64:**375-401.

70. **Viparellia F, Montia SM, Simonea GD, Innocentib A, Scozzafavab A, Xu Y, Morel FMM, Supuran CT.** 2010. Inhibition of the R1 fragment of the cadmium-containing ζ-class carbonic anhydrase from the diatom *Thalassiosira weissflogii* with anions. Bioorg. Med. Chem. Lett. **20:**4745-4748.

TABLES

Table 1. Current classification of carbonic anhydrases.

Class	Active site metal binding	Metal ligand	Structural composition	Distribution	Subclass
α	Zinc	3 × histidine	monomer	Mammals, a few bacteria, algae	NA
β	Zinc	2 × cysteine 1 × histidine	2 × N mer	Plants, Algae, Bacteria, and Archaea	plant, cab, carboxysome
γ	iron (*in vivo*), zinc/cobalt (*in vitro*), calcium	3 × histidine	trimer	Eukarya, Bacteria and Archaea	cam, camH
δ	zinc/cobalt	3 × histidine	monomer	Diatom, marine phytoplankton	NA
ζ	Zinc/cadmium	2 × cysteine 1 × histidine	monomer	Marine diatom only	NA

Table 2. Comparison of typical carbonic anhydrases from different organisms.

Organism	Denotation	Class	Optimal temperature for organism growth /°C	Optimal temperature for activity/°C	Inactivated temperature /°C	CO_2 hydration activity U/mg (W-A)	k_{cat} /s	Esterase activity	Ref
Human	NA	α	37	NA	NA	~3000	10^6	Y	(69)
Methanosarcina thermophila	Cam	γ	50	55	75	720	10^5	N	(6)
Methanobacterium thermoautotrophicum	Cab	β	65-75	70	90	624	10^4	N	(19)
Pyrococcus horikoshii	CamH	γ	98	NA	NA	NA	NA	NA	(29)
Metallosphaera sedula	Cams	β	70-75	NA	NA	292	NA	N	This study
Thalassiosira welssflogii	TWCA1	δ	20	NA	boiled	425	NA	Y	(15)
	CDCA1	ζ		NA	NA	NA	10^6	NA	(70)

Note: W-A means activity was measured using the Wilbur–Anderson method.

FIGURES

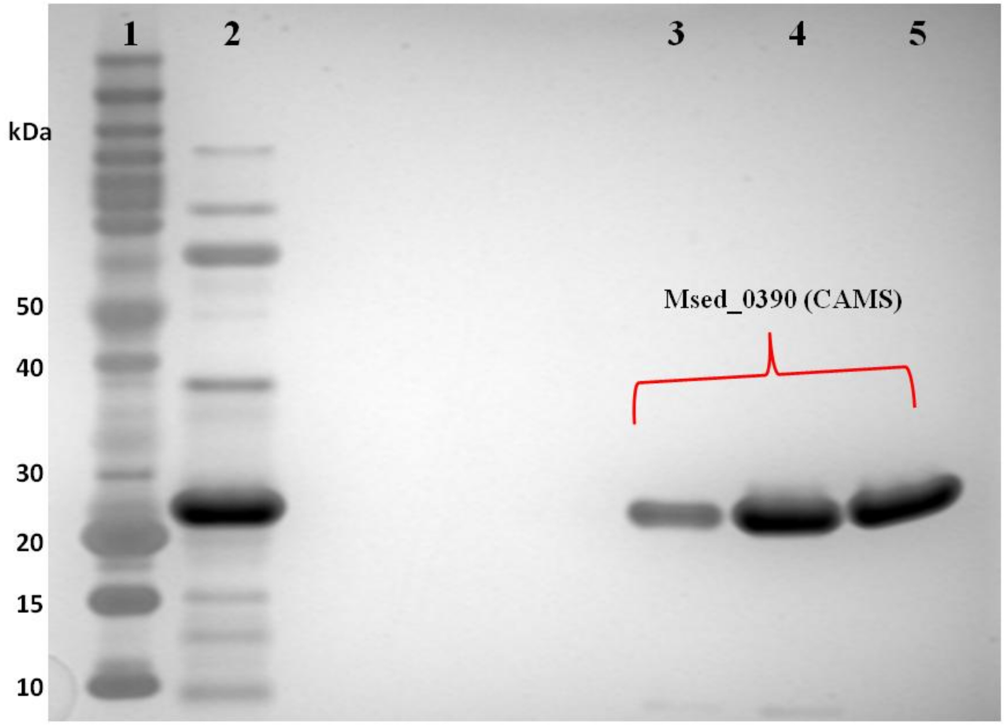


Figure 1. **Purification of recombinant CAMS (Msed_0390).** Lanes of SDS-PAGE: #1. Bench marker. #2. Heat-treated recombinant *E. coli* cell extract containing CAMS. #3, 4, 5. Purified CAMS fractions from IMAC.

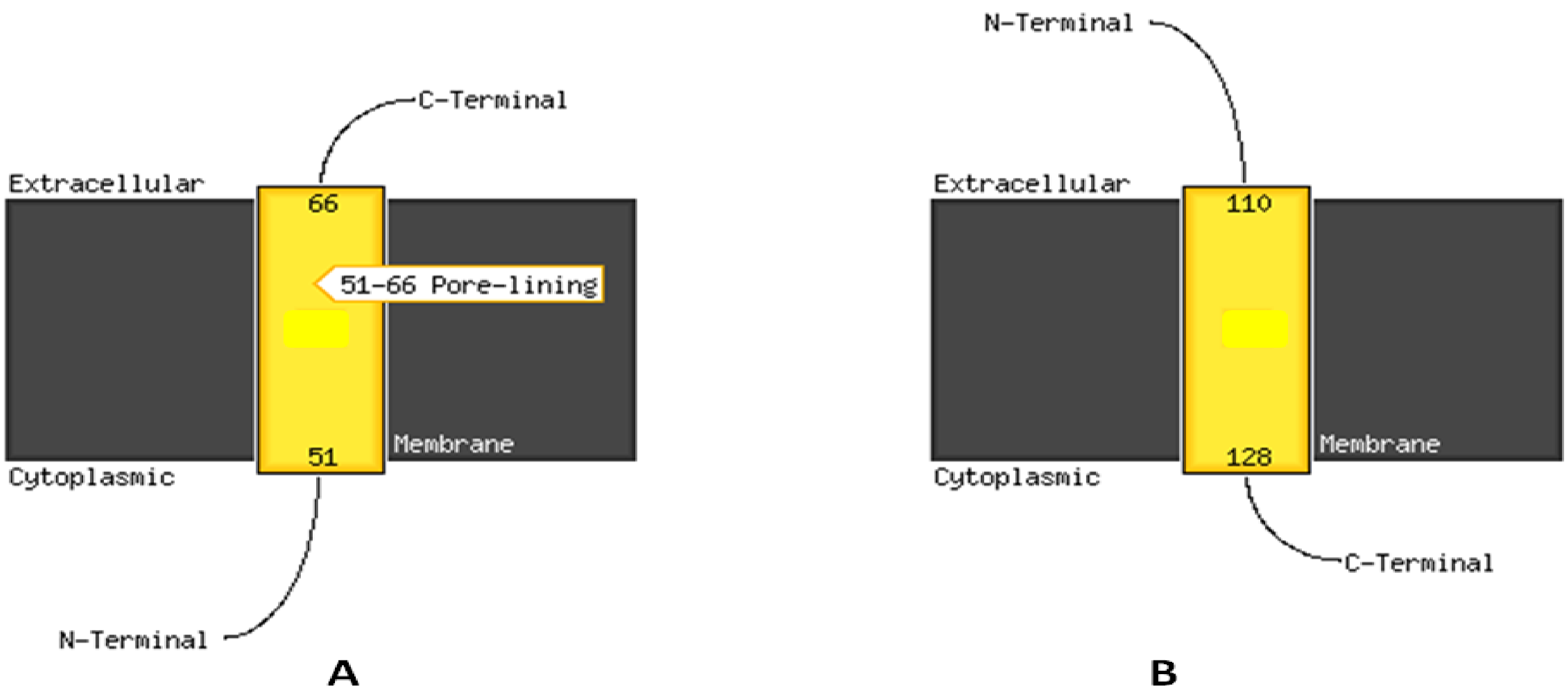
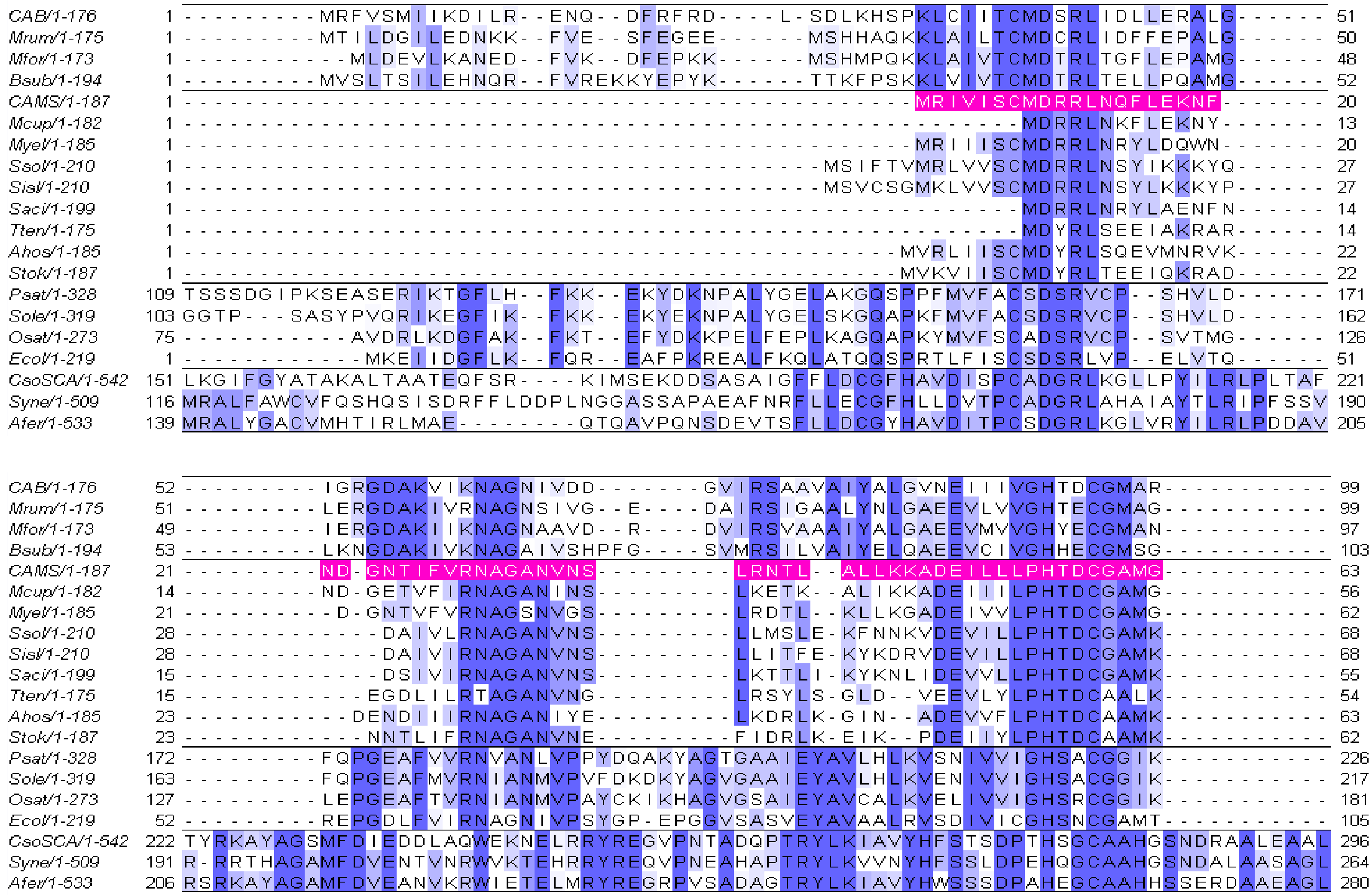
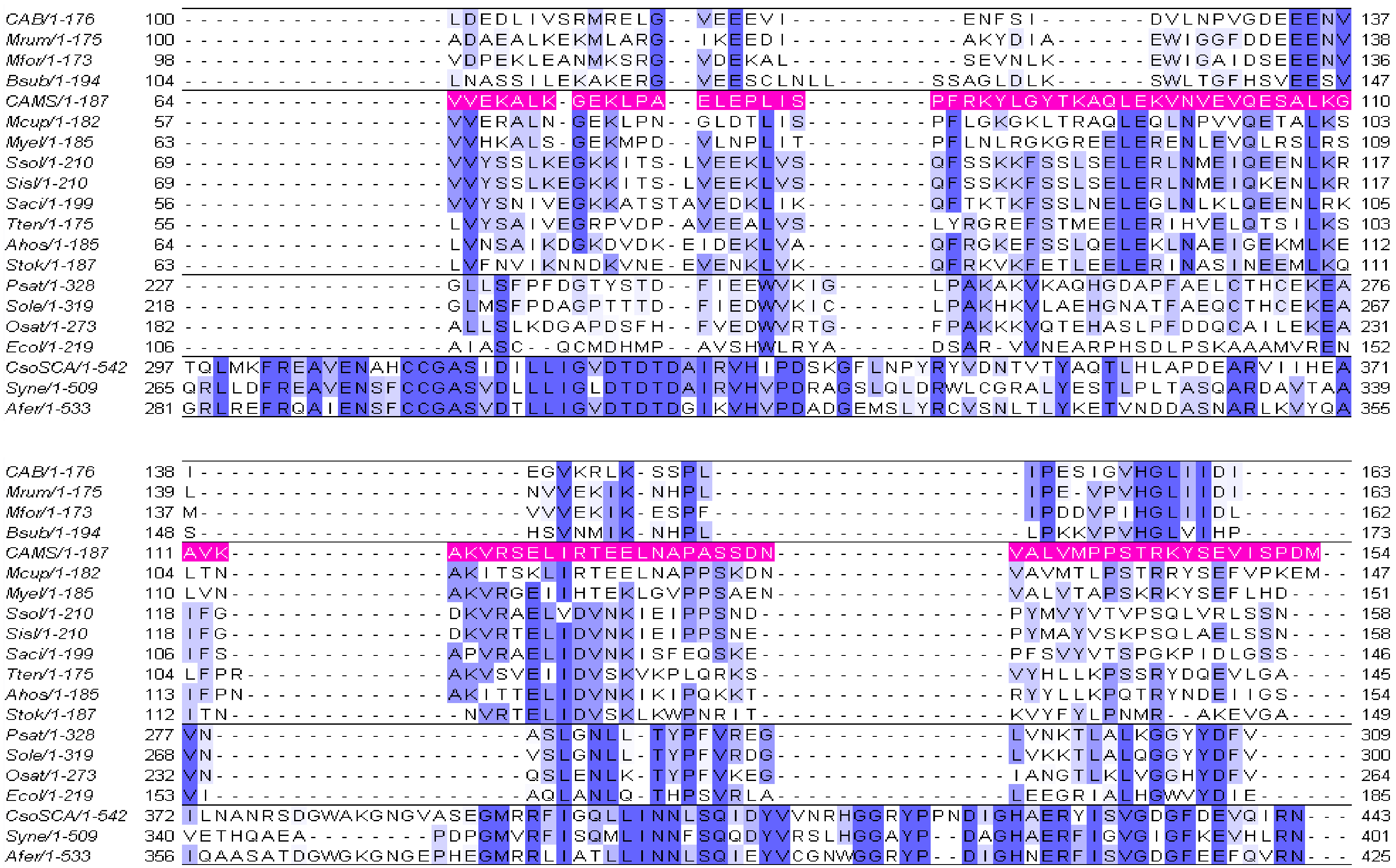
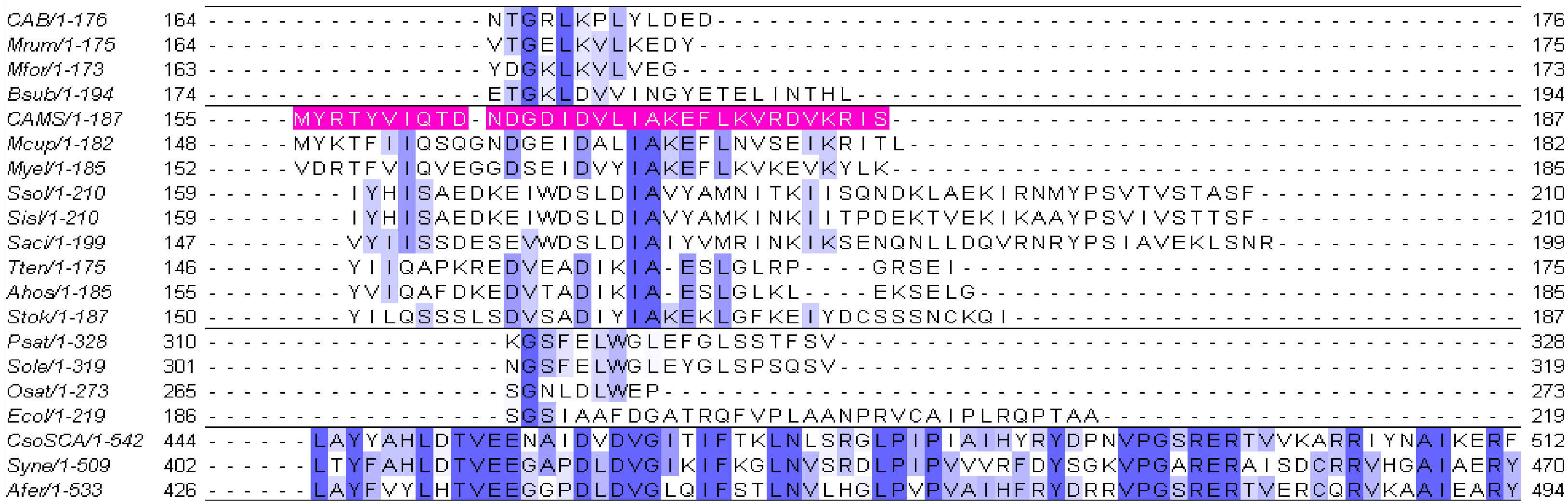


Figure 2. Prediction of transmembrane region of CAMS by PSIPRED (52)**, based on two different model.** (A) Schematic diagram of the MEMSATSVM. (B) Schematic diagram of MEMSAT3.

Figure 3. Multiple sequence alignment of CAMS with documented or putative carbonic anhydrases of cab, plant and carboxysome type. Sequences were aligned using Clustal W X version 2.0 (47)**, and displayed with Jalview 2.0** (48)**. Type 1(cab): CAB,** *Methanobacterium thermoautotrophicum,* **Mrum,** *Methanobrevibacter ruminantium,* **Mfor,** *Methanobacterium formicicum,* **Bsub,** *Bacillus subtilis***. Type 2: CAMS (cams),** *Metallosphaera sedula,* **Mcup,** *Metallosphaera cuprina,* **Myel,** *Metallosphaera yellowstonensis,* **Ssol,** *Sulfolobus solfataricus,* **Sisl,** *Sulfolobus islandicus,* **Saci,** *Sulfolobus acidocaldarius,* **Tten,** *Thermoproteus tenax,* **Ahos,** *Acidianus hospitalis,* **Stok,** *Sulfolobus tokodaii.* **Type 3(plant): Psat,** *Pisum sativum,* **Sole,** *Spinacia oleracea,* **Osat,** *Oryza sativa,* **Ecol,** *Escherichia coli.* **Type 4 (carboxysome): CsoSCA,** *Halothiobacillus Neapolitanus,* **Syne,** *Synechococcus sp.,* **Afer,** *Acidithiobacillus ferrooxidans.*


CAB/1-176 1 ----------MRFVSMIIKDILR--ENQ--DFRFRD----L-SDLKHSPKLCIITCMDSRLIDLLERALG------ 51
Mrum/1-175 1 ----------MTILDGILEDNKK--FVE--SFEGEE------MSHHAQKKLAILTCMDCRLIDFFEPALG------ 50
Mfor/1-173 1 ------------MLDEVLKANED--FVK--DFEPKK------MSHMPQKKLAIVTCMDTRLTGFLEPAMG------ 48
Bsub/1-194 1 ----------MVSLTSILEHNQR--FVREKKYEPYK------TTKFPSKKLVIVTCMDTRLTELLPQAMG------ 52
CAMS/1-187 1 --MRIVISCMDRRLNQFLEKNF-------- 20
Mcup/1-182 1 ---MDRRLNKFLEKNY-------- 13
Myel/1-185 1 --MRIIISCMDRRLNRYLDQWN-------- 20
Ssol/1-210 1 --MSIFTVMRLVVSCMDRRLNSYIKKKYQ------ 27
Sisl/1-210 1 --MSVCSGMKLVVSCMDRRLNSYLKKKYP------ 27
Saci/1-199 1 ---MDRRLNRYLAENFN------ 14
Tten/1-175 1 ---MDYRLSEEIAKRAR------ 14
Ahos/1-185 1 --MVRLIISCMDYRLSQEVMNRVK------ 22
Stok/1-187 1 --MVKVIISCMDYRLTEEIQKRAD------ 22
Psat/1-328 109 TSSSDGIPKSEASERIKTGFLH--FKK--EKYDKNPALYGELAKGQSPPFMVFACSDSRVCP--SHVLD------ 171
Sole/1-319 103 GGTP---SASYPVQRIKEGFIK--FKK--EKYEKNPALYGELSKGQAPKFMVFACSDSRVCP--SHVLD------ 162
Osat/1-273 75 -----------AVDRLKDGFAK--FKT--EFYDKKPELFEPLKAGQAPKYMVFSCADSRVCP--SVTMG------ 126
Ecol/1-219 1 ------------MKEIIDGFLK--FQR--EAFPKREALFKQLATQQSPRTLFISCSDSRLVP--ELVTQ------ 51
CsoSCA/1-542 151 LKGIFGYATAKALTAATEQFSR----KIMSEKDDSASAIGFFLDCGFHAVDISPCADGRLKGLLPYILRLPLTAF 221
Syne/1-509 116 MRALFAWCVFQSHQSISDRFFLDDPLNGGASSAPAEAFNRFLLECGFHLLDVTPCADGRLAHAIAYTLRIPFSSV 190
Afer/1-533 139 MRALYGACVMHTIRLMAE--------QTQAVPQNSDEVTSFLLDCGYHAVDITPCSDGRLKGLVRYILRLPDDAV 205
CAB/1-176 52 ---------IGRGDAKVIKNAGNIVDD-------GVIRSAAVAIYALGVNEIIIVGHTDCGMAR----------- 99
Mrum/1-175 51 ---------LERGDAKIVRNAGNSIVG--E----DAIRSIGAALYNLGAEEVLVVGHTECGMAG----------- 99
Mfor/1-173 49 ---------IERGDAKIIKNAGNAAVD--R----DVIRSVAAAIYALGAEEVMVVGHYECGMAN----------- 97
Bsub/1-194 53 ---------LKNGDAKIVKNAGAIVSHPFG----SVMRSILVAIYELQAEEVCIVGHHECGMSG----------- 103
CAMS/1-187 21 ---------ND-GNTIFVRNAGANVNS---------LRNTL--ALLKKADEILLLPHTDCGAMG----------- 63
Mcup/1-182 14 ---------ND-GETVFIRNAGANINS---------LKETK--ALIKKADEIIILPHTDCGAMG----------- 56
Myel/1-185 21 ----------D-GNTVFVRNAGSNVGS---------LRDTL--KLLKGADEIVVLPHTDCGAMG----------- 62
Ssol/1-210 28 -------------DAIVLRNAGANVNS---------LLMSLE-KFNNKVDEVILLPHTDCGAMK----------- 68
Sisl/1-210 28 -------------DAIVIRNAGANVNS---------LLITFE-KYKDRVDEVILLPHTDCGAMK----------- 68
Saci/1-199 15 -------------DSIVIRNAGANVNS---------LKTTLI-KYKNLIDEVVLLPHTDCGAMK----------- 55
Tten/1-175 15 ------------EGDLILRTAGANVNG---------LRSYLS-GLD--VEEVLYLPHTDCAALK----------- 54
Ahos/1-185 23 -----------DENDIIIRNAGANIYE---------LKDRLK-GIN--ADEVVFLPHTDCAAMK----------- 63
Stok/1-187 23 ------------NNTLIFRNAGANVNE---------FIDRLK-EIK--PDEIIYLPHTDCAAMK----------- 62
Psat/1-328 172 ---------FQPGEAFVVRNVANLVPPYDQAKYAGTGAAIEYAVLHLKVSNIVVIGHSACGGIK----------- 226
Sole/1-319 163 ---------FQPGEAFMVRNIANMVPVFDKDKYAGVGAAIEYAVLHLKVENIVVIGHSACGGIK----------- 217
Osat/1-273 127 ---------LEPGEAFTVRNIANMVPAYCKIKHAGVGSAIEYAVCALKVELIVVIGHSRCGGIK----------- 181
Ecol/1-219 52 ---------REPGDLFVIRNAGNIVPSYGP-EPGGVSASVEYAVAALRVSDIVICGHSNCGAMT----------- 105
CsoSCA/1-542 222 TYRKAYAGSMFDIEDDLAQWEKNELRRYREGVPNTADQPTRYLKIAVYHFSTSDPTHSGCAAHGSNDRAALEAAL 296
Syne/1-509 191 R-RRTHAGAMFDVENTVNRWVKTEHRRYREQVPNEAHAPTRYLKVVNYHFSSLDPEHQGCAAHGSNDALAASAGL 264
Afer/1-533 206 RSRKAYAGAMFDVEANVKRWIETELMRYREGRPVSADAGTRYLKIAVYHWSSSDPAHEGCAAHHSSERDAAEAGL 280




CAB/1-176 164 ----------------NTGRLKPLYLDED-- 176
Mrum/1-175 164 ----------------VTGELKVLKEDY--- 175
Mfor/1-173 163 ----------------YDGKLKVLVEG-- 173
Bsub/1-194 174 ----------------ETGKLDVVINGYETELINTHL-------------------------------------- 194
CAMS/1-187 155 -----MYRTYVIQTD-NDGDIDVLIAKEFLKVRDVKRIS------------------------------------ 187
Mcup/1-182 148 -----MYKTFIIQSQGNDGEIDALIAKEFLNVSEIKRITL----------------------------------- 182
Myel/1-185 152 -----VDRTFVIQVEGGDSEIDVYIAKEFLKVKEVKYLK------------------------------------ 185
Ssol/1-210 159 --------IYHISAEDKEIWDSLDIAVYAMNITKIISQNDKLAEKIRNMYPSVTVSTASF--------------- 210
Sisl/1-210 159 --------IYHISAEDKEIWDSLDIAVYAMKINKIITPDEKTVEKIKAAYPSVIVSTTSF--------------- 210
Saci/1-199 147 --------VYIISSDESEVWDSLDIAIYVMRINKIKSENQNLLDQVRNRYPSIAVEKLSNR-------------- 199
Tten/1-175 146 --------YIIQAPKREDVEADIKIA-ESLGLRP----GRSEI-------------------------------- 175
Ahos/1-185 155 --------YVIQAFDKEDVTADIKIA-ESLGLKL----EKSELG------------------------------- 185
Stok/1-187 150 --------YILQSSSLSDVSADIYIAKEKLGFKEIYDCSSSNCKQI----------------------------- 187
Psat/1-328 310 -----------------KGSFELWGLEFGLSSTFSV--------------------------------------- 328
Sole/1-319 301 -----------------NGSFELWGLEYGLSPSQSV--------------------------------------- 319
Osat/1-273 265 -----------------SGNLDLWEP--- 273
Ecol/1-219 186 -----------------SGSIAAFDGATRQFVPLAANPRVCAIPLRQPTAA------------------------ 219
CsoSCA/1-542 444 ------LAYYAHLDTVEENAIDVDVGITIFTKLNLSRGLPIPIAIHYRYDPNVPGSRERTVVKARRIYNAIKERF 512
Syne/1-509 402 ------LTYFAHLDTVEEGAPDLDVGIKIFKGLNVSRDLPIPVVVRFDYSGKVPGARERAISDCRRVHGAIAERY 470
Afer/1-533 426 ------LAYFVYLHTVEEGGPDLDVGLQIFSTLNVLHGLPVPVAIHFRYDRRVPGSRERTVERCQRVKAAIEARY 494

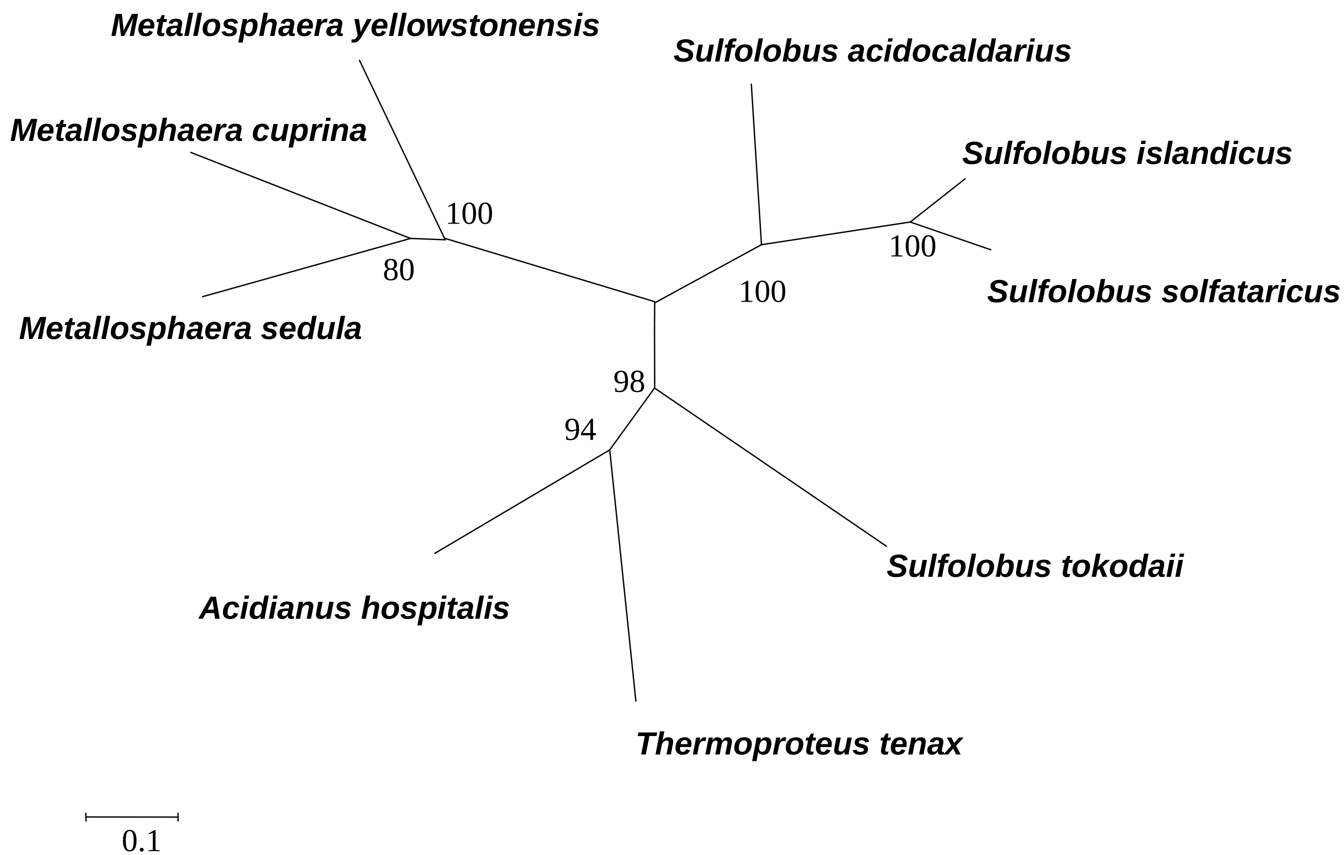


Figure 4. Unrooted phylogenetic tree for CAMS homologues within Sulfolobales. Nucleotide sequence alignments were performed with Clustal W (47) and used to construct a Neighbor Joining Tree with MEGA, version 5 (49). Bootstrap values are displayed at nodes at percentages of 1000 replicates.

Chapter 3

Characterization of Recombinant Acetyl-CoA/propionyl-CoA Carboxylase (E1), Malonyl-CoA reductase (E2), Malonic semialdehyde reductase (E3), and Protein Biotin Ligase (BPL) from the Extremely Thermoacidophilic Archaeon *Metallosphaera sedula* and *in vitro* Conversion of HCO_3^- and Acetyl-CoA into 3-Hydroxypropionate

ABSTRACT

Sub-pathway 1 (SP1) of the 3-Hydroxypropionate/4-Hydroxybutyrate (3HP/4HB) cycle from the extreme thermoacidophile *Metallosphaera sedula* is composed of the heteromultimeric acetyl-CoA/propionyl-CoA carboxylase (E1) (encoded by Msed_0147, 0148, 1375), malonyl-CoA reductase (E2) (encoded by Msed_0709), and malonic semialdehyde reductase (E3) (encoded by Msed_1993). These three enzymes sequentially convert HCO_3^- and acetyl-CoA into 3-hydroxypropionate (3-HP). All three enzymes were produced in recombinant form to demonstrate SP1 function *in vitro* and to explore ways in which the sub-pathway could be optimized *in vivo*. E2 and E3 were heterologously expressed in *Escherichia coli*, purified to homogeneity and biochemically characterized. However, co-expression of the three subunits of E1 in *E. coli* led to extremely low expression levels of soluble α unit (biotin carboxylase), as well as failure by the recombinant host to biotinylate the β sub-unit (biotin carboxyl carrier protein, BCCP) that is required for E1 function. Therefore, identification and recruitment of a biotin protein ligase (BPL), an accessory enzyme needed to post-translationally modify the β sub-unit of E1, was pursued. Previous work by other groups addressed the substrate specificity of BPLs acting on non-cognate BCCPs from different organisms. Of significance here, the BPL from *E. coli* failed to biotinylate BCCP from *Sulfolobus tokodaii*, a closely related species to *M. sedula*. Substrate specificity appeared to relate to a characteristic local sequence motif surrounding the canonical lysine of BCCP (-E-X-M-K-M-). Based on this assumption, this sequence motif was used to examine BCCPs from *E. coli*, *M. sedula*, *S. tokodaii* and *Pyrococcus furiosus*, and indicated that the native

BPL of *P. furiosus* was probably ineffective in biotinylating E1 β sub-unit from *M. sedula*. Hence, a BPL was identified in the *M. sedula* genome (Msed_2010), and evaluated for possible inclusion in *P. furiosus*, which was the designated metabolic engineering host for the 3-HP/4-HB cycle to be used for production of fuels and chemicals from CO_2. The three subunits encoding E1 in *M. sedula* were expressed recombinantly in *P. furiosus*, along with co-expression of the *M. sedula* BPL. As a result, an active form of E1 could be produced, which was confirmed by biochemical characterization of E1, purified through a His-tag attached at the N-terminus of the γ sub-unit. E1 was then combined with purified recombinant E2 and E3 *in vitro*, and 3-HP was produced from HCO_3^- and acetyl-CoA. Both gas chromatography and high performance liquid chromatography-based methods were used to detect 3-HP formation from *in vitro* assembly of recombinant E2 and E3 using acetyl-CoA or malonyl-CoA as substrate.

Introduction

The 3-hydroxypropionate (3-HP)/4-hydroxybutyrate (4-HB) cycle in the extreme thermoacidophile *Metallosphaera sedula* catalyzes the incorporation of carbon dioxide (CO_2) into central metabolism intermediates (1). This cycle can also be used as a basis for fuels and chemical production in a metabolic engineering host (2,3). The first part of this cycle, referred to here as Sub-Pathway 1 (SP1), involves three enzymes: heteromultimeric acetyl-CoA/propionyl-CoA carboxylase (E1) encoded by Msed_0147, 0148, 1375 (4), malonyl-CoA reductase (E2) encoded by Msed_0709) (5), and malonic semialdehyde reductase (E3) encoded by Msed_1993 (6). E1, the carboxylase, is a heteromultimeric assemblage (α_4,β_4,γ_4), of which the β sub-unit requires post-translational biotinylation by an enzyme referred to as a biotin protein ligase (BPL). Additional details on these enzymes are provided below:

E1: Acetyl-CoA/propionyl-CoA carboxylase: This enzyme was initially identified by Fuchs and co-workers by purification of the native form from *M. sedula* (4). Their study confirmed this enzyme functions as a 560 kDa complex, comprised of biotin carboxylase (α unit), biotin carboxyl carrier protein (β unit), and carboxyl transferase (γ unit), encoded by Msed_0147, Msed_0148 and Msed_1375, respectively. To date, there have been no reports of a soluble, active form of the acetyl-CoA carboxylase produced recombinantly from any Sulfolobus species.

E2: Malonyl-CoA reductase: This enzyme catalyzes the conversion of malonyl-CoA to malonic semialdehyde, while consuming NADPH as reductant. The native form of this protein was purified from *M. sedula* and the recombinant form was obtained by cloning the gene encoding a homolog from *Sulfolobus tokodaii* in *E. coli* (5). Subsequently, the *M.*

sedula genome sequence indicated that this enzyme was encoded by Msed_0709. Transcriptomic analysis showed the up-regulation of this ORF under autotrophic growth (7). To date, there have been no reports on the biochemical characteristics of the *M. sedula* version of malonyl-CoA reductase.

E3: Malonic semialdehyde reductase: Similar to E2, malonic semialdehyde reductase is also NADPH-specific, catalyzing the conversion of malonic semialdehyde to 3-hydroxypropionate. The native form of this protein was purified from *M. sedula*, and the gene encoded by Msed_1993 was confirmed to be this enzyme through heterologous expression in *E. coli* (6).

BPL: Biotin protein ligase, an accessory enzyme for activating E1.

Biotin, or vitamin H, is required by all forms of life (8). The free form of biotin does not play any vital metabolic role until it is covalently bound to protein. Biotin is synthesized by plants, most bacteria and some fungi, and covalently attached at the active site of certain enzymes that transfer carboxyl-groups from bicarbonate to organic acids to form cellular metabolites (8). These biotin-dependent enzymes include acetyl-CoA carboxylase, propionyl-CoA carboxylase, pyruvate carboxylase, β-methyl-crotonyl-CoA carboxylase, transcarboxylase, and oxaloacetate decarboxylase (8,9). The attachment of biotin to the active site of biotin-dependent enzymes is catalyzed by biotin protein ligase (BPL) (10). Since most organisms have fewer than five protein species to be biotinylated, this post-translational modification is extraordinarily specific (8). Of interest here is acetyl-CoA/propionyl-CoA carboxylase in *M. sedula*, and the need to biotinylate its β sub-unit to activate the enzyme complex.

As described in Chapter 1, the efficiency of CO_2 fixation step in SP1 of the 3-HP/4-HB CO_2 fixation cycle relies on domain-to-domain (or subunit-to-subunit) interaction (i.e., subunits of E1 acetyl-CoA carboxylase), and enzyme-to-enzyme interaction (i.e., acetyl-CoA carboxylase and biotin protein ligase). The multiple-domain biotin-dependent E1 (acetyl-CoA carboxylase) consists of the biotin carboxylase (BC), carboxyl transferase (CT) and biotin carboxyl carrier protein (BCCP). The BCCP functions as substrate of BC (for accepting carboxyl group from bicarbonate) and CT (for carboxyl transfer to acetyl-CoA), only after having its canonical lysine residue biotinylated via biotin protein ligase (BPL). Hence, the understanding of recognition and specificity of biotinylation on BCCP can potentially provide insightful ways to improve E1 functionality.

Classification and functions of BPL.

BPLs have been grouped into four sub-classes, according to their domain structures and their participation in biotin metabolism (11) (see **Figure 1**). Class II has an N-terminal DNA-binding domain that enables BPL to act as a transcriptional repressor for the biotin operon, more than as a biotin transfer/attachment enzyme (10). For example, *E. coli* BirA is a bi-functional class II BPL that has regulatory properties. It is a monomeric protein of 35.5 kDa with three distinct domains (12): an N-terminal DNA binding domain with helix-turn-helix (HTH) structure, a central catalytic domain, and a C-terminal domain with structural similarity to the SH3 domain of tyrosine kinases (13), and suggested to interact with ATP and apo-BCCP biotin domain (14). BirA can switch from the biotin ligation function to the repressor function in response to the intracellular biotin supply and biotin acceptor protein (e.g. apo-BCCP) levels (15,16). This mediation process (see **Figure 2**) can be ascribed to the competition between apo-BCCP and operator DNA for the BirA: biotinyl-5'-AMP complex. The complex can either be attacked by the ε-group of the specific lysine of apo-BCCP to form biotinylated BCCP (holo-BCCP), or, under the low supply of apo-BCCP or excess biotin, remain firmly bound and can repress biotin biosynthesis by binding BirA N-terminal HTH domains to the promoter of biotin operon (bioABCFD) (17). Such a dual role is also found in *Bacillus subtilis* BPL (18). It has been demonstrated that Ser32, Arg33, and Ala 34 of *E. coli* BirA and Glu23, Trp38, and Gly 59 of the *B. subtilis* BPL play critical roles in their repressor activity (19,20).

As a representative of class I, BPL from *Pyrococcus horikoshii* OT3 (PhBPL) lacks the N-terminal DNA-binding domain (11) and possesses only ligase activity, but the overall fold of PhBPL is highly similar with the catalytic and C-terminal domain of *E. coli* BirA. The three-dimensional structure of PhBPL reveals a homodimer as the functional unit. Each subunit contains a larger N-terminal catalytic domain and a smaller C-terminal domain (20). *E. coli* BirA dimerizes only after the addition of the biotin substrate. The dimerization of *E. coli* BirA, complexed with biotinyl-5'-AMP, contributes to the transcriptional repression function: excess biotin may facilitate the dimerization, which in turn, inhibits the biosynthesis of biotin (21). Since PhBPL is devoid of regulatory function, its dimeric structure might contribute to thermostability (20,22).

There are two other classes of BPL that have different domain architecture from Class I and II (see **Figure 1**). Both the human and yeast BPL belong to class IV, polypeptides consisting of 600 to 700 amino acids. They contain a large additional N-terminal domain of some sequence similarity, but not a HTH DNA-binding domain as described in *E. coli* BirA.

There is, as yet, little knowledge about the function of this region (23). The Class III BPL from a plant *Arabidopsis thaliana* is intermediate in size between the bacterial and other eukaryotic enzymes, with a predicted molecular mass of 41.1 kDa (24). Besides the catalytic domain and C-terminal domain, it carries an N-terminal organelle targeting sequence (chloroplast or mitochondrial). The absence of a DNA-binding domain, and thereby, the repressor function, in these eukaryotic BPLs is consistent with their biotin metabolism (10). Human and yeast can acquire biotin from exogenous sources, without relying on a bacterial-like biotin synthesis network, so a repressor function is not expected for these BPLs. Though plants do synthesize biotin, their 2,000-fold higher intracellular free biotin level than that of bacteria implies a strong repression mechanism is unnecessary (24).

As mentioned previously, biotinylation ends with the formation of an amide linkage between the carboxyl group of biotin and the ε-group of a specific lysine residue of apo-BCCP in a two-step reaction (10). In the first step, BPL catalyzes the nucleophilic attack by an oxygen atom of the biotin carboxylate group on the α-phosphate group of ATP, leading to the formation of biotinyl-5'-AMP. This intermediate remains bound in the active site, and in the second step, is attacked by the ε-amino group of the specific lysine on apo-BCCP to form an amide bond, with release of AMP. Hence, the biotin moiety remains attached throughout the lifetime of the protein acceptor (20,25).

Structure and mechanism of BPLs.

Although the *E. coli* BirA has been extensively studied by crystallography and biochemical analysis, its reaction mechanism remains unknown (20). Structural study of BirA is complicated by a number of issues: first, several interesting parts of *E. coli* BirA, thought to be responsible for substrate binding, were disordered in the crystal structures (12,26). Secondly, the complex crystals of *E. coli* BirA, with biotin or with biotinyl-5'-AMP, present experimental difficulties (27). Therefore, the exact locations of the nucleotide ligands on *E. coli* BirA cannot be determined, nor can the catalytic residues be identified.

Wilson *et al.* proposed a biotin-binding site and an ATP-binding site by attaching biotinyl-lysine to the *E. coli* BirA crystals, assuming the complex can mimic the enzyme-product state (12). A glycine-rich sequence (GRGRRG), in a partially disordered loop in the catalytic domain of the BirA crystal structure, was thought to be the ATP-binding site, close to the postulated biotin-binding site. This GRGRXG motif, where X is any amino acid residue, is highly conserved among BirA sequence homologs, and has similarity to mononucleotide binding sequences in protein kinases (19). However, Kwon and Beckett (28) presented strong experimental evidence disagreeing with that prediction. In their experiment, single-site mutants, G115S, R118G, and R119W, of the *E. coli* BirA were biochemically characterized with respect to binding of small molecule substrates and the biotinyl-5'-AMP intermediate. It was demonstrated that, rather than an ATP-binding site, this motif is actually for binding biotin and the biotinyl-5'-AMP intermediate in the biotinylation reaction.

Kinetic measurements indicate biotin and ATP are bound to BirA, with biotin binding first (27). Biotin binding causes a large structural change thought to facilitate ATP binding. Responsible for regulatory function, the N-terminal DNA binding domain of BirA is connected to the rest of the molecule through a hinge, showing minimal interaction in the crystal structure. However, it may still contribute to the tight binding of biotin and the biotinyl-5'-AMP to the catalytic domain (27). Such large conformational changes that BirA potentially undergoes during substrate binding, as concluded by Wilson (12), is consistent with the failure of attempts to co-crystallize BirA-biotin or BirA-biotinyl-5'-AMP complexes. The diffusion of the ligands into BirA crystals always causes cracking of the crystals. Only when Bagautdinov *et al.* determined the crystal structures of phBPL and its co-crystallized complexes with biotin, ADP, and biotin-5'-AMP (20), could a detailed structural view be obtained to explain the reaction mechanism of the first-step of biotinylation. This also identified the important catalytic residues and exact locations of bound substrate molecules and the intermediate biotin-5'-AMP involved in the reaction.

As Bagautdinov *et al.* described (20), the structure of phBPL is arranged to properly position the substrates (biotin and ATP), and minimize the activation energy barrier for the withdrawal of electrons from the phosphate groups of ATP and the biotin carboxylate group. This involves the suitable configuration of catalytic binding site and substrate pockets. The biotin-binding pocket of phBPL has a hydrophobic wall composed of five amino acid residues, with a hydrophilic bottom composed of 14 amino acid residues. Biotin is settled in the pocket via its carboxyl group and ureido nitrogen to interact with the hydrophilic interior. It uses its hydrophobic tail and thiophene ring to interact with the hydrophobic wall. Hence, orientation of biotin is fixed by this pocket. The position of adenine binding was found spatially adjacent to the biotin-binding pocket, held by a network of electrostatic and polar interactions. Orientation of the adenine ring is defined by aromatic-aromatic interactions, so the close placement of the reactive ends of biotin and ATP can be achieved. An active site β2-β3 loop becomes fully ordered upon binding ligands, possibly acting as a partial lid to prevent premature release of the ligands from the active site. Moreover, the pocket for biotinyl-5'-AMP binding was found to mimic the combination of that for ADP and biotin; hence, conformational change of phBPL during reaction can be minimized. This is also supported by the similarity between the overall structure of the un-liganded and the liganded forms of phBPL, which are only different in an ordering of the active site loop upon ligand binding. Also, some positively charged residues Arg48, Arg51, Lys111 and Arg233, were proposed to facilitate the formation of biotinyl-5'-AMP, as they are located within hydrogen-bonding distances from the ligands, co-stabilizing the negatively charged transition state intermediate via electrostatic interaction. Lys111, in particular, plays the most critical role. It can increase the electrophilicity of the α-phosphate group of ATP by withdrawing an electron so as to trigger the nucleophilic attack by the carboxylic group of biotin. In addition, Lys111 is thought to get involved in the second half reaction of biotinylation, by facilitating the nucleophilic attack of the active site lysine of BCCP.

Substrate specifity of biotinylation.

Before this study, there had been no reports concerning the BPL in *M. sedula*, although in a closely related species, *S. tokodaii*, the BPL was identified and its substrate specificity for biotinylation was determined. *S. tokodaii* can biotinylate its own BCCP, but not *E. coli* BCCP; likewise, BPL from *E. coli* can biotinylate its own BCCP but not *S. tokodaii* BCCP (29,30). It was discovered that the local sequence surrounding the canonical lysine of BCCP is conserved in many organisms, including archaea, in the form of -E-X-M-K-M- (e.g. *E. coli*) , though in some cases one or both methionines may be replaced with leucine or valine (31). For *S. tokodaii* BCCP, the C-terminal methionine is replaced with serine, significantly different from methionine or leucine in terms of size and nature of the side chain group (29).

Actually, this substitution occurs in the Crenarchaeal Sulfolobaceae, in species such as *S. acidocaldarius*, *S. metallicus,* and *Acidianus brierleyi,* which affects the substrate specificity of biotinylation, but is not a sole determinant, as concluded by Li *et al* (29). They converted the serine to methionine in *S. tokodaii* BCCP (S136M) and the methionine to serine in *E. coli* BCCP (M123S). Both mutants can then be biotinylated by heterologous BPLs, but remain as poor substrates. On the other hand, the mutated BCCPs dramatically decrease their affinity to their homologous BPLs.

So far, there has been no available three-dimensional structure of Sulfolobaceae BCCP and BPL showing their mutual recognition and interaction to fully explain the substrate specificity of biotinylation. Therefore, a functional BPL from *M. sedula* must overcome the potential deficiencies expected if expressed in *E. coli* or *P. furiosus* as recombinant host for active E1.

Materials and Methods

Cloning and heterologous expression of *M. sedula* E1 genes in *E. coli*

E1-acetyl-CoA/propionyl-CoA carboxylase. The *M. sedula* genes encoding three different subunits of E1 were cloned into pET46 Ek/LIC, pRSF-2 Ek/LIC, and pCDF-2 Ek/LIC (EMD4 Biosciences), respectively, which are compatible vectors enabling co-proliferation and co-expression of these genes in the *E. coli* strain (see **Table 1**). These vectors all contain a ligation independent cloning (LIC) site that entails the directional cloning of PCR products, without the need for restriction enzyme digestion or ligation reactions. The vectors containing target genes were co-transformed as needed (such as expression of β and co-expression of αβγ, αβ, and βγ), into competent *E. coli* Rosetta (DE3) cells, capable of encoding rare tRNAs required for translation of *M. sedula* enzymes. Cells were grown in 1L LB media, containing 100 μg ml^{-1} of ampicillin (for pET46), 30 μg ml^{-1} of kanamycin (for pRSF-2), 50 μg ml^{-1} streptomycin (for pCDF-2), and 34 μg $ml^{-1}$of chloramphanicol according to the different plasmids carried by *E. coli* strains. Expression was induced with 0.1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) (optical density at 600 nm, 0.6); cells were harvested 4 to 5 h after induction (37°C) by centrifuging and then were stored at -20 °C until they were used.

E2-malonyl-CoA reductase. The gene encoding the malonyl-CoA reductase in the *M. sedula* genome (Msed_0709) was amplified by using forward primer 5'-AAAAAA<u>CATATG</u>AGGAGAACGCTAAAGGCCGCAATTC -3', introducing an NdeI site at the initiation codon, and reverse primer 5'-AAAAAA<u>CTCGAG</u>TCATCTCTTGTCTATGTAGCCC -3', introducing an XhoI site after the stop codon. The PCR conditions were as follows: initial denaturation for 1 min at 95°C, 40 cycles of denaturation for 45 sec at 95°C, primer annealing for 30 sec at 61°C, and extension for 70 sec at 72°C, and a final extension for 5 min at 72°C. The PCR product was purified and cloned into pET21b (+) vector, resulting in plasmid pET21b-mse0709. The plasmid was transformed into Competent *E. coli* Rosetta (DE3) cells, capable of encoding rare tRNAs required for translation of *M. sedula* enzymes. Cells were grown in 1L LB media containing 100 μg ml^{-1} of ampicillin and 34 μg ml^{-1} of chloramphanicol. Expression was induced with 15 μM isopropyl-β-D-thiogalactopyranoside (IPTG) (optical density at 600 nm, 0.6); cells were harvested 4 to 5 h after induction (37°C) by centrifuging and then stored at -20 °C until they were used.

E3-malonic semialdehyde reductase. The gene encoding the malonic semialdehyde reductase in the *M. sedula* genome (Msed_1993) was amplified by using forward primer 5'-CCCCCCCAT ATGACTGAAAAGGTATCTGTAGTTG -3', introducing an NdeI site at the initiation codon, and reverse primer 5'-CCCCCCAAGCTT TTTTTCCCAAACTAGTT TGTATA-3', introducing an HindIII site after the stop codon. The PCR conditions were as follows: initial denaturation for 1 min at 95 °C, 40 cycles of denaturation for 45 sec at 95 °C, primer annealing for 30 sec at 61.9 °C, and extension for 58 sec at 72°C, and a final extension for 5 min at 72°C. The PCR product was purified and cloned into pET21b (+) vector, resulting in plasmid pET21b-mse1993. The plasmid was transformed into Competent *E. coli* Rosetta (DE3) cells, capable of encoding rare tRNAs required for translation of *M. sedula* enzymes. Cells were grown in 1L LB media containing 100 μg ml^{-1} of ampicillin and 34 μg ml^{-1} of chloramphanicol. Expression was induced with 0.2 mM isopropyl-β-D-thiogalactopyranoside (IPTG) (optical density at 600 nm, 0.6); cells were harvested 4 to 5 h after induction (37°C) by centrifuging and then stored at -20 °C until they were used.

Biotin protein ligase. The gene encoding the putative biotin protein ligase in the *M. sedula* genome (Msed_2010) was amplified by using forward primer 5'-GACGACGACAAGATGCAGATAGTTAAACTGGACAGAGTT-3', and reverse primer 5'-GAGGAGAAGCCCGGTTACGTCGTCTCAAGCCTT-3'. The PCR conditions were as follows: initial denaturation for 1 min at 95°C, 40 cycles of denaturation for 45 sec at 95°C, primer annealing for 30 sec at 60 °C, and extension for 41 sec at 72°C, and a final extension for 5 min at 72°C. The PCR product was purified and cloned into pET-46 Ek/LIC vector. The resulted plasmid was transformed into Competent *E. coli* Rosetta (DE3) cells, capable of encoding rare tRNAs required for translation of *M. sedula* enzymes. Cells were grown in 1L LB media containing 100 μg ml^{-1} of ampicillin and 34 μg ml^{-1} of chloramphanicol. Expression was induced with 0.1mM isopropyl-β-D-thiogalactopyranoside (IPTG) (optical density at 600 nm, 0.6); cells were harvested 4 to 5 h after induction (37°C) by centrifuging and then stored at -20 °C until they were used.

Cloning and heterologous expression of *M. sedula* E1 in *Pfu*

PCR products of the promoter of the *P. furiosus* PEP synthase gene (Ppep, consisting of a 123 bp DNA fragment immediately upstream of PF0043), the *M. sedula* biotin protein ligase (BPL) gene (Msed_2010), and the *M. sedula* carbonic anhydrase (CA) gene (Msed_0390) were combined into a synthetic operon via splice-overlap extension (SOE) PCR (32). As part of the artificial operon design, a 14-bp sequence containing the S-layer gene ribosome binding site (RBS) was inserted (via primers) in between the BPL and CA genes (5'-ggaggtggagaaaa). PCR products of 0.5-kb upstream and downstream flanking regions to genome region 5 (between convergent genes PF1232 and PF1233) were combined via Gibson Assembly (using an NEB Gibson Assembly kit) with the *pyrF* pop-out marker cassette (3), Pslp-E1αβ, his6E1γ, and with the Ppep-BL-CA operon, to construct the plasmids pGL033 (see **Figure 3**) for chromosomal insertion of PpepBLCA-Pslp-E1αβ-his6E1γ at genome region 5 in *P. furiosus* COM1. The Pslp-E1αβ-his6E1γ artificial operon contained a *P. furiosus* RBS from the gene encoding pyruvate ferredoxin oxidoreductase subunit γ (PF0791, 5'- ggaggtttgaag) upstream of the E1 γ subunit gene, and a 6x His tag flanked by two alanine codons (5'-gcacatcaccaccaccatcacgct) was also inserted after the start codon of the E1γ gene to facilitate protein purification. The *P. furiosus* genome regions (3 and 5) used for chromosomal integration of the constructs contain little to no transcriptional activity as determined from analysis of tiling array data (33). *P. furiosus* COM1 was transformed using uracil prototrophic selection on defined medium, essentially as previously described (34), using linearized plasmids pGL033 to construct strains MW112 (See **Figure 3**). Strains were further purified by two consecutive transfers on solid medium, and the final strains were verified by PCR and sequencing of the regions containing the chromosomal insertions. Strains used and constructed in this study are listed in **Table 2**.

MW112 growth.

The MW112 strain was cultured as previously described (34) in a seawater-based medium containing 5 g/L maltose, 5 g/L yeast extract, 5 g/L tryptone and 0.25 mg/L biotin. Cultures in 2-L fermenters were first grown at 95°C, sparged with 20% CO_2/80% N_2 (v/v) (15mL/min) and agitated (250 rpm) until cell density reached 10^8 cells/mL, at which point the temperature was rapidly switched to 72°C (within 10 min) for *M. sedula* enzymes optimal activity and maintained at the lower temperature for up to 18 hours. Cells were harvested by centrifuging and then stored at -80°C until used.

Purification of recombinant SP1 and E1 enzymes

Recombinant, non-tagged malonyl-CoA reductase. *E. coli* cells were re-suspended in 50 mM Tris/HCl, 5 mM $MgCl_2$, pH 7.8 and disrupted by sonication for 10 min at 60% Amplitude (10 sec on, 10 sec off). Cell extract was incubated at 75°C for 15 min to denature *E. coli* native proteins, and the supernatant, after centrifuging (9,400 × g, 20 min, 4°C), was collected for purification. Ammonium sulfate was added to the heat-treated cell extract to a final concentration of 1 M. The resulting precipitated protein was removed by centrifuging at 11,000 x g for 10 min, and the supernatant was directly applied to a Phenyl-Sepharose HIC column (GE Healthcare Life Sciences; volume, 20 mL) at a flow rate of 1 ml min^{-1}. The column has been equilibrated with 100 mM sodium phosphate, 1 M ammonium sulfate, pH 7.8. After washing the column with 4 bed volumes of this buffer at flow rate of 1 ml min^{-1}, proteins were eluted with a 210 ml decreasing linear gradient of 1 M to zero ammonium sulfate at 0.7 ml min^{-1}. The activity eluted between 250 mM and 380 mM salt, and the active fractions were pooled and concentrated via Amicon® Ultra-15 Centrifugal Filter Units (10 kDa) (EMD Milipore).

Recombinant C-terminal 6×His-tagged malonic semialdehyde reductase. *E. coli* cells were re-suspended in 20 mM sodium phosphate, 0.5 M NaCl, pH 7.4, and disrupted by sonication for 10 min at 60% Amplitude (10 sec on, 10 sec off). Cell extract was incubated at 65°C for 20 min to denature *E. coli* native proteins, and the supernatant, after centrifuging (16,000 × g, 15 min, 4°C), was collected for purification. The heat-treated cell extract was loaded on a HiTrap IMAC HP column (GE Healthcare Life Sciences; volume, 5 mL), equilibrated with 20 mM sodium phosphate, 0.5 M NaCl, at pH 7.4. The column was washed with five bed volumes of equilibration buffer and developed with a step gradient to 300 mM imidazole at a flow rate of 1 ml min^{-1}. The most active fractions eluted with imidazole were pooled and dialyzed in to 20 mM Tris, 1 mM DTT, 100 mM KCl, pH 8.0 to remove imidazole. The purified protein solution was concentrated via Amicon® Ultra-15 Centrifugal Filter Units (3 kDa) (EMD Milipore) and stored at -80°C.

Recombinant N-terminal 6×His-tagged biotin protein ligase. *E. coli* cells were re-suspended in 20 mM sodium phosphate, 0.5M NaCl, pH 7.4, and disrupted by sonication for 10 min at 60% Amplitude (10 sec on, 10 sec off). Cell extract was incubated at 65°C for 20 min to denature *E. coli* native proteins, and the supernatant after centrifuge (16,000 × g, 15 min, 4°C) was collected for purification. The heat-treated cell extract was loaded on a HiTrap IMAC HP column (GE Healthcare Life Sciences; volume, 5 mL) equilibrated with 20 mM sodium phosphate, 0.5 M NaCl, at pH 7.4. The column was washed with five bed volumes of equilibration buffer and developed with a step gradient to 300 mM imidazole at a flow rate of 1 ml min^{-1}. The most active fractions eluted with imidazole were pooled and dialyzed in 50mM Tris/HCl, pH 8.2 to remove imidazole. The purified protein solution was concentrated via Amicon® Ultra-15 Centrifugal Filter Units (3kDa) (EMD Milipore) and stored at -80°C.

Recombinant N-terminal 6×His-tagged (at γ unit) acetyl-CoA carboxylase. *Pfu* cells were re-suspended and lysed in a low salt buffer 50 mM Tris/HCl, pH 8.2 by shaking for 2 h at room temperature. The cell lysate was centrifuged at 10,000 × g for 20 min, and the supernatant was collected and sterile filtered before being loaded on a HiTrap IMAC HP column (GE Healthcare Life Sciences; volume, 1 mL) equilibrated with 50 mM Tris, 100 mM Nacl, pH 8.0. The column was washed with five bed volume equilibration buffer and developed with a step gradient to 500 mM imidazole at a flow rate of 1 ml min^{-1}. The most active fractions eluted with imidazole were pooled and dialyzed in 50 mM Tris/HCl, pH 8.2 to remove imidazole. The purified protein solution was concentrated via Amicon® Ultra-15 Centrifugal Filter Units (3kDa) (EMD Milipore) and stored at -80°C, 100 mM Nacl, pH 8.0. The column was washed with five bed volume equilibration buffer and developed with a step gradient to 500 mM imidazole at a flow rate of 1 ml min^{-1}. The most active fractions eluted with imidazole were pooled and dialyzed in 50 mM Tris/HCl, pH 8.2 to remove imidazole. The purified protein solution was concentrated via Amicon® Ultra-15 Centrifugal Filter Units (3kDa) (EMD Milipore) and stored at -80°C.

Molecular mass determination

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was performed, as described by Laemmli (35), using 4-12% gels (Invitrogen). Proteins were visualized by Coomassie blue staining (36). The native molecular mass of the enzyme was estimated by gel filtration chromatography. Protein from the IMAC or HIC steps was applied to a Superdex 200 HR 16/60 gel filtration column (GE Healthcare Life Sciences; volume, 20 ml), which had been equilibrated with 100 mM sodium phosphate, pH 7.8. The flow rate was 0.5 ml min^{-1}. The column was calibrated with the following molecular mass standards (Sigma): thyroglobulin (669 kDa), apoferritin (443 kDa), β-amylase (200 kDa), alcohol dehydrogenase (150 kDa), bovine serum albumin (66 kDa), carbonic anhydrase from bovine erythrocytes (29 kDa), cytochrome c (12.4 kDa), and aprotinin (6.5 kDa).

Enzyme assays

Enzymatic assay of malonyl-CoA reductase. Malonyl-CoA reductase activity was monitored in a spectrophotometric assay based on previous reports (5), but with some modifications. Malonyl-CoA-dependent oxidation of NADPH was followed spectrophotometrically at 340 nm (ε (340 nm) NADPH = 6,200 M^{-1} cm^{-1}) and 70°C. The assay mixture contained 50 mM Tris/HCl (pH 7.8), 5 mM $MgCl_2$, 5 mM 1,4-dithioerythritol, 0.5 mM NADPH, 0.4 mM malonyl-CoA, and cell extract, partially purified enzyme, or purified enzyme. The addition of malonyl-CoA started the reaction. The apparent K_m values were determined by varying the concentration of malonyl-CoA (0.01 to 0.4 mM), while keeping the co-substrate NADPH at saturating concentrations (0.5 mM). One unit of enzyme activity refers to 1 µmol of NADPH oxidized per min, corresponding to 1 µmol of malonyl-CoA reduced to malonate semialdehyde per min. Protein was quantitated by the method of Bradford (37), using bovine serum albumin as the standard.

Enzymatic assay of malonic semialdehyde reductase. Malonic semialdehyde reductase activity was monitored using an E2-E3 coupled assay based on (6) with some modifications. Malonic semialdehyde was enzymatically synthesized from malonyl-CoA using recombinant malonyl-CoA reductase. The malonic semialdehyde-dependent oxidation of NADPH was followed spectrophotometrically at 340 nm at 65°C. The assay mixture contained 100 mM MOPS-KOH (pH 7.6), 5 mM $MgCl_2$, 5 mM 1,4-dithioerythritol, 0.5 mM NADPH, 20 U recombinant malonyl-CoA reductase, and 0.2 mM malonyl-CoA. After 5 min of incubation to allow the completion of E2 reaction, the E3 reaction was started by adding purified recombinant malonic semialdehyde reductase.

Enzymatic assay of acetyl-CoA carboxylase. An E1-E2 coupled spectrophotometric assay was developed, based on (38) with some modification. Purified recombinant acetyl-CoA carboxylase was added to the assay mixture containing 80 mM Mops/NaOH pH 8.0, 0.5 mM acetyl-CoA, 5 mM $MgCl_2$, 5 mM 1,4-dithioerythritol, 5 mM ATP, 10 mM $NaHCO_3$, 0.5 mM NADPH and purified recombinant malonyl-CoA reductase. The acetyl-CoA-dependent reduction of NADPH was followed spectrophotometrically at 340 nm at 70°C. In this coupled enzyme assay, the product of acetyl-CoA carboxylase, malonyl-CoA, is further reduced to malonic semialdehyde and one molecule of NADPH is oxidized. A phosphate colorimetric assay was also used to measure E1 activity. Since the reaction catalyzed by E1 needs to consume ATP and free phosphate ion is released, a phosphate colorimetric assay was used to measure the specific activity of purified recombinant acetyl-CoA carboxylase acting on acetyl-CoA. The assay follows the protocol of BioVision phosphate colorimetric assay kit. Briefly, recombinant purified E1 was added to the substrate mixture containing 75 mM MOPS/NaOH pH 8.0, 0.5 mM acetyl-CoA, 5 mM $MgCl_2$, 5 mM 1,4-dithioerythritol, 2 mM ATP, and 10 mM $NaHCO_3$ incubated at 70°C. Samples were taken periodically from start point to 6 min and mixed with phosphate reagent (malachite green and ammonium molybdate). After incubation for 30 min at room temperature, the phosphate ion from samples forms chromogenic complex with phosphate reagent and was measured spectrophotometrically at 650 nm.

Biotinylation assays

Biotin quantification. The amount of biotin incorporated into proteins was analyzed using Pierce Biotin Quantitation Kit. In this colorimetric assay (see **Figure 4**), a solution containing the biotinylated protein was added to a pre-mix of 4'-hydroxyazobenzene-2-carboxylic acid (HABA) and avidin. The biotin has higher affinity to avidin so as to displace the weakly bound HABA and the absorbance at 500 nm decreases. The change in absorbance is proportional to the amount of biotin in the sample by the extinction coefficient of the HABA-avidin complex (ε = 34,000 M^{-1} cm^{-1}). The biotin: protein molar ratio was calculated as shown in **Table 3**.

***In vitro* biotinylation assay.** Biotin protein ligase (BPL) activity was measured by following biotinylation of purified recombinant E1-β or E1-βγ. Assay conditions were adapted according to (39), and the assay contained 30 mM Tris, pH 8.2, 100 mM KCl, 5 mM $MgCl_2$, 0.5 mM DTT, 0.3 mM ATP, 10 μM biotin and 1 μM purified recombinant putative BPL (Msed_2010). The reaction was initiated by the addition of purified recombinant E1-β or E1-βγ to a final concentration of 2.5 μM and incubated for 30 min at 60°C and 70°C or 4 h at 37 °C, respectively. As a control, E1-β or E1-βγ were incubated with BPL storage buffer instead. After incubation, samples were filtered through YM 10 to remove excess biotin and re-dissolved in 50 mM Tris/HCl, pH 8.2.

Visualization of biotinylated protein. A biotin capture method based on streptavidin sepharose beads (GE Healthcare) was developed to show biotinylated protein on SDS-PAGE. The products from *in vitro* biotinylation (including control) was loaded onto streptavidin sepharose beads and only the biotinylated protein was bound to the streptavidin. Heating the beads at 95°C for 5 min denatures the streptavidin and breaks the binding so the biotinylated protein is released from the beads.

Mass spectrometry characterization of proteins

The co-expressed recombinant E1 (from *E. coli*) was run on SDS-PAGE with two predominant bands (~20 kDa and 60 kDa). An in-gel tryptic digestion was performed on each band and peptides were extracted for interrogation by nano-LC-MS/MS on an LTQ-Orbitrap XL mass spectrometer in triplicate. The resulting data files were processed and searched against a *M. sedula* protein database using the MASCOT searching algorithm and

ProteoIQ was used for comparative analysis. Normalized Spectral Abundance Factor (NSAF) normalization was performed to obtain quantitative information for each protein.

Assembly of recombinant E1 to E3 *in vitro*

Recombinant E1 to E3 were added to substrate mixture (80 mM MOPS/NaOH pH 8.0, 2.5 mM acetyl-CA, 5 mM $NaHCO_3$, 5 mM ATP, 5mM NADPH, 5 mM $MgCl_2$, and 5 mM DTT) and incubated at 70°C for 20 min. Recombinant E2 to E3 were added to substrate mixture (80 mM MOPS/NaoH pH 8.0, 2.5 mM Malonyl-CA, 5 mM NADPH, 5 mM $MgCl_2$, and 5 mM DTT) and incubated at 70°C for 20 min.

Gas chromatography (GC) method for 3-HP detection

The resulting products from *in vitro* assembly of enzymes were dried by using rotary evaporator at 50°C. Pyridine (100 µL) was added to re-suspend dry samples, and then 50 µL N,O-Bis(trimethylsilyl)trifluoroacetamide (BSTFA) was added to derivatize samples, and the precipitates was removed by filtration. The supernatant was directly injected into the GC. Quantitative GC analysis was carried out on a HP6890 GC, equipped with a flame ionization detector and HP-1 column (30 m × 0.32 mm × 0.25 µm). The injection temperature was 200˚C, the detector temperature was 270˚C, and the helium flow rate was 2 mL min^{-1}. The column was held for 3 min at 120˚C, raised at 5˚C min^{-1} to 200˚C, followed by 10˚C min^{-1} to 260˚C, and kept isothermal at 260˚C for 5 min.

High performance liquid chromatography-based (HPLC) methods for 3-HP detection.

3-HP (H0297, 30%, w/v in water) was obtained from TCA America (http: //www.tciamerica. net/). For HPLC analysis, 3-HP was derivatized with 2,4-Dibromoacetophenone (DBAP) as follows. The sample was acidified with H_2SO_4 to pH 2 and 3-HP was extracted by adding 1.5 mL diethyl ether to a 500 μL sample, centrifuging for 10 min at 6,000 × g to separate the phases, removing the top ether layer, adding 50 μL 20 mM $NaHCO_3$ to neutralize, and then evaporated overnight. The dried sample was re-suspended in 50 μL acetonitrile, and the pH was adjusted to 9-10 by adding 10 μL 100 mM KOH. Derivatization was started by adding 100 μL acetonitrile, 50 μL 1 μM 15-crown-5 ether and 200 μL 100 mM 2, 4-DBAP. The sample was then incubated at 80°C for 60 min. The derivatized samples were analyzed by HPLC at 254 nm wavelength (Waters, Alliance, PDA detector). The column and run conditions were as follows: column, Shodex C-18; solvent system, A: 0.05% TFA, B: acetonitrile; gradient 0-15 min, 30-80% B, flow rate: 1 mL/min; temperature: 30°C.

Bioinformatics

Nucleotide and amino acid sequences of CAs were obtained from databases at NCBI (http://www.ncbi.nlm.nih.gov). The alignments of amino acid sequences were performed with Clustal Ω (40), and displayed with ESPript 3.0 (41). The analysis of domain architecture was performed by using pfam (42).

Results and Discussion

Biochemical characterization of recombinant E2 and comparison with native E2 and homologs from *S. tokodaii.* Initial efforts to clone and express E2 with an N-terminal His-tag in *E. coli* resulted in insoluble protein products. The protein, re-solubilized and refolded from inclusion bodies, failed to bind to the IMAC column. Re-design of the primers to remove the His-tag led to increased production of soluble E2, though significant amount of this protein still existed as inclusion bodies. The non-tagged recombinant E2 was purified by heat-treatment, hydrophobic interaction chromatography and size exclusion chromatography. Michaelis-Menten kinetic parameters of the purified enzyme were comparable to those of its homologs from *S. tokodaii* (5) (see **Table 4**). However, no RNA association was found for recombinant E2 from *M. sedula*, in contrast to reports for the recombinant E2 homolog from *S. tokodaii*. Gel filtration analysis indicated that E2 is a homotrimer. Molecular mass and subunit composition for all the SP1 enzymes in this study are listed in **Table 5**. The temperature optimum of purified recombinant E2 was 85°C, with a half-life at 72°C of 5.5 hours, 2 min at 95°C, the optimal temperature for *Pfu* growth (see **Figure 5-7**).

Biochemical confirmation of recombinant E3. The activity of purified recombinant E3 was confirmed based on an E2-E3 coupled assay. Since the E3 substrate, malonic semialdehyde, is not commercially available, the assay mixture used malonyl-CoA and the substrate. Recombinant E2 utilizes NADPH to synthesize malonic semialdehyde, and the addition of purified recombinant E3 consumed additional NADPH (see **Figure 8**).

Inability to express active E1 expression in *E. coli* and failure of post-modification (biotinylation). Initial efforts to express E1 subunits individually in *E. coli* resulted in insoluble E1-α and E1-γ units. Therefore, co-expression of three subunits was tried. Since both α and γ units run as 57 kDa bands on SDS-PAGE, proteomics analysis was used to determine levels of expression and indicated extremely low levels of the soluble α subunit were produced (see **Figure 9**). An additional drawback of *E. coli* as host to express functional E1 was that the *M. sedula* β unit was not biotinylated in *E. coli*, as indicated by a biotin quantitation assay (see **Table 3**). This was consistent with previous findings that the

BPLs of *S. tokodaii* and *E. coli* do not biotinylate BCCP from each other, potentially because of the characteristic local sequence motif surrounding the canonical lysine of BCCP (-M-K-M-) (29,30). Amino acid sequence comparisons between BCCP of *S. tokodaii* and *M. sedula* showed that the two enzymes were similar and likely explains the failure of *E. coli* to biotinylate the *M. sedula* enzyme. The low expression of E1-α in *E. coli* led to attempts to produce E1 in an alternative host, in this case *Pfu*, which would be the designated host for the pathway of interest. However, amino acid sequence comparison revealed similarity between *Pfu* BCCP and *E. coli* BCCP for the –M-K-M- motif, suggesting that the *Pfu* BPL probably would not function efficiently to biotinylate *M. sedula* E1. Thus, the recruitment of the BPL from *M. sedula* to achieve expression of active BCCP in *P. furiosus* was pursued.

Identification and biochemical characterization of biotin protein ligase from *M. sedula*. In view of the fact that a biotin protein ligase (BPL) had been identified in *S. tokodaii* (Stok BPL) (29), this sequence was used to query the NCBI database. Msed_2010 appeared to be homolog to Stok BPL, with 56% identity at the amino acid level. Hence, this ORF was cloned in *E. coli*, and the overexpressed protein was purified. To test its function, an *in vitro* biotinyhlation assay was performed by incubating this protein with apo-BCCP (purified *M. sedula* E1-β expressed in *E. coli*). The protein encoded in Msed_2010 (Msed BPL) could biotinylate recombinant E1-β (with a 6× His tag on N-terminal) either at 37°C or at 60°C-70°C. The co-expressed E1-βγ (with a 6× His tag on N-terminal of γ) produced in *E. coli* could also be biotinlyated by Msed BPL (see **Figure 10**). This showed that the His-tag at the N-terminus of the γ subunit could be retained and used for purification of intact E1-αβγ expressed in *P. furiosus*. In comparing the Msed BPL and its close homolog Stok BPL with other BPLs, significant differences can be seen in domain architecture. Apparently, the versions from the extreme thermoacidophiles only have the BPL-LiA-LiB domain, while the homolog from *P. furiosus* (PF BPL) has an additional BPL-C domain. This suggests that the *M. sedula* BPL has no regulatory function, unlike the homolog in *E. coli* (Ec BPL), since it lacks an N-terminal DNA-binding domain present in Ec BPL.

Alignment of BCCPs sequences focusing on the characteristic local motif surrounding the canonical lysine (-E-X-M-K-M-) for *E. coli*, *M. sedula*, *S. tokodaii* and *Pfu*, indicated similarity between *E. coli* and *Pfu*, but a common replacement of methionine with serine in *M. sedula* and *S. tokodaii* (see **Figure 11**). Therefore, the native BPL of *Pfu* is probably ineffective in biotinylating the E1 β sub-unit from *M. sedula*. By comparison of both BPLs and BCCPs among *M. sedula*, *E. coli* and *Pfu*, the BPL newly identified from *M. sedula* genome (Msed_2010), was concluded to be a necessary element in the designated metabolic engineering host *Pfu*, for the 3-HP/4-HB cycle to be used for production of fuels and chemicals from CO_2.

Biochemical characterization of recombinant E1 heterologously expressed in *P. furiosus*. Both the NADPH-based E1-E2 couple assay and the phosphate colorimetric assay confirmed the activity of recombinant E1, the latter showing specific activity of 0.65 μmol/min/mg, lower than the native E1 from *M. sedula*, which had a specific activity of 3.2 μmol/min/mg (38). The recombinant E1 produced in *Pfu* should be further studied, especially with respect to expression levels for each subunit, to inform genetic efforts to enhance expression, such as different promoter design. SDS-PAGE showed that E1-β was expressed at relatively low levels compare to the other two bigger subunits, possibly explaining the difference in specific activity between native and recombinant forms (see **Figure 12-14**). Further proteomics analysis could resolve this issue and could incorporate 2-D gels to separate the α and γ units which have nearly identical M_r (57 kDA) but different pIs (pI of α is 6.71, while pI of γ is 6.07).

Assembly of recombinant E1, E2, E3 *in vitro* to produce 3-HP. Both the GC-based method and the HPLC-based method can show the 3-HP formation via *in vitro* assembly of recombinant E1, E2 and E3. (see **Figures 15-16** and **Table 6**). Here, *in vitro* SP1 has been demonstrated, which provides the basis for assembly of this sub-pathway into metabolically engineered host *Pfu*. Alternatively, an HPLC-based method was also used. This method should be further optimized for better separation of 3-HP from substrates, in order to analyze 3-HP in much more complex mixtures, in particular the *Pfu* culture medium.

REFERENCES

1. **Zarzycki J, Brecht V, Müller M, Fuchs G, Müler M.** 2009. Identifying the missing steps of the autotrophic 3-hydroxypropionate CO_2 fixation cycle in *Chloroflexus aurantiacus*. Proc. Natl. Acad. Sci. U. S. A. **106:**21317-22.

2. **Hawkins AS, McTernan PM, Lian H, Kelly RM, Adams MW.** 2013. Biological conversion of carbon dioxide and hydrogen into liquid fuels and industrial chemicals. Curr. Opin. Biotechnol. **24:**1-9.

3. **Keller MW, Schut GJ, Lipscomb GL, Menon AL, Iwuchukwu IJ, Adams MWW, Leuko TT, Thorgersen MP, Nixon WJ, Hawkins AS, Kelly RM.** 2013. Exploiting microbial hyperthermophilicity to produce an industrial chemical, using hydrogen and carbon dioxide. Proc. Natl. Acad. Sci. U. S. A. **110:**5840-5.

4. **Menendez C, Bauer Z, Huber H, Gad'on N, Stetter KO, Fuchs G.** 1999. Presence of acetyl coenzyme A (CoA) carboxylase and propionyl-CoA carboxylase in autotrophic Crenarchaeota and indication for operation of a 3-hydroxypropionate cycle in autotrophic carbon fixation. J. Bacteriol. **181:**1088-98.

5. **Alber B, Olinger M, Rieder A, Kockelkorn D, Jobst B, Hugler M, Fuchs G, Hügler M.** 2006. Malonyl-coenzyme A reductase in the modified 3-hydroxypropionate cycle for autotrophic carbon fixation in archaeal *Metallosphaera* and *Sulfolobus* spp. J. Bacteriol. **188:**8551-9.

6. **Kockelkorn D, Fuchs G.** 2009. Malonic semialdehyde reductase, succinic semialdehyde reductase, and succinyl-coenzyme A reductase from *Metallosphaera sedula*: enzymes of the autotrophic 3-hydroxypropionate/4-hydroxybutyrate cycle in Sulfolobales. J. Bacteriol. **191:**6352-62.

7. **Auernik KS, Kelly RM.** 2009. Physiological versatility of the extremely thermoacidophilic archaeon *Metallosphaera sedula* supported by heterotrophy, autotrophy and mixotrophy transcriptomes. Appl. Environ. Microbiol. **76:**931-5.

8. **Chapman-Smith A, Cronan JE.** 1999. The enzymatic biotinylation of proteins: A post-translational modification of exceptional specificity. Trends Biochem. Sci. **24:**359-63.

9. **Samols D, Thornton CG, Murtif VL, Kumar GK, Haase FC, Wood HG.** 1988. Evolutionary conservation among biotin enzymes. J. Biol. Chem. **263:**6461-6464.

10. **Chapman-Smith A, Jr JEC, Cronan JE.** 1999. *In vivo* enzymatic protein biotinylation. Biomol. Eng. **16:**119-25.

11. **Mukhopadhyay B, Purwantini E, Wolfe RS.** 2001. Oxaloacetate synthesis in the *Methanarchaeon Methanosarcina* barkeri: Pyruvate carboxylase genes and a putative *Escherichia coli*-type bifunctional biotin protein ligase gene (*bpl / birA*) exhibit a unique organization. J. Bacteriol. **183:**3804-3810.

12. **Wilson KP, Shewchuk LM, Brennan RG, Otsuka AJ, Matthews BW.** 1992. *Escherichia coli* biotin holoenzyme synthetase/bio repressor crystal structure delineates the biotin- and DNA-binding domains. Proc. Natl. Acad. Sci. U. S. A. **89:**9257-9261.

13. **Noble MEM, Musacchio A, Saraste M, Courtneidge SA, Wierenga, RK.** 1993. Crystal structure of the SH3 domain in human Fyn; comparison of the three-dimensional structures of SH3 domains in tyrosine kinases and spectrin. EMBO J. **12:**2617-2614.

14. **Chapman-Smith A, Mulhern TD, Whelan F, Cronan JE, Wallace JC.** 2001. The C-terminal domain of biotin protein ligase from *E. coli* is required for catalytic activity. Protein Sci. **10:** 2608-2617.

15. **Beckett D**. 2005. The *Escherichia coli* biotin regulatory system: a transcriptional switch. J. Nutr. Biochem. **16:**411-415.

16. **Cronan JE Jr.** 1989. The *E. coli* bio operon: transcriptional repression by an essential protein modification enzyme. Cell **58:**427-429.

17. **Abdel-Hamid AM, Cronan JE.** 2007. Coordinate expression of the acetyl coenzyme A carboxylase genes, accB and accC, is necessary for normal regulation of biotin synthesis in *Escherichia coli*. J. Bacteriol. **189:**369-76.

18. **Bower S, Perkins J, Yocum RR, Serror P, Sorokin A, Rahaim P, Howitt CL, Prasad N, Ehrlich SD, Pero J**. 1995. Cloning and characterization of the Bacillus subtilis birA gene encoding a repressor of the biotin operon. J. Bacteriol. **177:**2572-75.

19. **Buoncristiani MR, Howard PK, Otsuka AJ**. 1986. DNA-binding and enzymatic domains of the bifunctional biotin operon repressor (BirA) of *Escherichia coli*. Gene **44:**255-261.

20. **Bagautdinov B, Kuroishi C, Sugahara M, Kunishima N**. 2005. Crystal structures of biotin protein ligase from *Pyrococcus horikoshii* OT3 and its complexes: structural basis of biotin activation. J. Mol. Biol. **353:**322–33.

21. **Weaver LH, Kwon K, Becket D, Matthews BW.** 2001. Competing protein:protein interactions are proposed to control the biological switch of the *E. coli* biotin repressor. Protein Sci. **10:** 2618-2622.

22. **Lokanath NK, Shiromizu I, Ohshima N, Nodake Y, Sugahara M, Yokoyama S, Kuramitsu S, Miyano M, Kunishima N.** 2004. Structure of aldolase from *Thermus thermophilus* HB8 showing the contribution of oligomeric state to thermostability. Acta. crystallographica. Section D, Biological crystallography **60:**1816-1823.

23. **Cronan JE Jr, Wallace JC.** 1995. The gene encoding the biotin-apoprotein ligase of *Saccharomyces cerevisiae*. FEMS Lett. **130:**221-30.

24. **Tissot G, Pepin R, Job D, Douce R, Alban C.** 1998. Purification and properties of the chloroplastic form of biotin holocarboxylase synthetase from *Arabidopsis thaliana* overexpressed in *Escherichia coli*. Eur. J. Biochem. **258:**586-596.

25. **Chapman-Smith A, Cronan JE.** 1999. Symposium: Nutrition, biochemisty and molecular biology of biotin molecular biology of biotin attachment to proteins. 1477-484.

26. **Weaver LH, Kwon K, Beckett D, Matthews BW.** 2001. Corepressor-induced organization and assembly of the biotin repressor: A model for allosteric activation of a transcriptional regulator. Proc. Natl. Acad. Sci. U. S. A. **98:**6045-6050.

27. **Beckett D, Matthews BW.** 1997. *Escherichia coli* repressor of biotin biosynthesis. Methods Enzymol. **279:**362-376.

28. **Kwon K, Beckett D.** 2000. Function of a conserved sequence motif in biotin holoenzyme synthetases. Protein Sce. **9:**1530-9.

29. **Li Y-Q, Sueda S, Kondo H, Kawarabayasi Y.** 2006. A unique biotin carboxyl carrier protein in archaeon *Sulfolobus tokodaii*. FEBS lett. **580:**1536-40.

30. **Sueda S, Li Y-Q, Kondo H, Kawarabayasi Y.** 2006. Substrate specificity of archaeon *Sulfolobus tokodaii* biotin protein ligase. Biochem. Biophys. Res. Commun. **344:**155–9.

31. **Clarke DJ, Coulson J, Baillie R, Campopiano DJ.** 2003. Biotinylation in the hyperthermophile *Aquifex aeolicus*. Isolation of a cross-linked BPL:BCCP complex. Eur. J. Biochem. **270:**1277-1287.

32. **Horton RM, Hunt HD, Ho SN, Pullen JK, Pease LR.** 1989. Engineering hybrid

genes without the use of restriction enzymes: Gene splicing by overlap extension. Gene **77:**61–68.

33. **Yoon SH, Reiss DJ, Bare JC, Tenenbaum D, Pan M, Slagel J, Moritz RL, Lim S, Hackett M, Menon AL, Adams MWW, Barnebey A, Yannone SM, Leigh JA, Baliga NS.** 2011. Parallel evolution of transcriptome architecture during genome reorganization. Genome Res. 21:1892-904.

34. **Lipscomb GL, Stirrett K, Schut GJ, Yang F, Jenney FE, Scott RA, Adams MWW,Westpheling J.** 2011. Natural competence in the hyperthermophilic archaeon *Pyrococcus furiosus* facilitates genetic manipulation: Construction of markerless deletions of genes encoding the two cytoplasmic hydrogenases. Appl. Environ. Microbiol. **77:**2232-2238.

35. **Laemmli UK.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227:**680-685.

36. **Zehr BD, Savin TJ, Hall RE.** 1989. A one-step, low background coomassie staining procedure for polyacrylamide gels. Anal. Biochem. **182:**157-159.

37. **Bradford MM.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72:**248-254.

38. **Hügler M, Krieger RS, Jahn M, Fuchs G.** 2003. Characterization of acetyl-CoA/propionyl-CoA carboxylase in *Metallosphaera sedula*. Eur. J. Biochem. **270:**736-744.

39. **Chapman-Smith A, Morris TW, Wallace JC, Cronan JE Jr.** 1999. Molecular recognition in a post-translational modification of exceptional specificity. Mutants of the biotinylated domain of acetyl-CoA carboxylase defective in recognition by biotin protein ligase. J. Biol. Chem. **274:**1449-1457.

40. **Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li WZ, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG.** 2011. Fast, scalable generation of high quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. **7:**539.

41. **Gouet P, Robert X, Courcelle E.** 2003. ESPript / ENDscript: extracting and rendering sequence and 3D information from atomic structures of proteins. Nucl. Acids Res. **31:** 3320-23.

42. **Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, Pang N,**

Forslund K, Ceric G, Clements J, Heger A, Holm L, Sonnhammer ELL, Eddy SR, Bateman A, Finn RD. 2012. The pfam protein family database. Nucl. Acids Res. **40:**D290-D30.

TABLES

Table 1. Cloning of E1 in *E. coli*			
Subunit of E1	Vector	Forward primer	Reverse primer
E1- α unit Msed_0147	pET46	5'- GAC GAC GAC AAG ATG CCA CCC TTT AGT AGA -3'	5'- GAG GAG AAG CCC GGT TAT CCC CAC ATC ACC CTA -3'
E1- β unit Msed_0148	pRSF-2	5'- GAC GAC GAC AAG ATG AAA CTG TAT AGG GTT CA -3'	5'- GAG GAG AAG CCC GGT TAC TTT ATC ACC ACT AG -3'
E1- γ unit Msed_1375	pCDF-2	5'- GAC GAC GAC AAG ATG ACT GCA ACT TTT GAA AAA C -3'	5'- GAG GAG AAG CCC GGT TAG AGG GGT ATA TTT CCA -3'

Table 2. *P. furiosus* Strains used in this study.

Strain	Parent	Genotype/Description	Source
COM1	DSM3638	Δ*pyrF*	(34)
MW112	COM1	Δ*pyrF* P_{gdh}*pyrF*-P_{pep}-BPL-CA-P_{slp}-E1αβ-his_6E1γ	This work

Table 3. Biotin quantification assay for recombinant E1 produced from *E. coli*.

Sample	MW g/mol	A_{500} H\A	A_{500} H\A\B	ΔA	m mol biotin per m mol protein
Biotinylated HRP (1mg/ml) (Positive Control)	40,000	1.02	0.78	0.14	1.64
Biotin (0.65mg/ml) (Positive Control)	244	1.07	0.15	0.81	N/A
E1-β (cloned in pET46) (16mg/ml)	20,000	1.03	0.94	-0.01	None detectable
Co-express E1 (pET46-α,pRSF-β,pCDF-γ) (2mg/ml)	Protein Mixture	1.05	0.98	-0.03	None detectable

A_{500} H\A: Absorbance (500 nm) of HABA\advidin premix.
A_{500} H\A\B: Absorbance (500 nm) of premix after adding samples.
$\Delta A = (A_{500}\ H\backslash A \times 0.9) - A_{500}\ H\backslash A\backslash B$

The positive ΔA indicates the existence of biotin in the sample, since it can displace HABA from advidin. ΔA close to zero indicates no existence of biotin in the sample.

# of m mol biotin per m mol protein can be calculated following the instructions of Pierce Biotin Quantitation Kit.

Table 4. Comparison of recombinant E2 from *S. tokodaii* and *M. sedula*						
	Spec Act U/mg	Km /µM	Molecular Weight	Subunit composition	RNA association	Ref
S. tokodaii	44 (65°C)	40 (65°C)	160	Homotrimers or tetramer	Y	(5)
M. sedula	43 (70°C)	66 (70°C)	126	Homotrimer	N	This work

Table 5. Molecular weight determination for all the SP1 enzymes in this study.					
Enzyme	**Native molecular mass (kDa)**	**Subunit molecular mass (kDa)**	**Suggested composition**	**Expression host**	**Comments**
E1	507	α 57	$(\alpha\beta\gamma)_4$	*P. furiosus*	α unit was insoluble, when expressed in *E. coli*, β unit was not biotinylated via *E. coli* biotin protein ligase
		β 18.6			
		γ 57			
E2	126	39	α_3	*E. coli*	No RNA association found as reported by Fuchs
E3	171	35	α_5	*E. coli*	No homotrimer as reported by Fuchs
Carbonic anhydrase (CA)	22	22	α_1	*E. coli*	Proposed as a new class of CA
Biotin protein ligase (BPL)	57	26	α_2	*E. coli*	Required for biotinylation of E1-β unit and normal function of E1

Table 6. ***In vitro*** **assembly of E1 to E3.**				
Reaction	**Substrate**	**3-HP formation**	**Conversion rate**	**Enzymes Utilization**
E2-E3	2.5 mM malonyl-CoA	2 mM	80%	E2 2 μg E3 40 μg
E1-E2-E3	2.5 mM acetyl-CoA	0.6 mM	24%	E1 4 μg E2 2 μg E3 40 μg

FIGURES

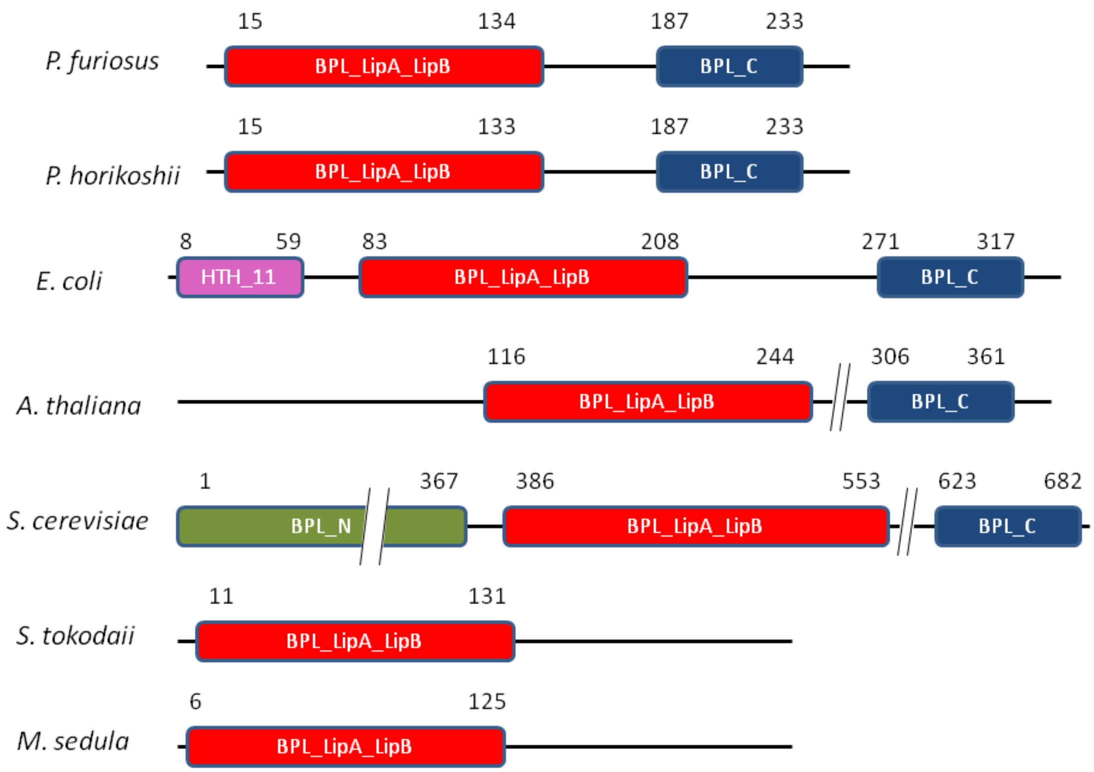


Figure 1. Domain architecture of different classes of biotin protein ligase. The domain structures were obtained from pfam (42). **HTH**:helix-turn-helix domain. **BPL_LiA_LiB**: includes biotin protein ligase, lipoate-protein ligase A and B. **BPL_C.** C-terminus domain.

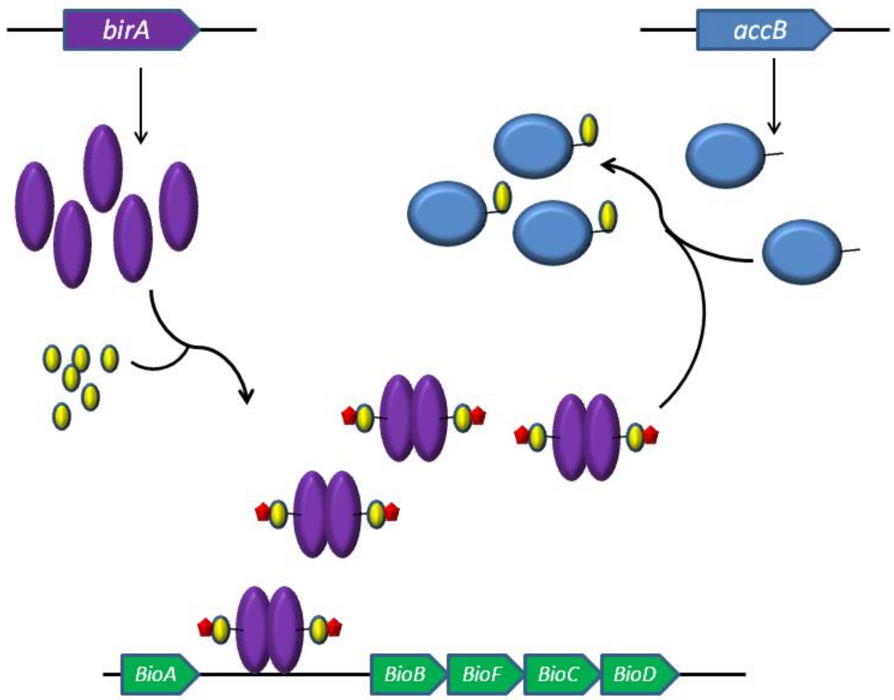


Figure 2. BirA repressor function. BirA can switch from the biotin ligation function to the repressor function in response to the intracellular biotin supply and biotin acceptor protein (e.g. apo-BCCP) levels. BirA is represented by purple ovals ( ), biotin is represented by yellow circles ( ), the AMP moiety is represented by red pentagons ( ), BCCP is encoded by accB and represented by dark blue ovals ( ). The arrows indicate transcription from the leftward and rightward bio promoters (Adapted from (17)).

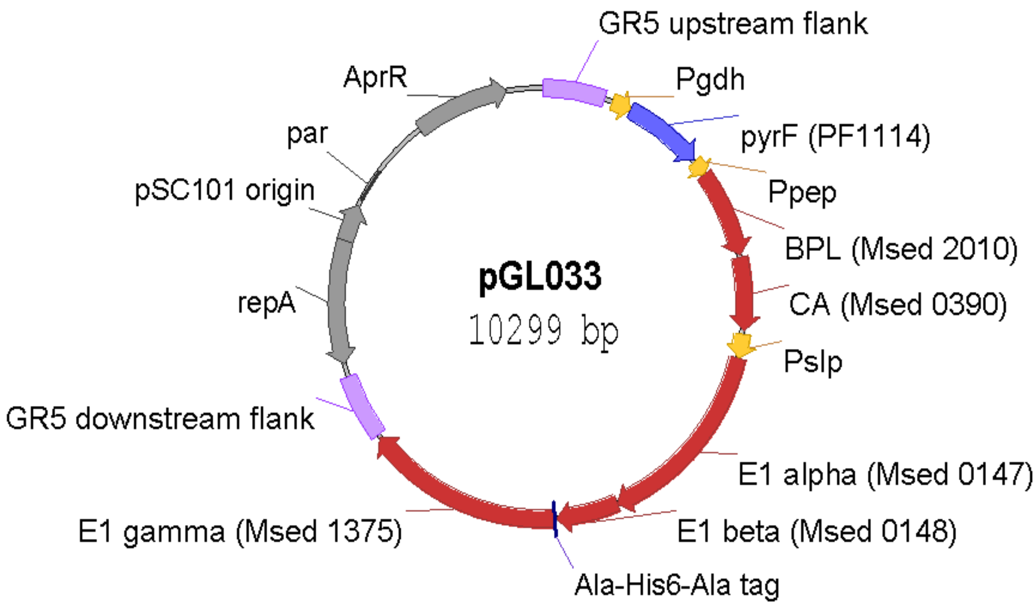


Figure 3. Plasmids used for chromosomal insertion of synthetic operons into *P. furiosus* COM1 to construct MW112 for recombinant E1 production. Plasmid elements are color-coded as follows: homologous recombination regions for *P. furiosus* chromosomal insertion, lavender; *P. furiosus* promoter regions, yellow; *P. furiosus pyrF* genetic marker, blue; *M. sedula* genes, red; his-tag, dark blue; *E. coli* plasmid maintenance elements, grey.

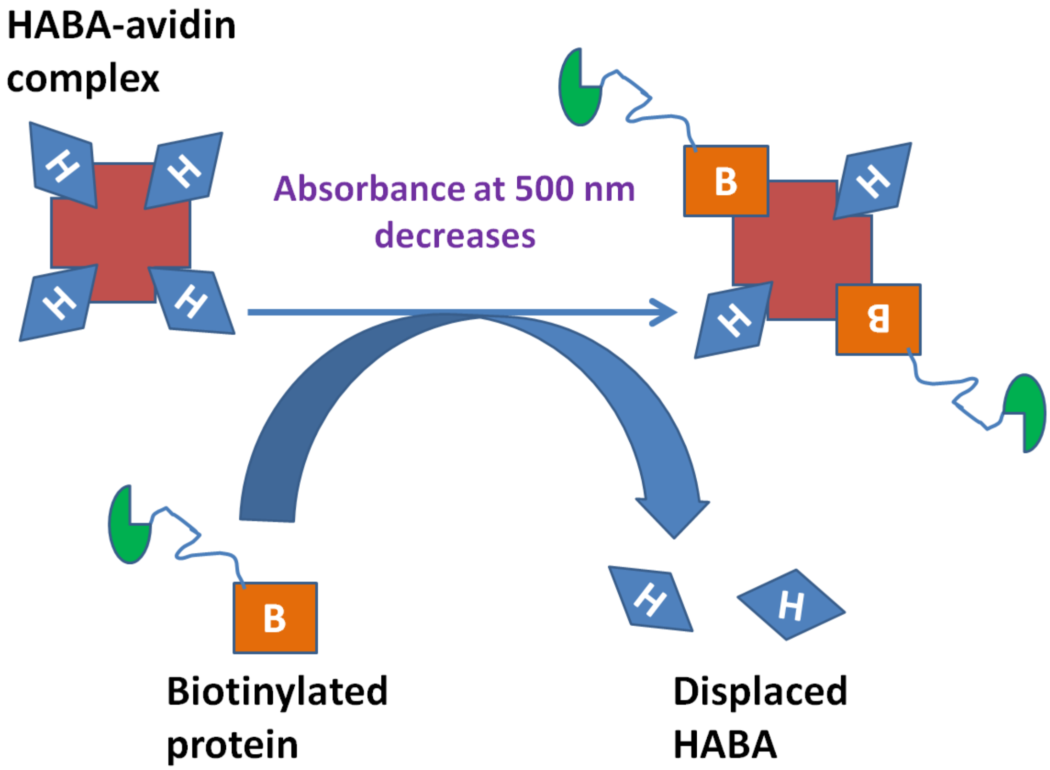


Figure 4. Chemical basis of the HABA biotin assay. (Adapted from http://www.piercenet.com/product/pierce-biotin-quantitation-kit)

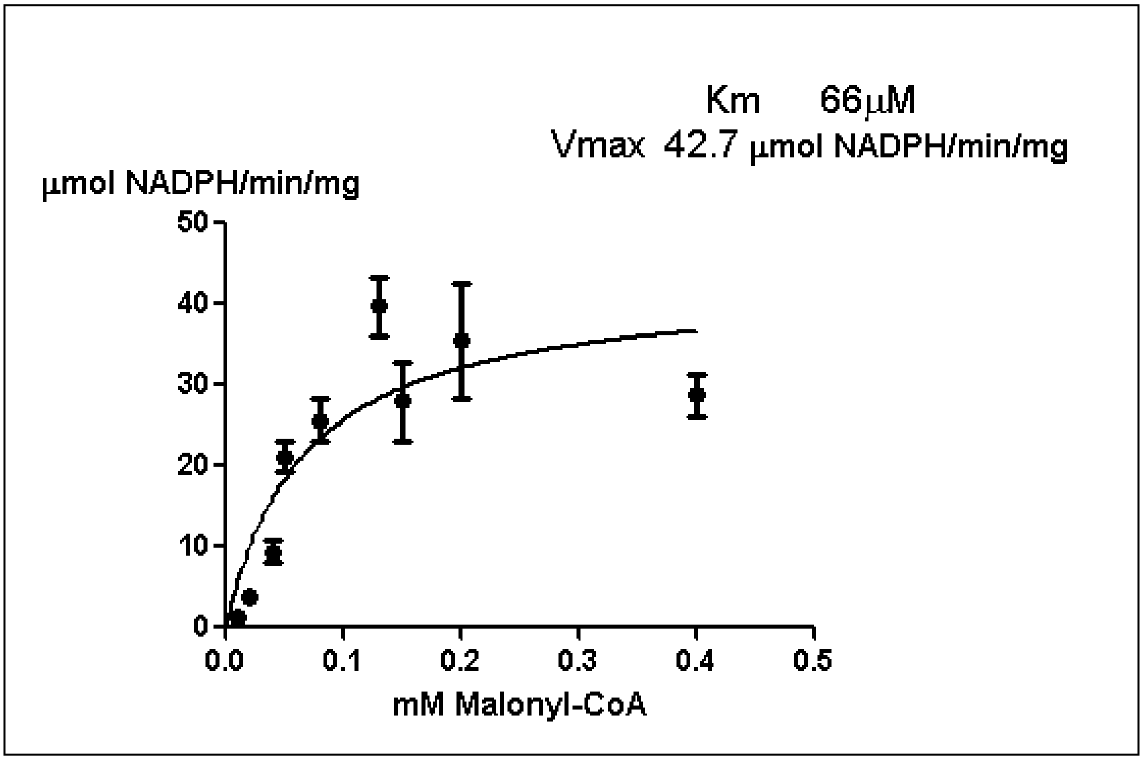


Figure 5. Kinetics for malonyl-CoA reductase (E2).

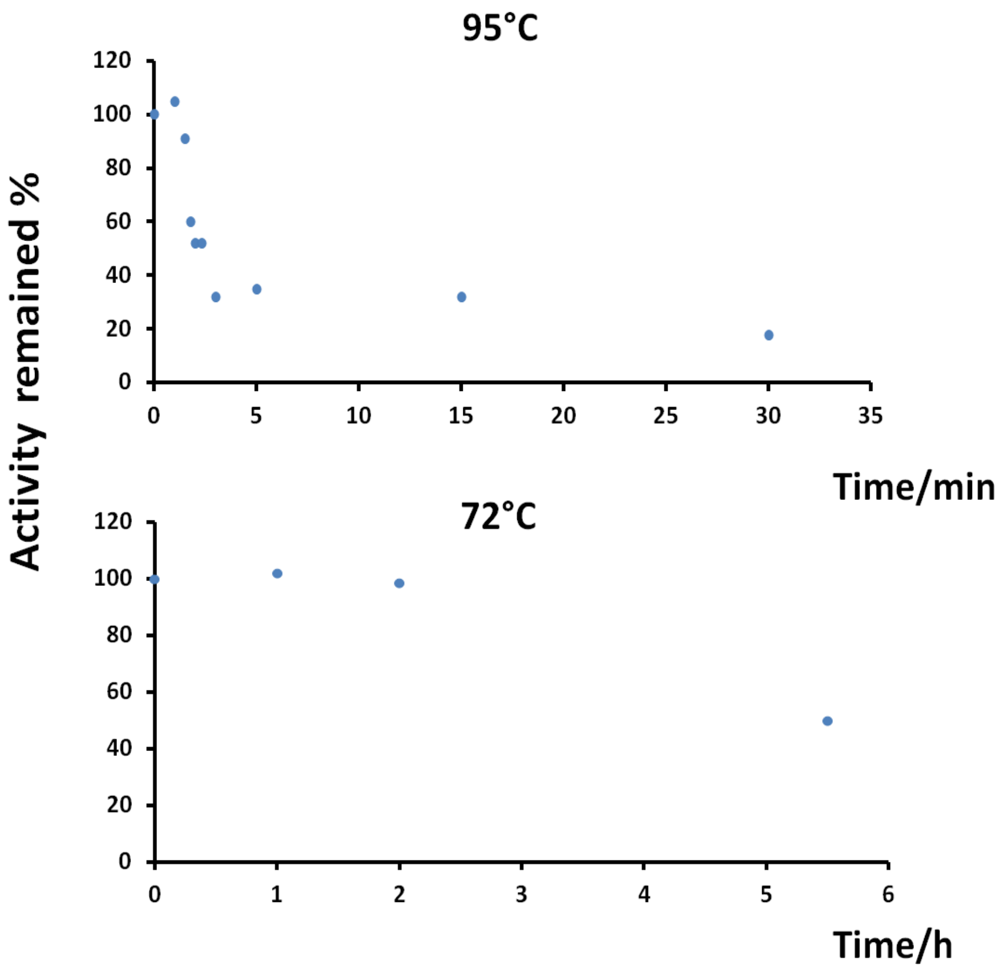


Figure 6. Thermostability of Malonyl-CoA reductase (E2).

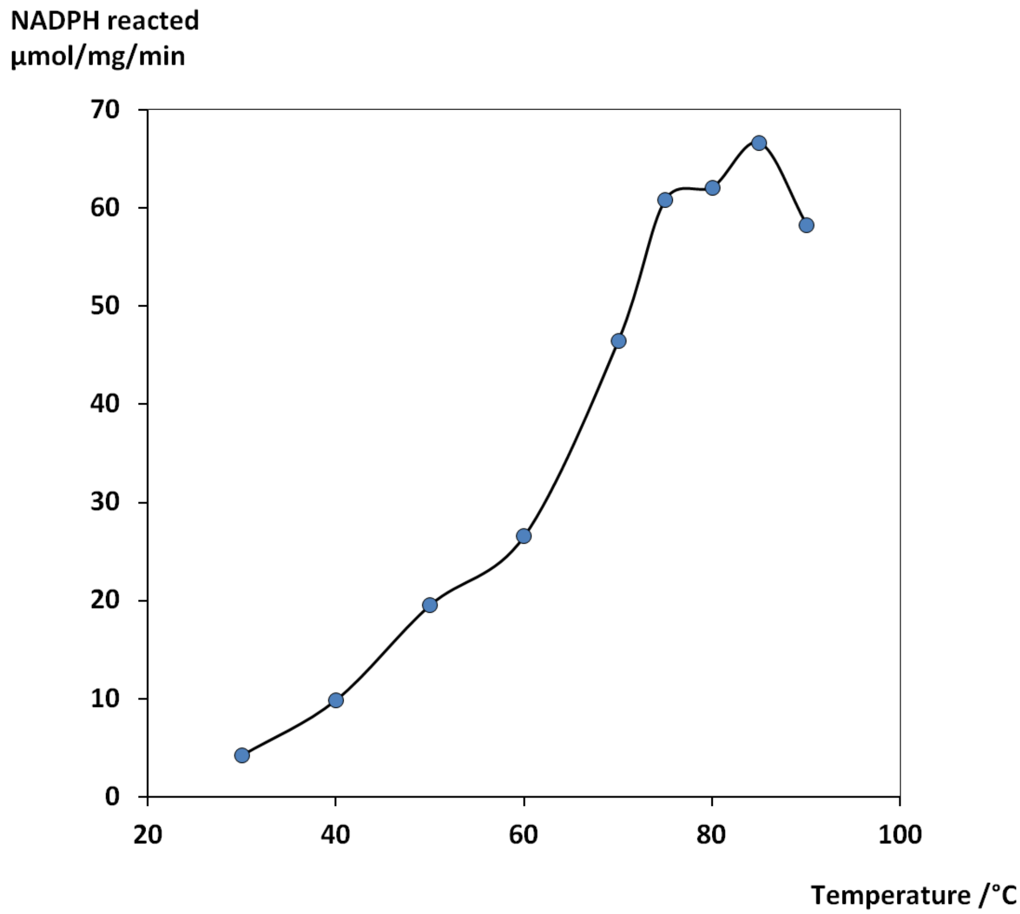


Figure 7. Optimal temperature for Malonyl-CoA reductase (E2).

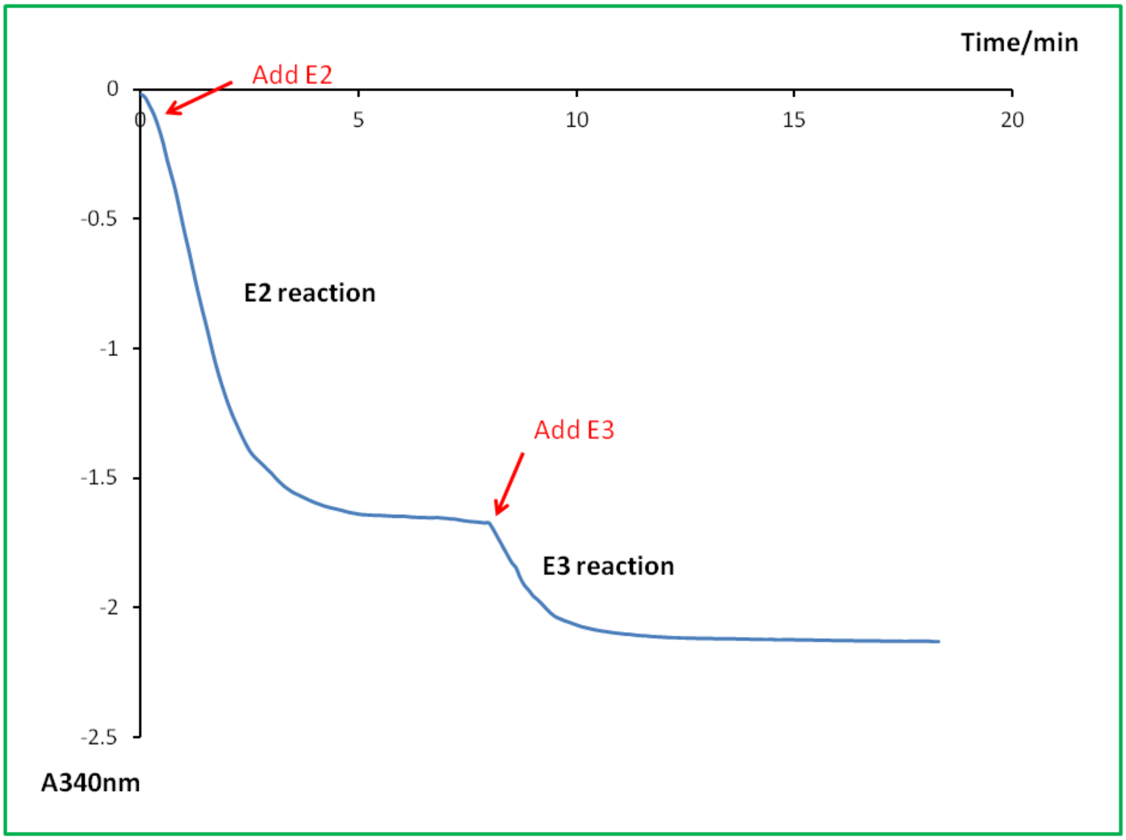


Figure 8. E2 and E3 activity detected from an E2-E3 coupled assay. E2 was added at time zero to substrate (malonyl-CoA) and the reaction was monitored by the decrease of NADPH (represented by the first slope). At close to 10 min, the E2 reaction reached equilibrium and E3 was added to react with E2 reaction product (malonic semialdehyede), and then NADPH decreased further (represented by the second slope).

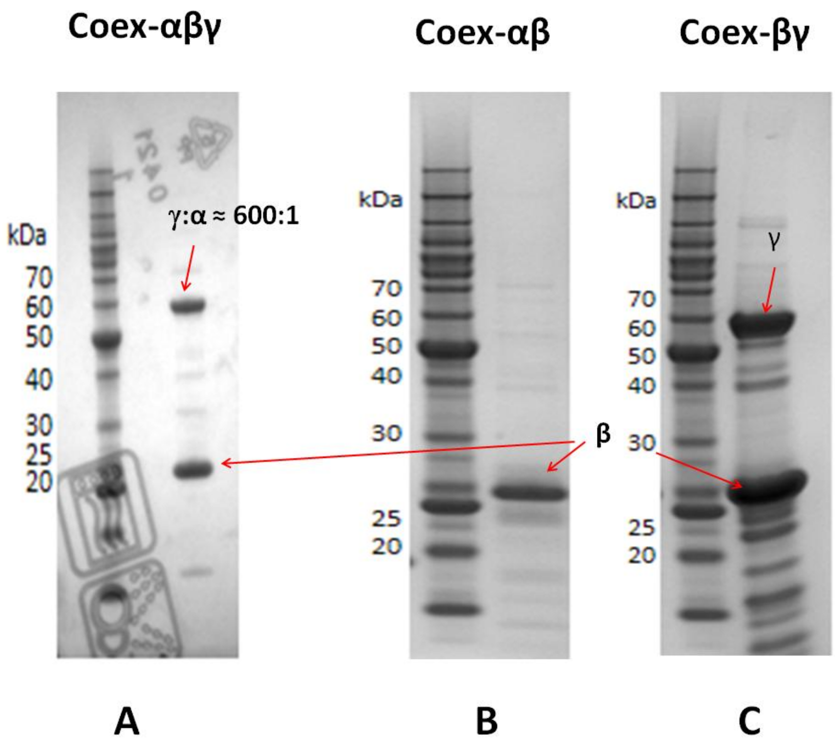


Figure 9. SDS-PAGE for E1 subunits heterologously expressed in *E. coli* and purified by IMAC. (A) αβγ-co-expressed (Msed_0147, 0148, 1375). (B) αβ-coexpressed (Msed_0147, 0148). (C) βγ-co-expressed (Msed_1375). The two bands from αβγ-co-expressed SDS-PAGE were analyzed by using nano-LC-MS/MS: the 20 kDa band was proved to be Msed_0148, and the 57kDa band was composed of α and γ with ratio of ~1:600.

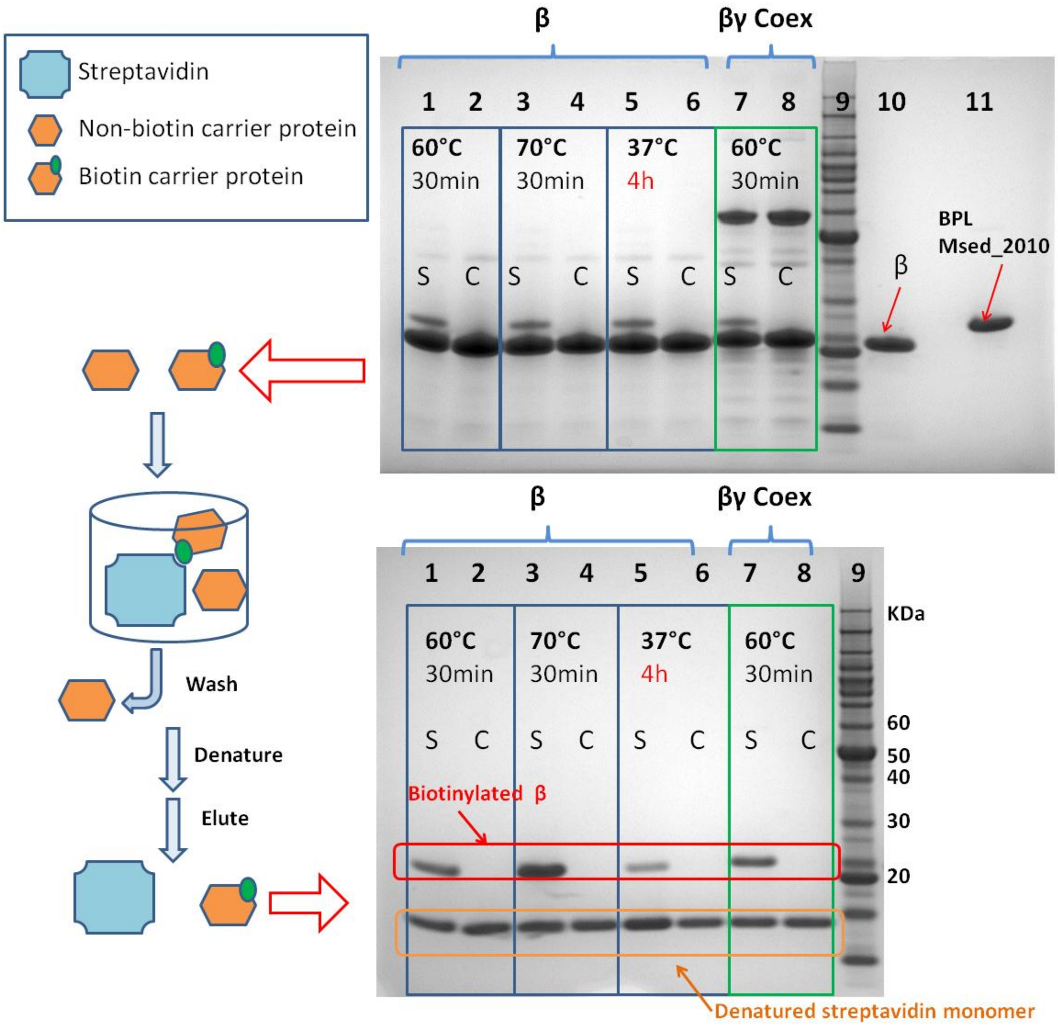


Figure 10. Identification of biotin protein ligase from *M. sedula*. Lanes of SDS-PAGE on top: #1, 3, 5. Mixture of E1-β and Msed BPL incubated at different temperature. #7. Mixture of E1-βγ and Msed BPL. #2, 4, 6. Control for #1, 3, 5 (not Msed BPL). #8. Control for #7 (Not Msed BPL). #10. E1-β. #11. Msed BPL. Lanes of SDS-PAGE on bottom: Elutes from streptavidin beads capture for mixture of E1-β/ E1-βγ and Msed BPL or control (corresponding to the top lanes).

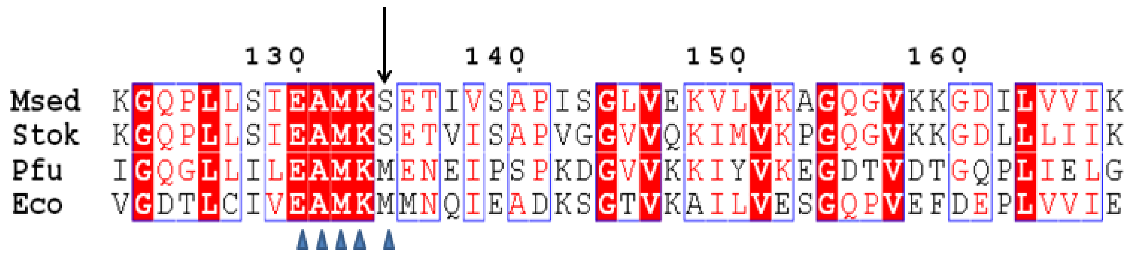


Figure 11. Alignment of partial amino acid sequences of BCCP around the canonical lysine residue (K). The conserved -E-X-M-K-M- motif is indicated by triangle below sequences. The mutated site is marked by an arrow. **Msed,** *Metallosphaera. sedula*, **Stok**, *Sulfolobus tokodaii*, **Pfu,** *Pyrococcus furiosus*, **Eco**, *Escherichia coli*. Alignment was performed by using Clustal Ω (40), and displayed with ESPript 3.0 (41).

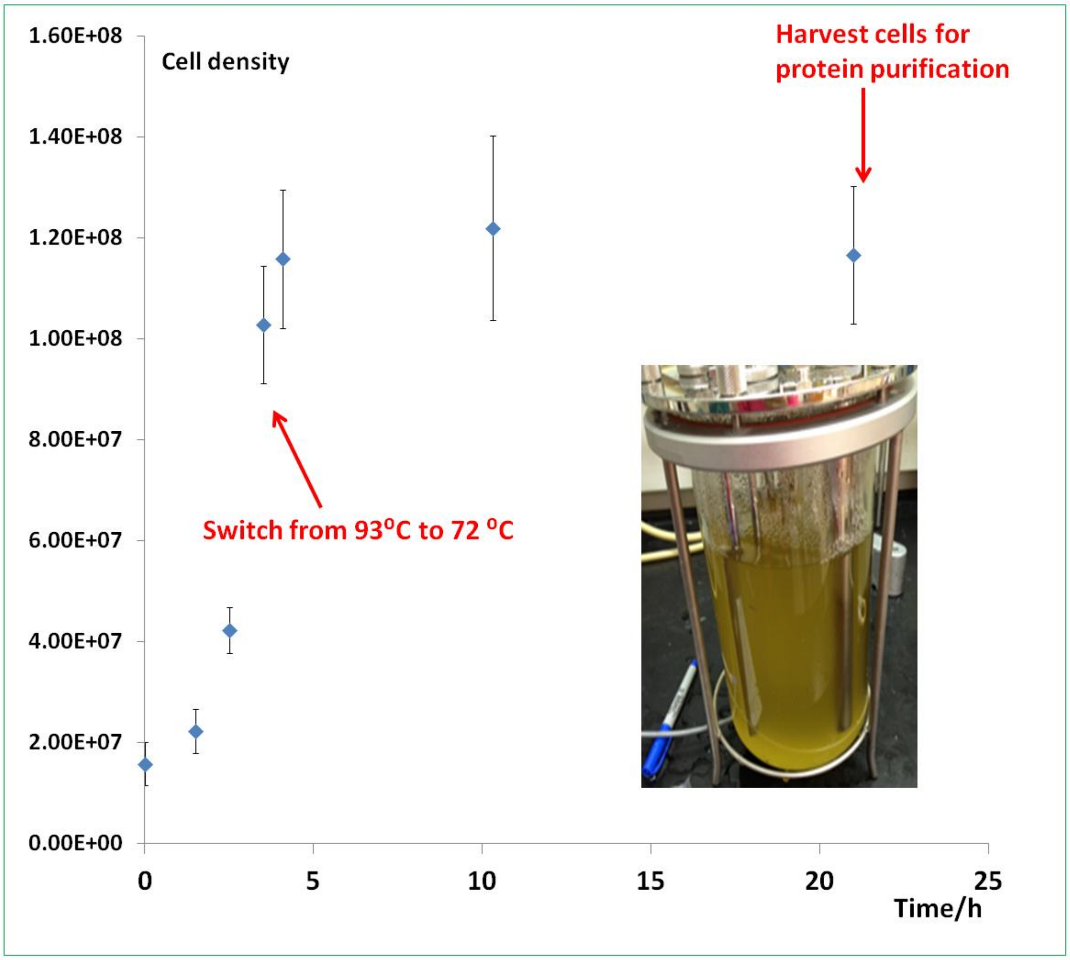


Figure 12. ***Pfu*** **MW 112 growing curve for recombinant E1 production.**

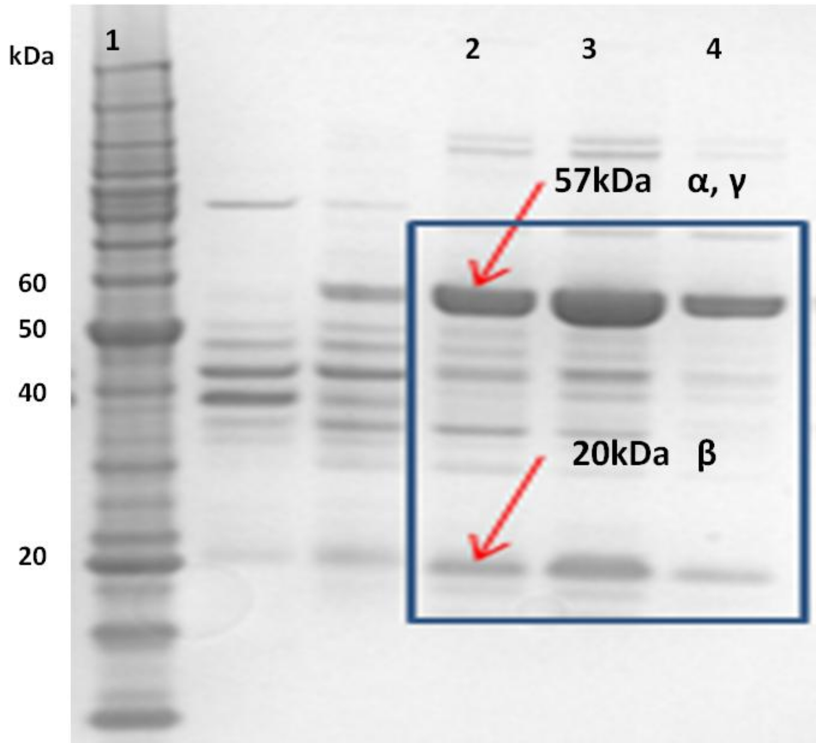


Figure 13. SDS-PAGE for IMAC purification of recombinant E1 from *Pfu* MW112. Subunits α, β, and γ are indicated by arrows. Lanes #1. Bench marker. #2-4. E1 fractions purified from IMAC.

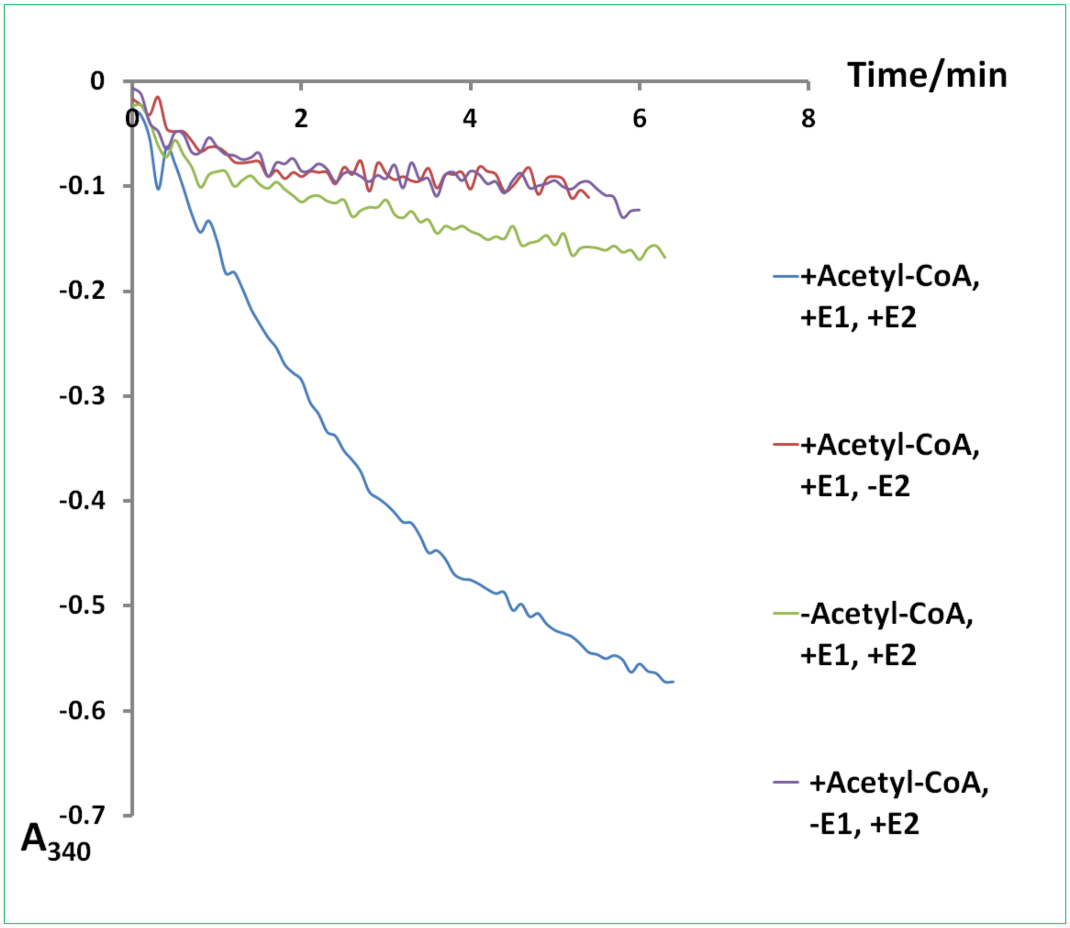
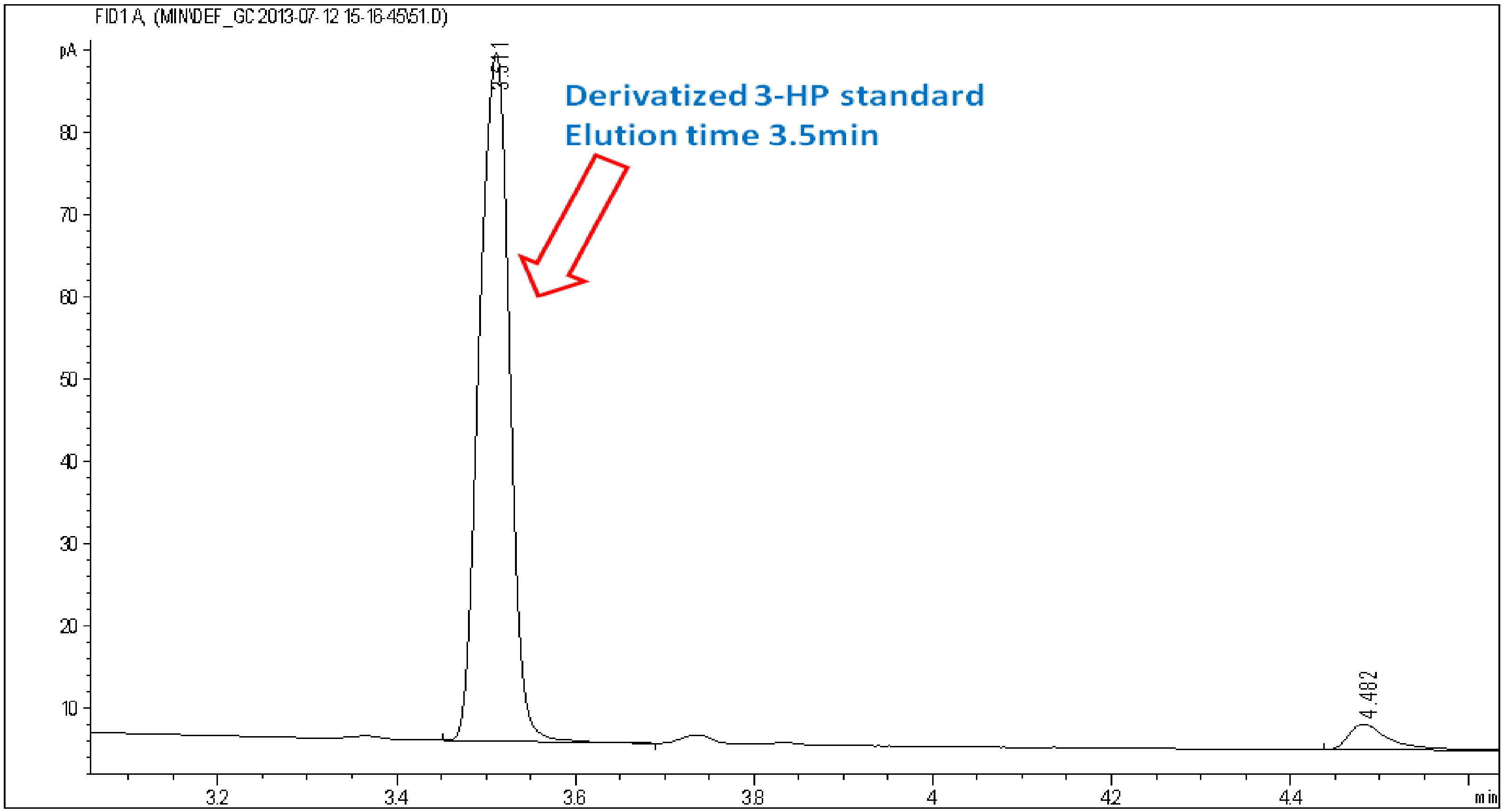


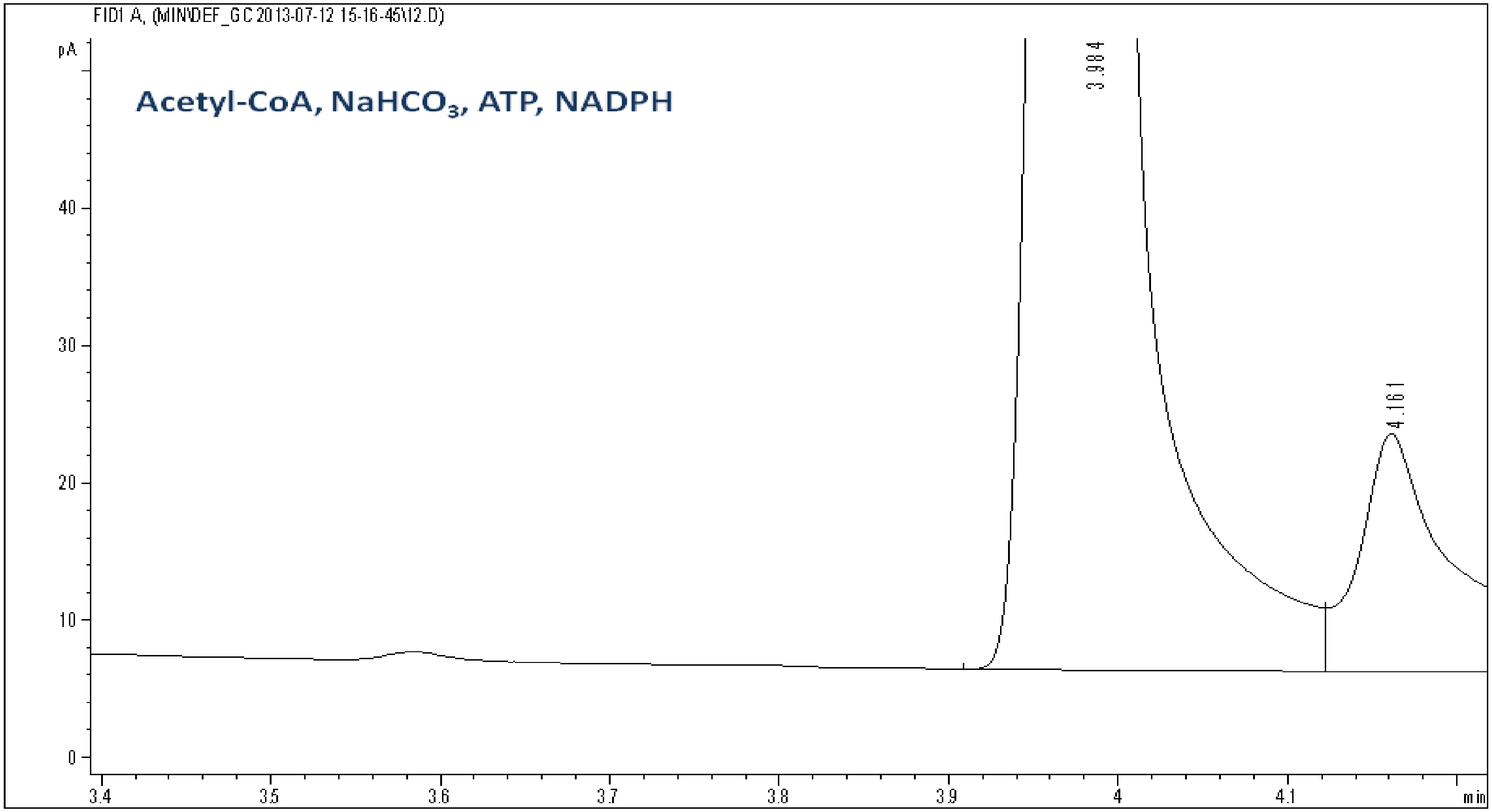
Figure 14. Recombinant E1 activity (produced from pfu strain MW 112 and IMAC purified) was confirmed in an E1-E2 coupled assay. In the reaction, both E1 and E2 was added into substrate mixture (acetyl-CoA, ATP and $NaHCO_3$). The decrease of the absorbance at 340 nm represents the consumption of NADPH by E2 reaction and the formation of malonyl-CoA by E1 reaction. Three controls were performed including adding E1 or E2 only to substrate mixture, or displacing acetyl-CoA from the substrate mixture. All the controls fail to show the NADPH consumption.

Figure 15. Demonstration of 3-HP formation from *In vitro* assembly of SP1 based on GC method after derivatization with BSTFA. (A) 3-HP standard chromatograph based on GC method after derivatization with BSTFA. (B) E1-E2-E3 *in vitro* assembly substrate mixture (No enzymes added) chromatograph based on GC method after derivatization with BSTFA. (C) E1-E2-E3 *in vitro* assembly reaction products chromatograph based on GC method after derivatization with BSTFA. (D) E2-E3 *in vitro* assembly substrate mixture (No enzymes added) chromatograph based on GC method after derivatization with BSTFA. (E) E2-E3 *in vitro* assembly reaction products chromatograph based on GC method after derivatization with BSTFA.

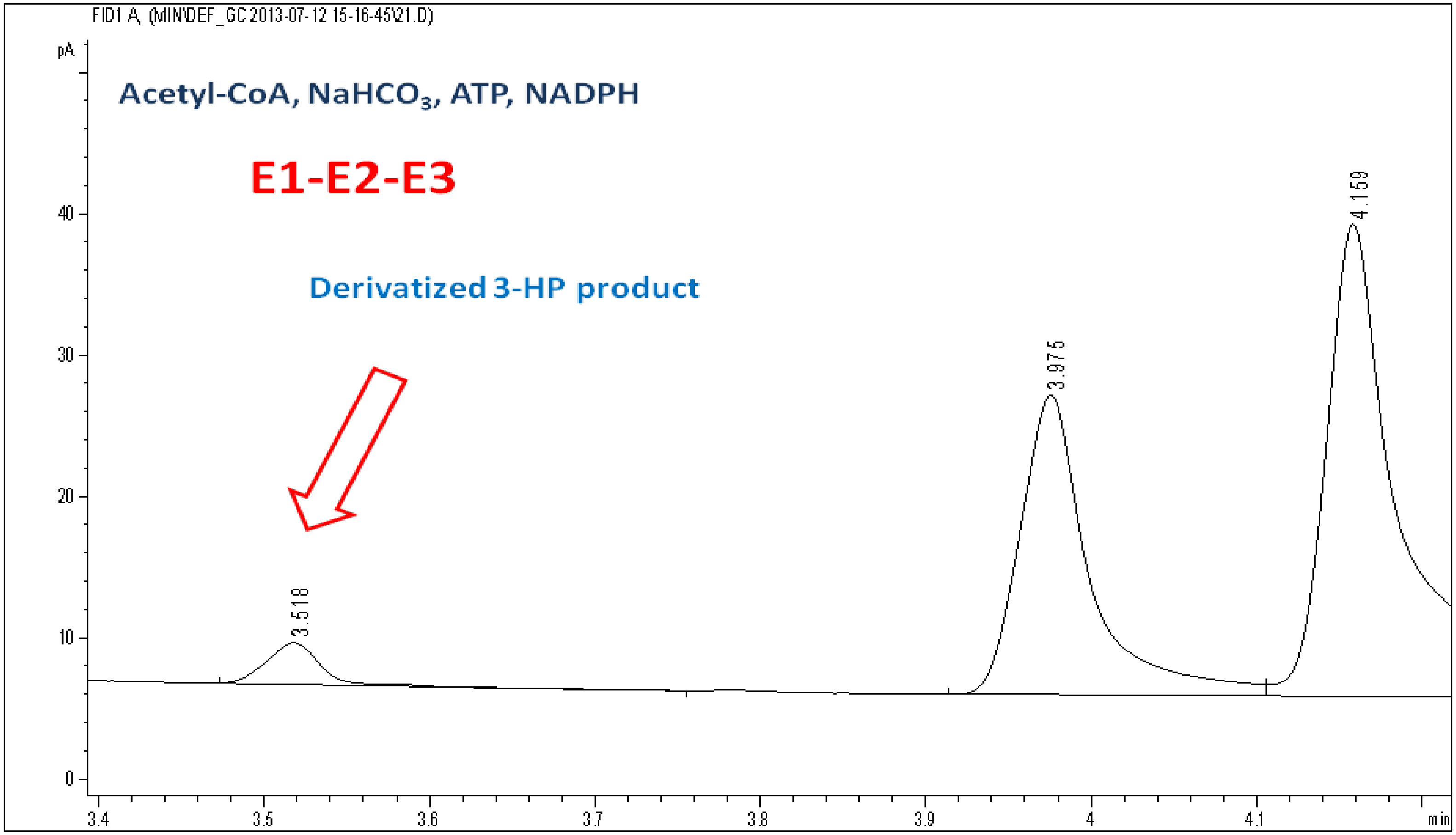
A


FID1 A, (MIN\DEF_GC 2013-07-12 15-16-45\51.D)
pA
Derivatized 3-HP standard
Elution time 3.5min
3.511
4.482
80
70
60
50
40
30
20
10
3.2
3.4
3.6
3.8
4
4.2
4.4
min

B

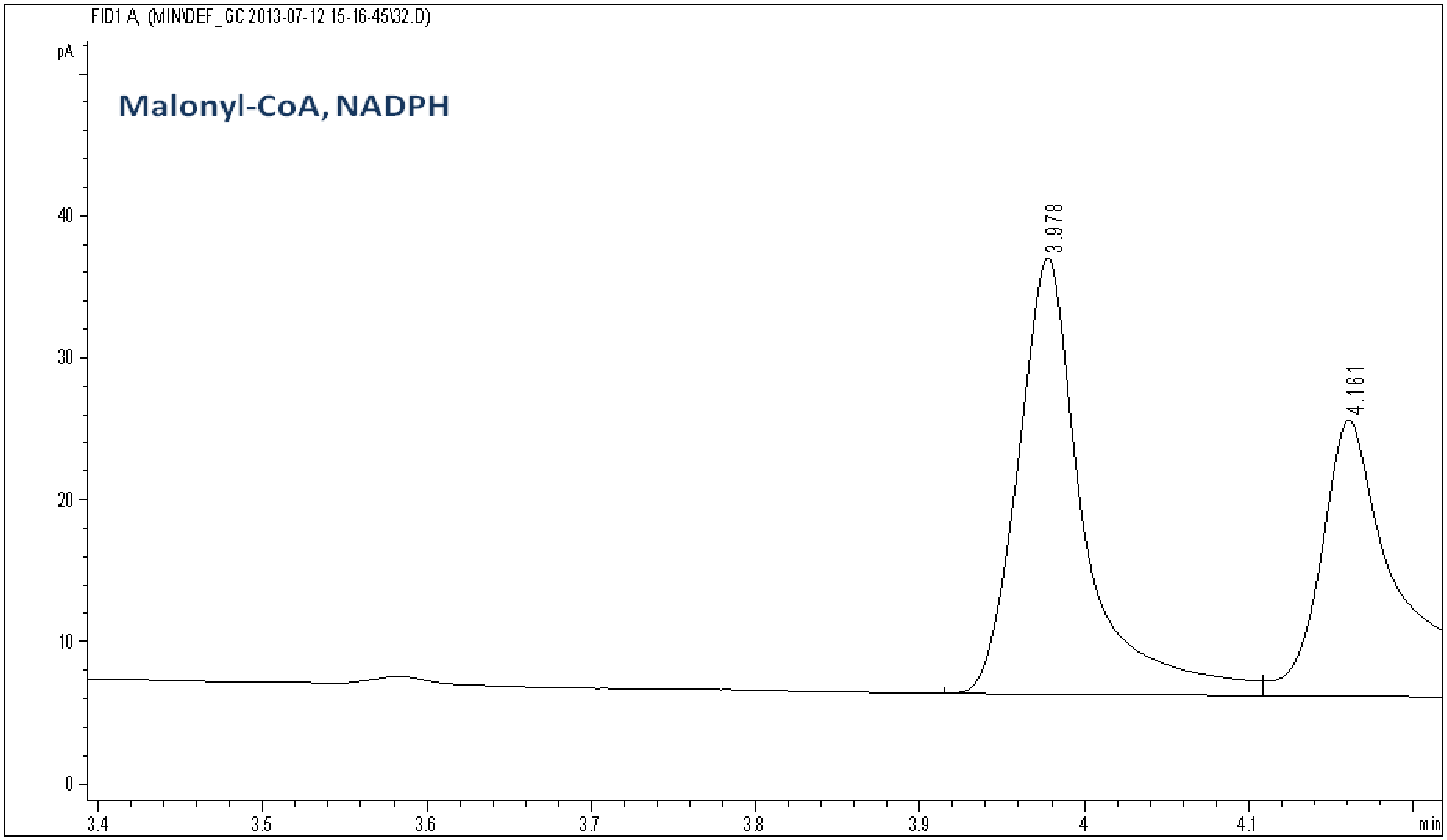

FID1 A, (MIN\DEF_GC 2013-07-12 15-16-45\12.D)
pA
Acetyl-CoA, NaHCO3, ATP, NADPH
3.984
4.161
40
30
20
10
0
3.4
3.5
3.6
3.7
3.8
3.9
4
4.1
min

C

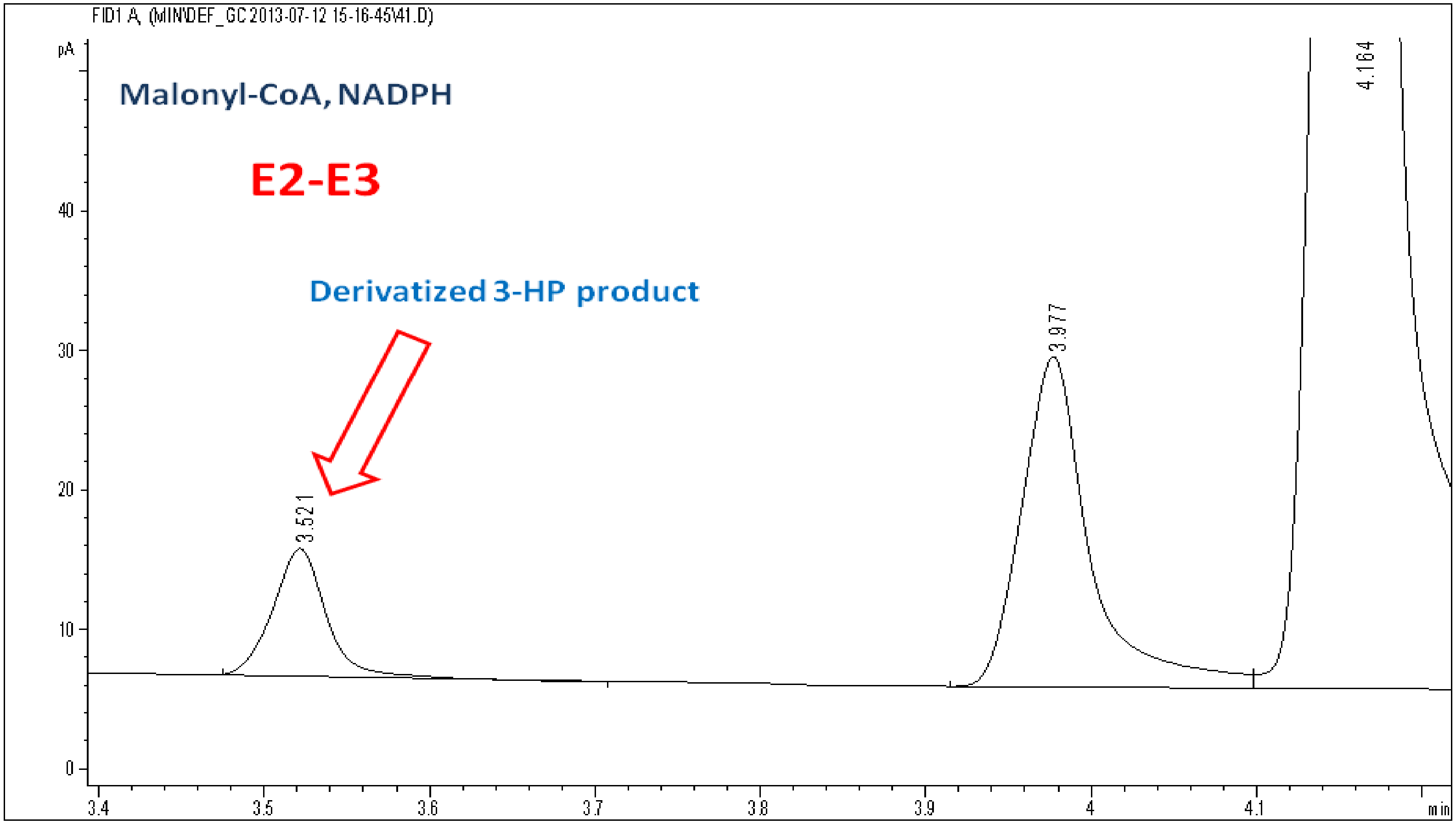


FID1 A, (MIN\DEF_GC 2013-07-12 15-16-45\21.D)
pA
Acetyl-CoA, NaHCO3, ATP, NADPH
E1-E2-E3
Derivatized 3-HP product
3.518
3.975
4.159
40
30
20
10
0
3.4
3.5
3.6
3.7
3.8
3.9
4
4.1
min

D


FID1 A, (MIN\DEF_GC 2013-07-12 15-16-45\32.D)
pA
Malonyl-CoA, NADPH
40
30
20
10
0
3.978
4.161
3.4
3.5
3.6
3.7
3.8
3.9
4
4.1
min

E



FID1 A, (MIN\DEF_GC 2013-07-12 15-16-45\41.D)
pA
Malonyl-CoA, NADPH
E2-E3
Derivatized 3-HP product
3.521
3.977
4.164
40
30
20
10
0
3.4
3.5
3.6
3.7
3.8
3.9
4
4.1
min

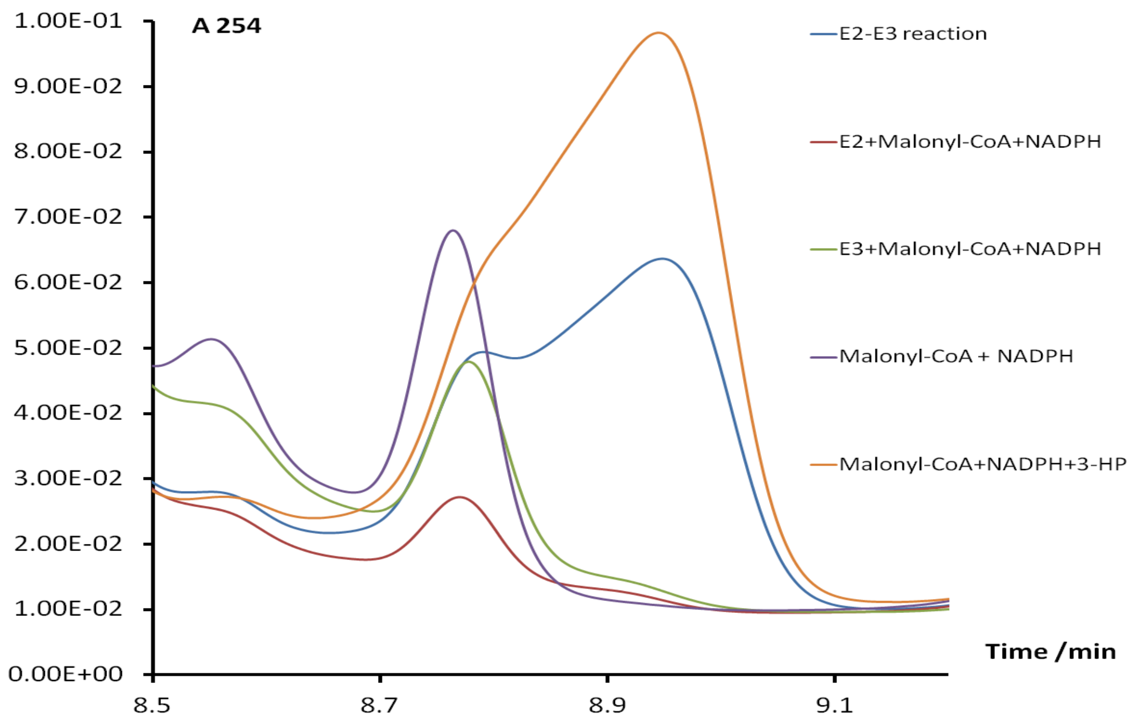


Figure 16. HPLC chromatograph for 3-HP formation via E2-E3 reaction *in vitro*

Chapter 4

Transcriptomic and Physiological Analysis of a Recombinant *Pyrococcus furiosus* Strain Metabolically Engineered to Produce 3-Hydroxypropionate from CO_2 and Maltose

Hong Lian[1], Aaron S. Hawkins[1], Andrew J. Loder[1], Arpan Mukherjee[1], Benjamin Zeldes[1], Declan Nishiyuma[1], Gina Lipscomb[2], Michael W.W. Adams[2], and Robert M. Kelly[1]

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

[2] Dept. of Biochemistry and Molecular Biology,
Univ. of Georgia, Athens, GA 30602

To be submitted to: *Metabolic Engineering* (May, 2014)

ABSTRACT

Acetyl-CoA carboxylase (ACC), malonyl-CoA reductase (MCR) and malonic semialdehyde reductase (MSR) from the extreme thermoacidophile *Metallosphaera sedula* (T_{opt} 73°C) were metabolically engineered into the hyperthermophilic archaeon *Pyrococcus furiosus* (Topt 98-100°C) to enable production of 3-hydroxypropionate (3-HP) from maltose and CO_2. *P. furiosus* was grown in a 2-liter fermentor at 95°C until late exponential phase, at which time the temperature was reduced to 72°C to initiate 3-HP production. The addition of genes encoding *M. sedula* carbonic anhydrase (CA) and biotin protein ligase (BPL), the latter of which biotinylates the subunit of ACC, led to two-fold higher 3-HP titers. Furthermore, CO_2 mass transfer was found to be rate-limiting, as 10-fold higher 3-HP concentrations (from 0.3 to 3 mM) were obtained when sparging and agitation rates were increased from 15 mL/min to 50 mL/min, and 250 rpm to 400 rpm, respectively. Transcriptomic analysis using a hybrid oligonucleotide microarray revealed that the *M. sedula* genes were among the most highly transcribed in the recombinant *P. furiosus* genome, both at 2.5h and 40h after the temperature shift. Transcriptomes of both strains do not show significant differences at 2.5 h, but were considerably different at 40 h. Of note, genes encoding monovalent cation/H+ antiporter subunits were up-regulated in MW76 compared to COM1 at 40 h. This may reflect an active process by MW76 to export 3-HP to maintain cytosolic homeostasis, in the face of intracellular acidification. From 2.5 h to 40h, COM1 exhibited significant down-regulation of genes involved in energy metabolism, amino acid biosynthesis, DNA metabolism and protein synthesis, while MW76 maintained relatively high levels of transcriptional activity. Synthesis

of biotin in MW76 was induced at 40 h, indicating that biotin was required for biotinylation of *M. sedula* ACC. The information here will be critical in designing further genetic refinements to the MW76 strain and to coordinate these changes to bioprocessing strategies.

Introduction

The list of potential metabolic engineering hosts has expanded in recent years as molecular genetics tools have become available for non-model microorganisms (1). Thus, it is no longer necessary to limit efforts to model species, such as *Escherichia coli* and *Saccharamyces cerevisiae*, especially if there are strategic opportunities that relate to the host's physiology that can be exploited to produce fuels and chemicals of interest. The development of genetic systems for several extremophiles is especially interesting, since the intrinsic physiological features of these microorganisms can be advantageous for industrial bioprocessing (2). This is especially true for thermophiles, given their reduced sensitivity to elevated temperatures and the consequential relaxation on equipment specifications and operating procedures associated with concerns related to contamination (3).

The development of genetic tools for extremophiles, such as the hyperthermophilic archaeon *Pyrococcus furiosus* (T_{opt} 98-100°C) (4), has created novel bioprocessing strategies that relate to temperature regulation. By engineering in genes encoding enzymes with significantly lower temperature optima than the host, heterologous pathways and cycles can function in the background of reduced host metabolic activity, potentially improving the bioenergetics efficiency (5). This prospect has been demonstrated for *P. furiosus* and opens up a range of possible biotechnological opportunities that exploit temperature (5,6).

The production of industrial chemicals and fuels in microbial hosts has been a central focus on metabolic engineering for nearly two decades (7). For the most part, these efforts have been based on sugars as the carbon and energy feedstock. Recently, however, CO_2 has been considered as a direct way to avoid the inherent bioenergetics limitations of photosynthetic processes and to potentially reduce greenhouse gases by recycling them into useful products (8). In fact, *P. furiosus* MW56 has been shown to produce 3-hydroxypropionate (3-HP) at approximately 0.3 mM from maltose and CO_2, when this strain was engineered to contain three enzymes (Acetyl-CoA carboxylase (ACC), malonyl-CoA reductase (MCR) and malonic semialdehyde reductase (MSR)) from the *Metallosphaera sedula* 3-hydroxypropionate/4-hydroxybutyrate (3-HP/4-HB) CO_2 fixation cycle (6). Maltose was the source of acetyl-CoA, which can accept CO_2 catalyzed by ACC, such that the carbon flux can be re-directed to 3-HP. The optimization of pathway design and fermentation process control have been effective tools to improve 3-HP production in other expression hosts such as *Escherichia coli* and *Klebsiella pneumoniae* (9). As described in Chapter 2 and 3, a carbonic anhydrase (Msed_0390) and biotin protein ligase (Msed_2010) were identified in the *M. sedula* genome, and these enzymes are expected to facilitate the delivery of inorganic carbon as substrate to ACC, and post-transcriptionally modify the β-subunit (i.e., biotinylation) of ACC to improve 3-HP production by engineered *P. furiosus* strains.

Here, we report efforts to improve 3-HP production by *P. furiosus* by inserting genes encoding carbonic anhydrase and biotin protein ligase from *M. sedula* in the hyperthermophilic host to improve CO_2 fixation and produce higher levels of ACC. This strain was then evaluated to examine host physiology under bioprocessing conditions, aided by genome-wide transcriptional response analysis. This study supports the prospect of *P. furiosus* as a newly-emerging hyperthermophilic model organism for liquid fuel or chemical production.

Materials and Methods

Plasmid construction. PCR products of the promoter of the *P. furiosus* PEP synthase gene (Ppep, consisting of a 123 bp DNA fragment immediately upstream of PF0043), the *M. sedula* biotin protein ligase (BPL) gene (Msed_2010), and the *M. sedula* carbonic anhydrase (CA) gene (Msed_0390) were combined into a synthetic operon via splice-overlap extension (SOE) PCR (10). As part of the artificial operon design, a 14-bp sequence containing the S-layer gene ribosome-binding site (RBS) was inserted (via primers) in between the BPL and CA genes (5'-ggaggtggagaaaa). The Ppep-BL-CA construct was cloned into pGL010 (6) at the SphI and AscI sites to construct plasmid pGL021 (see **Figure 1**) for insertion of the Ppep-BL-CA and Pslp-E1αβγ-E2-E3 artificial operons at genome region 3 in *P. furiosus* COM1 (between convergent genes PF0574 and PF0575). The *P. furiosus* genome regions (3 and 5) used for chromosomal integration of the constructs contains little or no transcriptional activity, as determined from analysis of tiling array data (11).

Strain construction. *P. furiosus* COM1 was transformed using uracil prototrophic selection on defined medium, essentially as previously described (12), using linearized plasmids pGL021 to construct strains MW76. (MW56 was constructed as described in (6)). Strains were further purified by two consecutive transfers on solid medium, and the final strains were verified by PCR and sequencing of the regions containing the chromosomal insertions. Strains used and constructed in this study are listed in **Table 1**. *M. sedula* ORFs carried by different strains are listed in **Table 2**.

Growth of *P. furiosus* for 3-HP production and transcriptomic analysis. The MW56, MW76 and COM1 strains were cultured, as previously described, in a seawater-based medium containing 5 g/L maltose, 5 g/L yeast extract, 5 g/L tryptone, and 0.25mg/L biotin (12). Cultures were grown in 2-L Applikon fermentors (Applikon Biotechnology, Foster City, CA) to cell late exponential phase at 95°C, sparging with 20/80 CO_2/N_2 (vol/vol) and agitated with a turbin type agitator. Once cell densities reached 1×10^8 cells/mL, the temperature was rapidly switched to 72°C (within 10 min), such that the polypeptides expressed from the *M. sedula* genes could fold properly and function near their temperature optima. Cultures were maintained for approximately 92.5 h. Two different growth conditions were used at 72°C: 1- gas feed at 15 mL/min), agitation at 250 rpm, no pH control; 2- gas feed at 50 mL/min, agitation at 400 rpm, pH controlled at 5.5-6.5 by addition of 10M NaOH. For analysis of 3-HP and acetate production and consumption of maltose, samples were periodically removed from fermentors, centrifuged (6,000 × g for 15 min), and the supernatant stored at -20°C until used.

Transcriptional response analysis. A hybrid oligonucleotide microarray was used containing 2032 unique probes for *P. furiosus* genes, along with 150 probes for heterologous genes from a variety of organisms (*M. sedula*, *Caldicellulosirupter* spp., *Thermotoga* spp., and *Thermoanaerobacter* spp.). The heterologous probes are specific to genes involved in key metabolic pathways of interest for present and future metabolic engineering work, and include genes from the 3HP/4HB cycle, butanol synthesis, carbohydrate transport and utilization, and pentose metabolism. The hybrid oligonucleotide microarray probes were designed using OligoArray 2.1 (13). Initial probe constraints were as follows: length, 58-60 base pairs; melting temperature, 82-88°C; GC content, 35-50%; pentapolymer repeats prohibited. These settings were sufficient for finding 2,111 unique probes, with the remaining probes designed by relaxing various constraints until unique probe sequences were attained.

A four-slide experimental design loop was used as shown in **Figure 2**. Tandem bioreactors operating under identical conditions were used to generate biological repeats. Samples of COM1 and MW76 (300 mL), were rapidly removed from the fermentors after incubation at 72°C for 2.5 hours and 40 hours, respectively. After chilling in a dry ice and ethanol bath at -20°C, cells were centrifuged at 4°C for 15 min (6000×g). RNA was extracted using the Qiagen RNAqueous kit (Qiagen, Valencia, CA), following the manufacturer's instructions. Equal amounts of RNA from the two biological repeats for each strain were pooled and cDNA was synthesized.

cDNA was hybridized as described previously (14). Microarray slides were scanned with the Axon 400B (Molecular Devices, Sunnyvale, CA) microarray scanner and the image data were processed with GenePix Pro 6.0 (Molecular Devices, Sunnyvale, CA) to translate to signal intensity, which was further analyzed with JMP Genomics version 5 (SAS, Cary, NC), using a mixed linear analysis of variance model as described previously (15). A ± 2.0 fold change (FC) or higher (where a log2 value of ±1 equals a 2-FC) with significance values at or above the Bonferroni correction, which was 5.5 (equivalent to a p-value of 3.0×10^{-6}, defined differential expression for this data set.)

Biochemical assays. Maltose concentrations in growth media were determined using a Sigma-Aldrich Maltose Assay Kit (Sigma-Aldrich, St. Louis, MO), following vendor protocols. 3-HP or acetate was derivatized with 2,4-Dibromoacetophenone (DBAP) as described previously with some modification (6): The sample was acidified with H_2SO_4 to pH 2 and 3-HP/acetate was extracted by adding 1.5 mL diethyl ether to 500 μL sample, centrifuging for 10 min at 6,000×g to separate the phases, removing the top ether layer, adding 50 μL of 20 mM $NaHCO_3$ to neutralize, and evaporating overnight. The dried sample was re-suspended in 50 μL acetonitrile, and the pH was adjusted to 9 to 10 by adding 10 μL of 100 mM KOH. Derivatization was started by adding 100 μL acetonitrile, 50 μL 1μM 15-crown-5 ether and 200 μL of 100 mM 2, 4-DBAP and the sample was incubated at 80°C for 60 min. The derivatized samples were analyzed by HPLC at a wavelength of 254 nm (Waters, Alliance, PDA detector). The column and run conditions were as follows: column, Shodex C-18; solvent system, A: 0.05% TFA, B: acetonitrile; gradient 0-15 min, 30-80% B, flow rate: 1 mL/min; temperature: 30°C.

Results and Discussion

Effect of recombinant *M. sedula* CA and BPL on 3-HP production in *P. furiosus*. The presence of *M. sedula* CA and BPL in MW76 led to an increase in the production of 3-HP by approximately two-fold compared with that of MW56 (which lacked CA and BPL) (see **Figure 3A**). MW56 and MW76 grew similarly to each other and to COM1, in terms of cell density profile (see **Figure 3B**). This improvement was attributed to the combined effect of CA facilitation of CO_2 uptake in MW76 and/or rapid conversion to HCO_3^-, the substrate for *M. sedula* ACC, and the biotinylation of the β-subunit of *M. sedula* ACC, thereby improving carboxylase function. Given the improved 3-HP production, MW76 was chosen for further efforts to improve product formation and to assess the physiological state of this strain during processing periods at 72°C.

3-HP production by *P. furiosus* MW76. Higher gas feed flow rates (50 mL/min vs. 15 mL/min), higher agitation rates (400 rpm vs. 250 rpm) and pH control led to 10-fold improvements in 3-HP production by *P. furiosus* MW76 (see **Figure 4**). At these conditions, MW76 and COM1 grew to similar peak cell densities at 72°C (6×10^{-8} cells/ml), having been at ~$1x10^8$ cells/ml when the temperature was shifted, produced similar amounts of acetate, and consumed essentially all maltose in the medium by ~5 h after initial inoculation and by 68 h following the temperature shift to 72°C (see **Figure 5**). Efforts to improve upon the 3 mM 3-HP levels for MW76 by increasing agitation rate (to 1,000 rpm), increasing N_2/CO_2 sparging rates (to 100 mL/min) and by adding additional maltose (to 25 g/L) were not successful. For these cases, although growth rates improved and 3-HP volumetric production increased, 3-HP titers did not increase beyond 2-3 mM (Hawkins and Kelly, unpublished data). Additional efforts are needed to determine the shear sensitivity of *P. furiosus* at high agitation/sparging rates, and whether different bioprocessing strategies (i.e., fed-batch addition of maltose) could improve 3-HP titers.

Another issue to consider is the possible toxic effect of 3-HP as it accumulates in the cytosol and growth medium. In general, the accumulation of organic acids in the cytoplasm exerts a toxic effect on microbial growth by disrupting the intracellular pH and anion pool (16). Essential enzymes and integrity of purine bases can be compromised, leading to cell death. It has been reported that for recombinant *E. coli* and *K. pneumoniae* the rate of 3-HP production decreases significantly when its concentration in the medium exceeds 200-300 mM; initial testing of *P. furiosus* at these 3-HP levels indicates a comparable effect (communication with Gerrit J. Schut, University of Georgia). Also, addition of 3-HP to the growth medium impairs the rate of cell growth, glycerol consumption and 3-HP production (17,18,19). The ability of organic acids to diffuse across the cell membrane can reduce their toxic effects to the cell, but this export process may re-direct metabolic energy from maintaining cell growth and essential cellular functions (20). It was estimated that the minimum inhibitory concentration of 3-HP was 15 g/L, which represents the smallest concentration of a chemical that prevents the growth of an organism under a set of conditions (21). As reported by Warnecke *et al.* (22), the toxic effects of 3-HP to the *E. coli* was related to inhibition of the chorismate and threonine super-pathways, and the supplement of metabolites in these pathways seem to improve 3-HP tolerance. OPX Biotechnologies has identified genes and pathways that are up-regulated in strains selected for tolerance to 3-HP, including transporters and pathways for the synthesis of the amino acids tyrosine, phenylalanine, and threonine (16, 23).

Transcriptome analysis. Transcriptome analysis was used to compare MW76 and COM1 growing with a gas feed at 50 mL/min, agitated at 400 rpm, and pH controlled at 5.5-6.5 by periodic addition of 10 M NaOH. Two different time points were sampled for this comparison: 2.5 h after temperature switching to 72°C, at which point no 3-HP production was detectable, and 40 h after switching to 72°C, when 3-HP accumulation in the culture medium was reaching over half of maximum levels obtained. The contrasts for the 4-slide loop (see **Figure 2**) are summarized in **Figure 6**, and selected ORFs with significant fold change were listed in **Table 3-7**.

(a). MW76 vs. COM1 contrast at 72°C, 2.5 h: COM1 and MW76 exhibited comparable transcriptomes at 2.5 h at 72°C, with only 23 ORFs differentially transcribed (>2-fold).

(b). MW76 vs. COM1 contrast at 72°C, 40 h: However, as the 72°C incubation extended to 40 h, the number of ORFs differentially transcribed (>2-fold) between MW76 and COM1 increased to 348, suggesting a different physiological state for the two strains. Among the 202 up-regulated ORFs (>2-fold) in MW76, 44 were involved in energy metabolism, 27 were related to protein synthesis, and seven corresponded to the *M. sedula* genes that had been metabolically engineered into the *P. furiosus* MW76 strain. The energy metabolism genes include: the V-type ATP synthase subunits I, K, C, F, B, responsible for ATP-proton motive force inter-conversion (PF0177, PF0178, PF0180, PF0181, and PF0183); phosphoenolpyruvate synthase, involved in glycolysis/gluconeogenesis (PF0043); and cytochrome-c3 hydrogenase subunit gamma (PF0892) and ferredoxin (PF1909), both of which are involved in electron transport for energy conservation. The protein synthesis genes include 30S ribosomal proteins (e.g., PF1809, PF1810, PF1644), 50S ribosomal proteins (e.g., PF1802, PF1807, PF1818, PF1820), translation initiation factors (e.g., PF1349, PF1817) and an RNA-associated protein (e.g., PF 1570).

At the 40 h time point, MW76 and COM1 had both reached ~6 x 10^8 cells/ml and 3-HP titers were ~2mM. At 40 h, genes encoding monovalent cation/H+ antiporter subunits (PF1423, PF1425, PF 1429, PF1451, and PF1449) were up-regulated in MW76 compared to COM1. This may reflect an active process by MW76 exported 3-HP to maintain cytosolic homeostasis, in the face of intracellular acidification. PF1739, the trehalose/maltose binding protein, was transcribed over 8-fold higher in MW76, possibly indicating increased carbon/energy requirements to maintain cellular functions and to direct organic carbon flux to 3-HP. Notably, 2-amino-3-ketobutyrate coenzyme A ligase (PF0265), and a hypothetical protein (PF0144), which are implicated in biotin synthesis, were up-regulated 4- to 8-fold in MW76, presumably corresponding to the extra demand for biotin, as a critical cofactor for *M. sedula* ACC for its carboxyl group transfer sub-unit (β).

(c). COM1 contrast at 72°C, 40 h vs. 2.5 h: At 40 h, the COM1 strain showed significantly reduced levels of transcription relative to 2.5 h. Genes involved in amino acid synthesis, energy metabolism, DNA metabolism, transcription, and protein synthesis were all down-regulated. Amino acids belonging to the pyruvate family, including 3-isopropylmalate dehydratase large subunit (PF1679), 3-isopropylmalate dehydratase small subunit) (PF1680), aspartate family, lysine biosynthesis protein (PF1681), acetyl-lysine deacetylase (PF1686), aspartate aminotransferase (PF1702) and glutamate family , N-acetyl-γ-glutamyl-phosphate reductase (PF1683), acetylglutamate/acetylaminoadipate kinase (PF1684) exhibited 4- to 32-fold down-regulation. The V-type ATP synthase subunits I, K, E, C, F, B (PF0177, PF0178, PF0179, PF0180, PF0181, and PF0183, respectively) were all down-regulated between 2- to 10-fold, indicating a downturn in energy generation. Genes encoding proteins involved in DNA replication, recombination and repair were also down-regulated in COM1 at 40 h, including: DNA topoisomerase VI subunit B, (PF1579), PF1881, DNA/RNA-binding protein alba (PF1881), 3-methyladenine DNA glycosylase (PF0511), DNA repair protein Rad25 (PF0126), DNA repair and recombination protein RadA (PF1926), histone a1 (PF1831). Ribosomal proteins and translation initiation factors related to protein synthesis were all down-regulated at 40 h in COM1, as were genes encoding the DNA-directed RNA polymerase subunit A, D, E, H, K (PF1562, PF1647, PF0256, PF1565, PF1642, respectively), as well as transcription factor (PF1295) and RNA 3'-terminal-phosphate cyclase (PF1549).

(d). MW76 contrast at 72°C, 40 h vs. 2.5 h: In contrast to COM1, MW 76 increased transcription for several genes involved in amino acid synthesis, including: aspartate family , threonine synthase (PF1180), acetyl-lysine deacetylase) (PF1686); serine family, sulfate adenylytransferase (PF1220), hypothetical protein (PF0211); aromatic amino acid family, anthranilate synthase component I (PF1709), anthranilate phosphoribosyltransferase (PF1710), indole-3-glycerol-phosphate synthase (PF1711). Most amino acid biosynthesis genes significantly down-regulated in COM1 (40 h vs. 2.5 h) showed little change in MW76 (40 h vs. 2.5 h). Also, the most up-regulated amino acid synthesis genes in MW76 were not differentially transcribed in COM1 during this period. The potentially high translational burden generated by the high transcription levels of the seven *M. sedula* genes at both 2.5 h and 40 h in MW76 could be responsible for higher amino acid requirements, possibly explaining the differences between MW76 and COM1 during incubation at 72°C.

(e). Other observations from the COM1 and MW76 transcriptomes. There was no transcriptomic evidence for possible inhibition of the chorismate and threonine super-pathways in MW76 (40 h vs 2.5 h). In fact, chorismate mutase (PF1701) and threonine synthase (PF1180) were up-regulated significantly. As mentioned previously, OPX Biotechnologies identified genes and pathways that are up-regulated in *E. coli* strains selected for tolerance to 3-HP, including transporters and pathways for synthesis of tyrosine, phenylalanine, and threonine (16, 23). Up-regulation of these amino acid biosynthesis genes could facilitate 3-HP tolerance in MW76.

It was clear that at 40 h energy metabolism in MW76 was not suppressed, as noted for COM1. Only 11 ORFs related to energy generation were down-regulated in MW76 (40 h vs. 2.5 h), compared with 54 ORFs in COM1. In fact, some MW76 ORFs were up-regulated more than two-fold, such as: ADP-dependent glucokinase (PF0312), phosphoenolpyruvate synthase (PF0043), and galactokinase (PF0445). Phosphoenolpyruvate synthase (PF0043) was up-regulated 3-fold (40 h vs. 2.5 h) in MW76, while this gene was down-regulated 6-fold (40 h vs. 2.5 h) in COM1. MW76 consumed maltose slightly faster than COM1, suggesting heightened metabolic levels and diverting some carbon to 3-HP (see **Figure 5**). It is also possible that MW76 requires more maltose as an energy source to meet the extra biosynthetic burden for producing *M. sedula* enzymes. It is interesting that MW76 maintained its cell density level longer than COM1, despite the burdens placed on this strain for 3-HP production and recombinant protein synthesis. The possibility that these burdens offset normal cellular processes that lead to stationary phase and cell death needs to be considered further.

Biotin synthesis appeared to increase in MW76 at 40 h, compared to 2.5 h, as indicated by a 3-fold up-regulation of a hypothetical protein (PF0144) predicted to be related to biotin biosynthesis pathways. In contrast, for this same time period, biotin synthesis in COM1 was lower, since the 2-amino-3-ketobutyrate coenzyme A ligase (PF0265) was significantly down-regulated. This difference likely relates to the biotin needed to post-translationally modify *M. sedula* ACC, as catalyzed by *M. sedula* BPL. During incubation at 72°C, the external supply of biotin in the medium likely becomes limited, so the biotin synthesis is initiated by MW76.

At last, the ORFs which were up-or down-regulated in both COM1 and MW76 (2.5h vs 40h) were listed in **Table 8**.

Conclusions. The results from this study lay the basis for further efforts to develop *P. furiosus* as a metabolic engineering strain for production of fuels and chemicals. The transcriptomic analysis provided important insights into the physiological state of this hyperthermophile during extended incubation at a substantially sub-optimal temperature (72°C) to form 3-HP via seven genes inserted from the *M. sedula* 3-HP/4-HB pathway. This information will be critical in designing further genetic refinements to the MW76 strain and to coordinate these changes to bioprocessing startegies.

REFERENCES

1. **Woolston BM, Edgar S, Stephanopoulos G.** 2013. Metabolic engineering: past and future. Annu. Rev. Chem. Biomol. Eng. **4:**259-88.

2. **Deive FJ, López E, Rodríguez A, Longo M, Sanromán MÁ.** 2012. Targeting the production of biomolecules by extremophiles at bioreactor scale. Chem. Eng. Technol. **35:**1565-1575.

3. **Turner P, Mamo G, Karlsson EN.** 2007. Potential and utilization of thermophiles and thermostable enzymes in biorefining. Microb. Cell Fact. **6:**9.

4. **Fiala G, Stetter KO.** 1986. *Pyrococcus furiosus* sp. nov. represents a novel genus of marine heterotrophic archaebacteria growing optimally at 100°C. Arch. Microbiol. **145:**56-61.

5. **Basen M, Sun J, Adams MWW**. 2012. Engineering a hyperthermophilic archaeon for temperature-dependent product formation. mBio **3:**e00053-12.

6. **Keller MW, Schut GJ, Lipscomb GL, Menon AL, Iwuchukwu IJ, Adams MWW, Leuko TT, Thorgersen MP, Nixon WJ, Hawkins AS, Kelly RM.** 2013. Exploiting microbial hyperthermophilicity to produce an industrial chemical, using hydrogen and carbon dioxide. Proc. Natl. Acad. Sci. U. S. A. **110:**5840-5.

7. **Berry DA**. 2010. Engineering organisms for industrial fuel production. Bioeng. Bugs. **1:**303-308.

8. **Conrado RJ, Haynes CA, Haendler BE, Toone EJ.** 2013. Electrofuels: A new paradigm for renewable fuels, p. 1037-64. *In* Lee JW(ed), Advanced Biofuels and Bioproducts. Springer, New York, N.Y.

9. **Kumar V, Ashok S, Park S.** 2013. Recent advances in biological production of 3-hydroxypropionic acid. Biotech. Adv. **31:**945-61.

10. **Horton RM, Hunt HD, Ho SN, Pullen JK, Pease LR.** 1989. Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. Gene **77:**61-68.

11. **Yoon SH, Reiss DJ, Bare JC, Tenenbaum D, Pan M, Slagel J, Moritz RL, Lim S, Hackett M, Menon AL, Adams MWW, Barnebey A, Yannone SM, Leigh JA, Baliga NS.** 2011. Parallel evolution of transcriptome architecture during genome reorganization. Genome Res. **21:**1892-904.

12. **Lipscomb GL, Stirrett K, Schut GJ, Yang F, Jenney FE, Scott RA, Adams MWW, Westpheling J.** 2011. Natural competence in the hyperthermophilic archaeon *Pyrococcus furiosus* facilitates genetic manipulation: construction of markerless deletions of genes encoding the two cytoplasmic hydrogenases. Appl. Environ. Microbiol.**77:**2232-2238.

13. **Rouillard JM, Zuker M, Gulari E.** 2003. OligoArray 2.0: Design of oligonucleotide probes for DNA microarrays using a thermodynamic approach. Nucleic Acids Res. **31:** 3057-3062.

14. **Chhabra SR, Shockley KR, Conners SB, Scott KL, Wolfinger RD, Kelly RM.** 2003. Carbohydrate-induced differential gene expression patterns in the hyperthermophilic bacterium Thermotoga maritima. J. Biol. Chem. **278:**7540-52.

15. **Tachdjian S, Kelly RM.** 2006. Dynamic metabolic adjustments and genome plasticity are implicated in the heat shock response of the extremely thermoacidophilic archaeon *Sulfolobus solfataricus*. J. Bacteriol. **188:**4553-4559.

16. **Warnecke T, Gill RT.** 2005. Organic acid toxicity, tolerance, and production in *Escherichia coli* biorefining applications. Microb. Cell Fact. **4:**25.

17. **Ashok S, Sankaranarayana M, Ko Y, Jae K-E, Ainala SK, Kumar V, Park S.** 2013. Production of 3-hydroxypropionic acid from glycerol by recombinant *Klebsiella pneumoniae* ΔdhaTΔyqhD which can produce vitamin B12 naturally. Biotechnol. Bioeng. **110:**511-24.

18. **Rathnasingh C, Raj SM, Jo J-E, Park S**. 2009. Development and evaluation of efficient recombinant Escherichia coli strains for the production of 3-hydroxypropionic acid from glycerol. Biotechnol. Bioeng. **104:**729-739.

19. **Vinod Kumar, Mugesh Sankaranarayanan, Meetu Durgapal, Shengfang Zhou, Yeounjoo Ko, Somasundar Ashok, Ritam Sarkar SP.** 2013. Simultaneous production of 3-hydroxypropionic acid and 1,3-propanediol from glycerol using resting cells of the lactate dehydrogenase-deficient recombinant *Klebsiella pneumoniae* overexpressing an aldehyde dehydrogenase. Bioresour. Technol. **135:**555-563.

20. **Van Maris AJA, Konings WN, van Dijken JP, Pronk JT.** 2004. Microbial export of lactic and 3-hydroxypropanoic acid: implications for industrial fermentation processes. Metab. Eng. **6:**245-255.

21. **Warnecke TE, Lynch MD, Lipscomb ML, Gill RT.** 2012. Identification of a 21

amino acid peptide conferring 3-hydroxypropionic acid stress-tolerance to *Escherichia coli*. Biotechnol. Bioeng. **109:**1347-52.

22. **Warnecke TE, Lynch MD, Karimpour-Fard A, Lipscomb ML, Handke P, Mills T, Ramey CJ, Hoang T, Gill RT.** 2010. Rapid dissection of a complex phenotype through genomic-scale mapping of fitness altering genes. Metab. Eng. **12:**241-250.

23. **Gill R, Lipscomb TE, Lynch MD.** 2008. Compositions and methods for enhancing tolerance for the production of organic chemicals produced by microorganisms. Patent no. US2008050921.

TABLES

Table 1. *P. furiosus* strains used in this study			
Strain	Parent	Genotype/Description	Source
COM1	DSM3638	Δ*pyrF*	(12)
MW56	COM1	Δ*pyrF* P_{gdh}*pyrF*-P_{slp}-E1αβγ-E2-E3	(6)
MW76	COM1	Δ*pyrF* P_{gdh}*pyrF*-P_{pep}-BPL-CA-P_{slp}-E1αβγ-E2-E3	This work

Table 2. Specific ORFs inserted into *P. furiosus* strains

Symbol	Enzyme	ORFs	MW56	MW76	COM1
E1	Acetyl-CoA/Propionyl-CoA carboxylase	Msed_0147, 0148, 1375	Y	Y	N
E2	Malonyl-CoA reductase	Msed_0709	Y	Y	N
E3	Malonic semialdehyde reductase	Msed_1993	Y	Y	N
BPL	Biotin protein ligase	Msed_2010	N	Y	N
CA	Carbonic anhydrase	Msed_0390	N	Y	N

Table 3. Transcription levels of *M. sedula* ORFs in MW76 at 2.5 h and 40 h incubation at 72°C (compared to COM1)

ORF		**Description**	**Log_2 (fold change)**	
			2.5 h	**40 h**
Msed_0147	E1- α	Biotin carboxylase	4.8	2.9
Msed_0148	E1- β	Biotin carboxyl carrier protein	4.4	5.1
Msed_0390	CA	Carbonic anhydrase	4.1	4.4
Msed_0709	E2	Malonyl-CoA reductase	4.7	5.5
Msed_1375	E1- γ	Carboxyltransferase	5.1	4.9
Msed_1993	E3	Malonic semialdehyde reductase	3.0	2.5
Msed_2010	BPL	Biotin protein ligase	4.3	4.0

Table 4. Selected *P. furiosus* ORFs whose expression is up-regulated in MW76 vs, COM1 after 40 h incubation at 72°C			
Pfu COM1 ORF	**Pfu WT ORF**	**Description**	**Log_2 (fold change)**
Biosynthesis of biotin			
PFC_00380	PF0265	2-amino-3-ketobutyrate coenzyme A ligase	2.8
PFC_07480	PF0144	hypothetical protein	2.0
Energy metabolism			
PFC_07315	PF0177	V-type ATP synthase subunit I	2.2
PFC_07310	PF0178	V-type ATP synthase subunit K	3.2
PFC_07300	PF0180	V-type ATP synthase subunit C	1.8
PFC_07295	PF0181	V-type ATP synthase subunit F	1.4
PFC_07285	PF0183	V-type ATP synthase subunit B	2.2
PFC_00035	PF0196	glucose-6-phosphate isomerase	1.4
PFC_00160	PF0220	D-arabino 3-hexulose 6-phosphate formaldehyde lyase	1.6
PFC_06370	PF1429	monovalent cation/H+ antiporter subunit C	2.3
PFC_06340	PF1423	monovalent cation/H+ antiporter subunit E	1.4
PFC_06480	PF1449	monovalent cation/H+ antiporter subunit B	1.4
PFC_03670	PF0892	cytochrome-c3 hydrogenase subunit gamma	1.4
PFC_09040	PF1909	ferredoxin	2.5
PFC_08720	PF1972	anaerobic ribonucleoside-triphosphate reductase activating protein	3.2
Protein synthesis			
PFC_00320	PF0253	30S ribosomal protein S24e	1.2
PFC_00885	PF0376	50S ribosomal protein LX	1.4
PFC_05610	PF1279	50S ribosomal protein L10e	1.5
PFC_05230	PF1204	seryl-tRNA ligase	1.7
PFC_05960	PF1349	translation initiation factor IF-2B subunit beta	1.4
PFC_07035	PF1561	50S ribosomal protein L30e	1.7
PFC_07080	PF1570	RNA-associated protein	2.8
PFC_10385	PF1644	30S ribosomal protein S9P	1.8
PFC_10195	PF1682	ribosomal protein s6 modification protein	1.8
PFC_08615	PF1990	transcription antitermination protein NusG	1.1
PFC_09580	PF1802	50S ribosomal protein L15P	1.5
PFC_09485	PF1821	30S ribosomal protein S19P	1.1

Table 4 Continued			
Pfu COM1 ORF	**Pfu WT ORF**	**Description**	**Log_2 (fold change)**
PFC_09575	PF1803	50S ribosomal protein L30P	1.4
PFC_09570	PF1804	30S ribosomal protein S5P	1.2
PFC_09555	PF1807	50S ribosomal protein L32e	1.5
PFC_09550	PF1808	50S ribosomal protein L6P	1.6
PFC_09545	PF1809	30S ribosomal protein S8P	1.8
PFC_09540	PF1810	30S ribosomal protein S14P	1.5
PFC_09505	PF1817	translation initiation factor Sui1	1.6
PFC_09500	PF1818	50S ribosomal protein L29P	2.2
PFC_09560	PF1806	50S ribosomal protein L19e	1.8
PFC_09490	PF1820	50S ribosomal protein L22P	1.8
PFC_09470	PF1824	50S ribosomal protein L4P	1.2
PFC_10065	PF1708	anthranilate synthase component II	1.3
PFC_08605	PF1992	50S ribosomal protein L1P	1.4
PFC_08600	PF1993	acidic ribosomal protein P0	1.6
PFC_08500	PF2012	elongation factor EF-2	1.7
Transport and binding proteins			
PFC_06350	PF1425	monovalent cation/H+ antiporter subunit G	1.6
PFC_06490	PF1451	monovalent cation/H+ antiporter subunit G	1.5
PFC_09900	PF1739	trehalose/maltose binding protein	3.3

.

Table 5. Selected *P. furiosus* ORFs whose expression is down-regulated in COM1 (40 h vs. 2.5h incubation at 72°C)			
Pfu COM1 ORF	**Pfu WT ORF**	**Annotation**	**Log_2 (fold change)**
Amino acid biosynthesis			
Glutamate family			
PFC_00095	PF0207	argininosuccinate synthase	-1.8
PFC_02100	PF0594	ornithine carbamoyltransferase	-2.8
PFC_10190	PF1683	N-acetyl-gamma-glutamyl-phosphate reductase	-3.2
PFC_10185	PF1684	acetylglutamate/acetylaminoadipate kinase	-3.5
Pyruvate family			
PFC_10215	PF1678	2-isopropylmalate synthase	-4.8
PFC_10210	PF1679	3-isopropylmalate dehydratase large subunit	-5.1
PFC_10205	PF1680	3-isopropylmalate dehydratase small subunit	-4.3
PFC_10200	PF1681	lysine biosynthesis protein	-4.3
PFC_10175	PF1686	acetyl-lysine deacetylase	-2.1
PFC_10095	PF1702	aspartate aminotransferase	-1.7
Aromatic amino acid family			
PFC_10155	PF1690	3-deoxy-7-phosphoheptulonate synthase	-1.7
PFC_10150	PF1691	3-dehydroquinate synthase	-2.0
PFC_10085	PF1704	hypothetical protein	-1.1
DNA metabolism			
PFC_07125	PF1579	DNA topoisomerase VI subunit B	-2.3
PFC_09180	PF1881	DNA/RNA-binding protein albA	-2.1
PFC_08925	PF1931	hypothetical protein	-2.1
PFC_08470	PF2018	hypothetical protein	-2.0
PFC_01675	PF0511	3-methyladenine DNA glycosylase	-1.1
PFC_06140	PF1383	hypothetical protein	-1.6
PFC_07570	PF0126	DNA repair protein rad25	-1.5
PFC_08950	PF1926	DNA repair and recombination protein RadA	-1.0
PFC_09435	PF1831	histone a1	-1.9
Energy metabolism			
PFC_07990	PF0043	phosphoenolpyruvate synthase	-2.7
PFC_07315	PF0177	V-type ATP synthase subunit I	-3.0
PFC_07310	PF0178	V-type ATP synthase subunit K	-3.6

Table 5 Continued			
Pfu COM1 ORF	**Pfu WT ORF**	**Annotation**	**Log_2 (fold change)**
PFC_07305	PF0179	V-type ATP synthase subunit E	-1.6
PFC_07300	PF0180	V-type ATP synthase subunit C	-2.6
PFC_07295	PF0181	V-type ATP synthase subunit F	-2.4
PFC_07285	PF0183	V-type ATP synthase subunit B	-3.3
PFC_00035	PF0196	glucose-6-phosphate isomerase	-2.9
PFC_02915	PF0751	flavoprotein	-2.0
PFC_02925	PF0753	2-oxoacid ferredoxin oxidoreductase subunit beta	-1.9
PFC_02930	PF0754	2-oxoacid ferredoxin oxidoreductase subunit alpha	-2.0
PFC_03150	PF0796	galactoside o-acetyltransferase	-1.8
PFC_04055	PF0965	pyruvate ferredoxin oxidoreductase subunit beta	-1.5
PFC_04060	PF0966	pyruvate ferredoxin oxidoreductase subunit alpha	-4.0
PFC_04065	PF0967	pyuvate ferredoxin oxidoreductase subunit delta	-2.5
PFC_04070	PF0968	2-ketoisovalerate ferredoxin oxidoreductase subunit beta	-2.6
PFC_04075	PF0969	2-ketoisovalerate ferredoxin oxidoreductase subunit alpha	-3.7
PFC_04080	PF0970	2-ketoisovalerate ferredoxin oxidoreductase subunit delta	-2.2
PFC_04085	PF0971	pyruvate/ketoisovalerate ferredoxin oxidoreductase subunit gamma	-3.2
PFC_06385	PF1432	mbh10 NADH dehydrogenase subunit	-3.1
PFC_06395	PF1434	membrane bound hydrogenase alpha	-2.3
PFC_06400	PF1435	mbh13 NADH dehydrogenase subunit	-2.1
PFC_06405	PF1436	NADH-plastoquinone oxidoreductase subunit	-2.2
PFC_06445	PF1444	NADH dehydrogenase subunit B	-2.5
PFC_06635	PF1480	formaldehyde:ferredoxin oxidoreductase wor5	-1.9
PFC_10400	PF1641	hypothetical protein	-1.0
PFC_10165	PF1688	transketolase	-2.2
PFC_09860	PF1747	hypothetical protein	-1.5
PFC_09825	PF1755	fumarate hydratase	-1.6
PFC_09740	PF1770	2-oxoglutarate ferredoxin oxidoreductase subunit gamma	-1.3
PFC_09735	PF1771	2-oxoglutarate ferredoxin oxidoreductase subunit alpha	-1.6
PFC_09730	PF1772	2-oxoglutarate ferredoxin oxidoreductase subunit beta	-1.8
PFC_09725	PF1773	2-oxoglutarate ferredoxin oxidoreductase subunit gamma	-1.9
PFC_09655	PF1787	acetyl-CoA synthetase I subunit beta	-1.8

Table 5 Continued			
Pfu COM1 ORF	**Pfu WT ORF**	**Annotation**	**Log_2 (fold change)**
PFC_09155	PF1886	sugar kinase	-1.2
PFC_09040	PF1909	ferredoxin	-3.5
PFC_08910	PF1935	pullulanase	-1.3
PFC_08800	PF1956	fructose-bisphosphate aldolase	-1.9
PFC_08775	PF1961	aldehyde:ferredoxin oxidoreductase	-1.3
PFC_08720	PF1972	anaerobic ribonucleoside-triphosphate reductase activating protein	-2.8
PFC_08705	PF1975	phosphoenolpyruvate carboxylase	-1.3
PFC_08535	PF2005	glycerol-3-phosphate dehydrogenase	-1.1
Transcription			
PFC_00330	PF0255	DNA-directed RNA polymerase subunit E"	-2.0
PFC_02110	PF0596	mRNA 3'-end processing factor	-1.3
PFC_05690	PF1295	transcription factor	-1.1
PFC_06975	PF1549	RNA 3'-terminal-phosphate cyclase	-1.0
PFC_00335	PF0256	DNA-directed RNA polymerase subunit E'	-2.2
PFC_07040	PF1562	DNA-directed RNA polymerase subunit A"	-3.0
PFC_07055	PF1565	DNA-directed RNA polymerase subunit H	-2.1
PFC_07065	PF1567	exosome complex RNA-binding protein Rrp42	-2.1
PFC_07070	PF1568	exosome complex exonuclease Rrp41	-2.9
PFC_10370	PF1647	DNA-directed RNA polymerase subunit D	-1.6
PFC_10395	PF1642	DNA-directed RNA polymerase subunit K	-1.9
Protein synthesis			
PFC_00315	PF0252	30S ribosomal protein S27ae	-1.9
PFC_02370	PF0646	hypothetical protein	-1.5
PFC_05680	PF1293	prolyl-tRNA ligase	-1.5
PFC_06055	PF1367	50S ribosomal protein L7Ae	-1.3
PFC_06060	PF1368	30S ribosomal protein S28e	-1.9
PFC_06105	PF1376	30S ribosomal protein S10P	-1.2
PFC_06120	PF1379	cell division protein pelota	-1.6
PFC_07425	PF0155	asparaginyl-tRNA ligase	-2.2
PFC_00320	PF0253	30S ribosomal protein S24e	-2.1
PFC_07020	PF1558	30S ribosomal protein S7P	-1.2
PFC_07035	PF1561	50S ribosomal protein L30e	-1.9
PFC_00885	PF0376	50S ribosomal protein LX	-2.0
PFC_05610	PF1279	50S ribosomal protein L10e	-2.6

Table 5 Continued			
Pfu COM1 ORF	**Pfu WT ORF**	**Annotation**	**Log_2 (fold change)**
PFC_06100	PF1375	elongation factor 1-alpha	-3.2
PFC_07080	PF1570	RNA-associated protein	-2.1
PFC_10385	PF1644	30S ribosomal protein S9P	-2.3
PFC_10365	PF1648	30S ribosomal protein S11P	-2.8
PFC_10360	PF1649	30S ribosomal protein S4P	-2.1
PFC_10195	PF1682	ribosomal protein s6 modification protein	-3.4
PFC_09580	PF1802	50S ribosomal protein L15P	-2.9
PFC_09575	PF1803	50S ribosomal protein L30P	-3.1
PFC_09570	PF1804	30S ribosomal protein S5P	-3.6
PFC_09565	PF1805	50S ribosomal protein L18P	-2.5
PFC_09555	PF1807	50S ribosomal protein L32e	-3.2
PFC_09550	PF1808	50S ribosomal protein L6P	-3.0
PFC_09545	PF1809	30S ribosomal protein S8P	-3.2
PFC_09540	PF1810	30S ribosomal protein S14P	-2.0
PFC_09525	PF1813	50S ribosomal protein L24P	-2.0
PFC_08610	PF1991	50S ribosomal protein L11P	-1.8
PFC_08615	PF1990	transcription antitermination protein NusG	-1.3
PFC_09505	PF1817	translation initiation factor Sui1	-3.6
PFC_09500	PF1818	50S ribosomal protein L29P	-2.3
PFC_09495	PF1819	30S ribosomal protein S3P	-2.8
PFC_09490	PF1820	50S ribosomal protein L22P	-2.1
PFC_09480	PF1822	50S ribosomal protein L2P	-2.5
PFC_09475	PF1823	50S ribosomal protein L23P	-2.4
PFC_09470	PF1824	50S ribosomal protein L4P	-2.6
PFC_08605	PF1992	50S ribosomal protein L1P	-2.8
PFC_09465	PF1825	50S ribosomal protein L3P	-1.7
PFC_08600	PF1993	acidic ribosomal protein P0	-3.5
PFC_08595	PF1994	50S ribosomal protein L12P	-2.3
PFC_08500	PF2012	elongation factor EF-2	-2.0
PFC_09485	PF1821	30S ribosomal protein S19P	-2.0
PFC_09515	PF1815	30S ribosomal protein S17P	-1.6
PFC_09560	PF1806	50S ribosomal protein L19e	-1.9
PFC_10020	PF1717	translation initiation factor IF-2 subunit gamma	-1.6
PFC_10065	PF1708	anthranilate synthase component II	-1.2

Table 5 Continued			
Pfu COM1 ORF	**Pfu WT ORF**	**Annotation**	**Log_2 (fold change)**
PFC_10320	PF1657	histidyl-tRNA synthetase	-1.4
PFC_10355	PF1650	30S ribosomal protein S13P	-1.1
PFC_10375	PF1646	50S ribosomal protein L18e	-1.9
PFC_10380	PF1645	50S ribosomal protein L13P	-1.5
PFC_10470	PF1627	glycyl-tRNA ligase	-1.8

Table 6. Selected *P. furiosus* ORFs whose expression is up-regulated in MW76 (40 h vs. 2.5 h incubation at 72°C)

Pfu COM1 ORF	Pfu WT ORF	Annotation	Log_2 (fold change)
Amino acid biosynthesis			
Serine family			
PFC_00115	PF0211	hypothetical protein	1.8
PFC_03430	PF0847	hypothetical protein	1.4
PFC_05320	PF1220	sulfate adenylyltransferase	1.1
PFC_09700	PF1778	serine hydroxymethyltransferase	2.2
Aspartate family			
PFC_05120	PF1180	threonine synthase	2.2
PFC_10175	PF1686	acetyl-lysine deacetylase	1.3
Glutamate family			
PFC_01365	PF0450	glutamine synthetase	1.7
Histidine family			
PFC_10295	PF1662	1-(5-phosphoribosyl)-5-[(5-phosphoribosylamino)methylideneamino] imidazole-4-carboxamide isomerase	1.4
Aromatic amino acid family			
PFC_00520	PF0291	prephenate dehydratase	1.5
PFC_10100	PF1701	chorismate mutase	2.3
PFC_10060	PF1709	anthranilate synthase component I	1.4
PFC_10055	PF1710	anthranilate phosphoribosyltransferase	3.4
PFC_10050	PF1711	indole-3-glycerol-phosphate synthase	3.3
Energy metabolism			
PFC_07990	PF0043	phosphoenolpyruvate synthase	1.7
PFC_00640	PF0312	ADP-dependent glucokinase	1.6
PFC_00985	PF0356	beta-galactosidase	1.2
PFC_00950	PF0363	beta-galactosidase	1.4
PFC_01320	PF0441	galactose-1-phosphate uridylyltransferase	1.9
PFC_01340	PF0445	galactokinase	1.8
PFC_02935	PF0755	non-phosphorylating glyceraldehyde-3-phosphate dehydrogenase	1.5
PFC_03670	PF0892	cytochrome-c3 hydrogenase subunit gamma	1.1
PFC_04055	PF0965	pyruvate ferredoxin oxidoreductase subunit beta	1.6
PFC_04105	PF0975	hypothetical protein	1.2

Table 6 Continued			
Pfu COM1 ORF	**Pfu WT ORF**	**Annotation**	**Log_2 (fold change)**
PFC_04645	PF1085	hypothetical protein	1.4
PFC_05920	PF1341	glycine cleavage system aminomethyltransferase T	2.1
PFC_09400	PF1838	hypothetical protein	2.2
PFC_08710	PF1974	thermosome, single subunit	2.6

Table 7. Selected *P. furiosus* ORFs whose expression is up-regulated in COM1 vs. MW76 after 2.5 h incubation at 72°C			
Pfu COM1 ORF	**Pfu WT ORF**	**Annotation**	**Log_2 (fold change)**
Amino acid biosynthesis			
Pyruvate family			
PFC_10215	PF1678	2-isopropylmalate synthase	3.5
PFC_10210	PF1679	3-isopropylmalate dehydratase large subunit	3.7
PFC_10205	PF1680	3-isopropylmalate dehydratase small subunit	3.2
Aspartate family			
PFC_10200	PF1681	lysine biosynthesis protein	3.0
PFC_10175	PF1686	acetyl-lysine deacetylase	2.6
Glutamate family			
PFC_10190	PF1683	N-acetyl-gamma-glutamyl-phosphate reductase	2.7
PFC_10185	PF1684	acetylglutamate/acetylaminoadipate kinase	3.1
PFC_10180	PF1685	acetylornithine/acetyl-lysine aminotransferase	2.9
Cellular processes			
PFC_00165	PF0221	hypothetical protein	1.0
PFC_05620	PF1281	superoxide reductase	1.0
PFC_04400	PF1033	peroxiredoxin	1.2
PFC_05650	PF1287	ABC transporter	1.3
Energy metabolism			
PFC_08710	PF1974	thermosome, single subunit	1.3
Protein synthesis			
PFC_10195	PF1682	ribosomal protein s6 modification protein	1.5
Transport and binding protein			
PFC_05255	PF1209	oligopeptide ABC transporter	2.0

Table 8. ***P. furiosus*** **ORFs up-regulated (red font) or down-regulated(blue font) in both COM1 and MW76 (2.5h vs 40h) at 72℃**

			COM1		MW76	
Pfu ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change COM1 (2.5h)-(40h)**	**-log 10(p value)**	**Average Log 2 fold change MW76 (2.5h)-(40h)**	**-log 10(p value)**
PF0011	PFC_08155	ABC transporter transmembrane component	-1.7	8.0	-2.0	9.5
PF0012	PFC_08150	hypothetical protein	-1.4	6.9	-1.2	6.0
PF0021	PFC_08110	DNA repair and recombination protein RadB	-1.3	8.9	-1.6	11.2
PF0024	PFC_08090	glycerate kinase	1.1	7.6	1.3	9.3
PF0025	PFC_08085	hypothetical protein	-1.6	10.2	-1.3	8.4
PF0028	PFC_08070	acetyl transferase	-1.4	10.2	-1.1	8.0
PF0072	PFC_07840	hypothetical protein	-1.6	6.1	-1.7	6.5
PF0107	PFC_07670	hypothetical protein	-1.8	11.0	-1.1	5.9

Table 8 Continued						
			COM1		MW76	
Pfu ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change COM1 (2.5h)-(40h)**	**-log 10(p value)**	**Average Log 2 fold change MW76 (2.5h)-(40h)**	**-log 10(p value)**
PF0173	PFC_07335	preprotein translocase subunit SecF	-1.1	8.0	-1.1	7.9
PF0181	PFC_07295	V-type ATP synthase subunit F	2.4	15.7	1.1	8.2
PF0196	PFC_00035	glucose-6-phosphate isomerase	2.9	17.3	1.2	7.5
PF0207	PFC_00095	argininosuccinate synthase	1.8	8.9	1.6	8.2
PF0209	PFC_00105	ribosomal protein s6 modification protein	-1.5	6.1	-1.8	7.9
PF0211	PFC_00115	hypothetical protein	-1.9	10.3	-1.8	10.1
PF0215	PFC_00135	enolase	1.7	14.9	1.2	10.6

Table 8 Continued

Pfu ORF#	Pfu COM1 ID	Annotation	Average Log 2 fold change COM1 (2.5h)-(40h)	-log 10(p value)	Average Log 2 fold change MW76 (2.5h)-(40h)	-log 10(p value)
PF0220	PFC_00160	D-arabino 3-hexulose 6-phosphate formaldehyde lyase	2.7	14.7	1.1	5.6
PF0252	PFC_00315	30S ribosomal protein S27ae	1.9	10.8	1.2	6.4
PF0254	PFC_00325	hypothetical protein	2.8	18.0	1.1	7.3
PF0262	PFC_00365	MFS family transporter	2.1	12.5	2.0	12.1
PF0293	PFC_00530	aminotransferase	1.0	9.0	1.2	10.9
PF0302	PFC_00585	co-factor modifying protein	2.0	13.4	1.6	10.7
PF0303	PFC_00590	hypothetical protein	-1.2	11.1	-1.0	10.0
PF0316	PFC_00655	hypothetical protein	-2.1	17.0	-1.8	15.2
PF0336	PFC_00760	hypothetical protein	-2.8	23.3	-1.3	13.8
PF0337	PFC_00765	flagellin	-1.5	10.9	-1.0	6.9
PF0340	PFC_00780	HTH transcription regulator	-1.1	6.5	-1.7	10.5
PF0363	PFC_00950	beta-galactosidase	-1.0	6.4	-1.4	9.2

Table 8 Continued

Pfu ORF#	Pfu COM1 ID	Annotation	Average Log 2 fold change COM1 (2.5h)-(40h)	-log 10(p value)	Average Log 2 fold change MW76 (2.5h)-(40h)	-log 10(p value)
PF0386	PFC_00835	hypothetical protein	1.5	8.6	1.2	6.6
PF0401	PFC_01090	methyltransferase	4.1	19.7	2.1	11.7
PF0402	PFC_01100	UDP-or dTTP-glucose 4-epimerase or 4-6-dehydratase	-1.0	6.0	-1.3	8.3
PF0413	PFC_01150	hypothetical protein	1.7	11.2	1.1	7.2
PF0426	PFC_01215	phosphoribosylaminoimidazole carboxylase ATPase subunit	1.5	9.7	1.3	7.8
PF0445	PFC_01340	galactokinase	-1.9	9.5	-1.8	8.8
PF0460	PFC_01415	hypothetical protein	1.7	6.9	1.4	5.7
PF0473	PFC_01480	serine/threonine protein kinase	-2.0	15.5	-1.5	11.6
PF0493	PFC_01580	hypothetical protein	-1.5	11.8	-1.4	11.1
PF0508	PFC_01660	integral membrane glycosyltransferase	-1.9	13.8	-1.4	10.5

Table 8 Continued						
Pfu ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change COM1 (2.5h)-(40h)**	**-log 10(p value)**	**Average Log 2 fold change MW76 (2.5h)-(40h)**	**-log 10(p value)**
PF0510	PFC_01670	nucleotide pyrophosphohydrolase	1.0	8.4	1.2	10.0
PF0560	PFC_01930	hypothetical protein	-1.7	9.5	-1.8	9.6
PF0566	PFC_01960	hypothetical protein	-1.2	8.6	-1.9	13.7
PF0591	PFC_02085	mannosyl-3-phosphoglycerate synthase	1.5	7.4	1.2	5.6
PF0615	PFC_02215	hydrogenase nickel incorporation protein	1.6	15.2	1.0	10.0
PF0641	PFC_02340	hypothetical protein	1.2	9.6	1.3	11.1
PF0655.1n	PFC_02420	hypothetical protein	-1.3	7.9	-1.4	8.4
PF0728	PFC_02785	hypothetical protein	-1.3	9.4	-1.1	7.7
PF0732	PFC_02805	daunorubicin resistance ATP-binding protein	-3.0	11.6	-2.0	7.6
PF0743	PFC_02875	hypothetical protein	-1.0	8.6	-1.3	10.9

Table 8 Continued						
Pfu ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change COM1 (2.5h)-(40h)**	**-log 10(p value)**	**Average Log 2 fold change MW76 (2.5h)-(40h)**	**-log 10(p value)**
PF0763	PFC_02965	hypothetical protein	1.5	12.3	1.1	8.8
PF0770	PFC_03005	glucose-1-phosphate uridylyltransferase	-1.8	10.7	-2.3	13.8
PF0802	PFC_03180	hypothetical protein	-2.1	13.6	-1.7	11.0
PF0812	PFC_03235	hypothetical protein	-1.6	5.9	-1.8	6.9
PF0815	PFC_03250	hypothetical protein	-1.4	8.2	-1.0	5.6
PF0825	PFC_03310	prolyl endopeptidase	1.5	7.6	1.2	5.7
PF0829	PFC_03330	hypothetical protein	-2.6	14.7	-1.3	8.4
PF0831	PFC_03340	hypothetical protein	-1.2	9.0	-1.0	7.3
PF0844	PFC_03415	hypothetical protein	2.2	17.3	1.1	9.1
PF0847	PFC_03430	hypothetical protein	-1.2	6.4	-1.4	7.9
PF0880	PFC_03600	DEXX-box atpase	-2.0	14.2	-1.6	11.8
PF0881	PFC_03605	ABC transporter	-1.4	8.5	-1.2	7.1

Table 8 Continued						
Pfu ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change COM1 (2.5h)-(40h)**	**-log 10(p value)**	**Average Log 2 fold change MW76 (2.5h)-(40h)**	**-log 10(p value)**
PF0889	PFC_03655	hypothetical protein	-1.2	6.4	-1.6	9.8
PF0910	PFC_03770	iron ABC transporter permease	-3.5	23.0	-3.3	22.3
PF0929	PFC_03865	hypothetical protein	1.7	12.1	1.2	8.8
PF0930	PFC_03870	hypothetical protein	1.5	10.9	1.2	8.2
PF0966	PFC_04060	pyruvate ferredoxin oxidoreductase subunit alpha	4.0	24.0	1.5	12.0
PF0972	PFC_04090	hypothetical protein	3.2	16.8	1.3	7.0
PF0973	PFC_04095	acetyl-CoA acetyltransferase	2.2	10.5	1.4	6.2
PF0975	PFC_04105	hypothetical protein	-1.2	5.8	-1.2	5.6
PF0979	PFC_04125	cobalamin adenosyltransferase	1.8	10.2	1.3	7.0
PF0994	PFC_04200	type II secretion system protein	1.1	7.1	1.1	7.0

Table 8 Continued						
Pfu ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change COM1 (2.5h)-(40h)**	**-log 10(p value)**	**Average Log 2 fold change MW76 (2.5h)-(40h)**	**-log 10(p value)**
PF1049	PFC_04480	Pyruvate formate lyase activating like protein	-1.5	8.8	-1.5	8.2
PF1059	PFC_04530	hypothetical protein	-1.1	6.5	-1.4	9.0
PF1077	PFC_04610	hypothetical protein	-1.1	6.2	-2.3	13.9
PF1080	PFC_04625	hypothetical protein	-1.3	7.0	-1.2	6.4
PF1107	PFC_04760	hit family protein	-1.3	6.6	-1.2	6.5
PF1128	PFC_04860	hypothetical protein	1.9	15.8	1.2	10.3
PF1130	PFC_04870	hypothetical protein	1.1	7.7	1.0	7.7
PF1239	PFC_05415	hypothetical protein	-1.2	7.4	-1.1	6.9
PF1243	PFC_05445	hypothetical protein	-2.4	6.3	-2.6	7.0
PF1258	PFC_05495	ribose-5-phosphate isomerase A	1.3	6.1	1.3	5.9
PF1271	PFC_05570	hypothetical protein	-2.4	13.6	-1.3	6.9
PF1281	PFC_05620	superoxide reductase	3.0	25.0	1.5	16.0

Table 8 Continued						
Pfu ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change COM1 (2.5h)-(40h)**	**-log 10(p value)**	**Average Log 2 fold change MW76 (2.5h)-(40h)**	**-log 10(p value)**
PF1295	PFC_05690	transcription factor	1.1	8.4	1.1	8.3
PF1310	PFC_05760	hypothetical protein	-1.8	11.6	-1.7	11.1
PF1313	PFC_05775	DNA processing SMF protein	-1.8	7.6	-1.4	5.5
PF1314	PFC_05780	hypothetical protein	-1.2	7.2	-1.5	9.0
PF1336	PFC_05895	cytosine permease	-1.4	9.6	-2.3	15.4
PF1341	PFC_05920	glycine cleavage system aminomethyltransferase T	-1.8	6.0	-2.1	7.2
PF1367	PFC_06055	50S ribosomal protein L7Ae	1.3	12.2	1.6	15.0
PF1368	PFC_06060	30S ribosomal protein S28e	1.9	18.9	1.7	17.3
PF1371	PFC_06075	hypothetical protein	-1.5	8.5	-1.7	9.2
PF1375	PFC_06100	elongation factor 1-alpha	3.2	16.9	1.5	8.2
PF1388	PFC_06165	hypothetical protein	-2.9	18.5	-1.0	7.9
PF1444	PFC_06445	NADH dehydrogenase subunit B	2.5	11.2	1.4	5.8

Table 8 Continued

Pfu ORF#	Pfu COM1 ID	Annotation	Average Log 2 fold change COM1 (2.5h)-(40h)	-log 10(p value)	Average Log 2 fold change MW76 (2.5h)-(40h)	-log 10(p value)
PF1460	PFC_06535	hypothetical protein	-1.1	5.7	-1.2	6.3
PF1468	PFC_06575	hypothetical protein	-3.0	22.8	-3.4	24.6
PF1496	PFC_06715	hypothetical protein	-1.6	11.2	-1.5	11.1
PF1536	PFC_06910	RNase III inhibitor	2.2	8.2	1.6	5.5
PF1541	PFC_06935	50S ribosomal protein L37e	-1.0	9.0	-1.5	13.9
PF1549	PFC_06975	RNA 3'-terminal-phosphate cyclase	1.0	9.5	1.4	13.5
PF1558	PFC_07020	30S ribosomal protein S7P	1.2	7.5	1.2	7.6
PF1562	PFC_07040	DNA-directed RNA polymerase subunit A"	3.0	16.8	1.3	7.5
PF1565	PFC_07055	DNA-directed RNA polymerase subunit H	2.1	14.9	1.6	11.7

Table 8 Continued						
Pfu ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change COM1 (2.5h)-(40h)**	**-log 10(p value)**	**Average Log 2 fold change MW76 (2.5h)-(40h)**	**-log 10(p value)**
PF1568	PFC_07070	exosome complex exonuclease Rrp41	2.9	18.6	1.5	10.5
PF1569	PFC_07075	exosome complex RNA-binding protein Rrp4	3.2	14.7	1.8	8.4
PF1615	PFC_10530	hypothetical protein	2.3	17.0	1.4	10.8
PF1619	PFC_10510	hypothetical protein	-1.3	9.5	-1.2	8.3
PF1622	PFC_10495	n-type ATP pyrophosphatase superfamily protein	-2.2	11.2	-1.8	9.5
PF1632	PFC_10445	hypothetical protein	-1.4	7.6	-1.5	8.4
PF1646	PFC_10375	50S ribosomal protein L18e	1.9	11.2	1.6	9.4
PF1648	PFC_10365	30S ribosomal protein S11P	2.8	17.0	1.8	11.8
PF1649	PFC_10360	30S ribosomal protein S4P	2.1	15.4	1.4	10.8
PF1656	PFC_10325	NADH oxidase	-1.8	11.1	-1.7	11.4

Table 8 Continued

Pfu ORF#	Pfu COM1 ID	Annotation	Average Log 2 fold change COM1 (2.5h)-(40h)	-log 10(p value)	Average Log 2 fold change MW76 (2.5h)-(40h)	-log 10(p value)
PF1657	PFC_10320	histidyl-tRNA synthetase	1.4	11.4	1.4	11.0
PF1662	PFC_10295	1-(5-phosphoribosyl)-5-[(5-phosphoribosylamino)methylideneamino] imidazole-4-carboxamide isomerase	-1.5	5.6	-1.4	5.7
PF1679	PFC_10210	3-isopropylmalate dehydratase large subunit	5.1	23.7	1.5	9.9
PF1680	PFC_10205	3-isopropylmalate dehydratase small subunit	4.3	25.0	1.1	8.2
PF1681	PFC_10200	lysine biosynthesis protein	4.3	16.5	1.7	6.5
PF1730	PFC_09955	thymidylate kinase	2.4	10.4	1.5	6.0
PF1731	PFC_09950	signal recognition particle protein Srp54	2.3	12.2	1.4	7.1

Table 8 Continued

Pfu ORF#	Pfu COM1 ID	Annotation	Average Log 2 fold change COM1 (2.5h)-(40h)	-log 10(p value)	Average Log 2 fold change MW76 (2.5h)-(40h)	-log 10(p value)
PF1735.2n	PFC_09925	hypothetical protein	1.9	14.0	1.3	9.5
PF1737	PFC_09915	hypothetical protein	-1.2	6.4	-1.8	10.1
PF1742	PFC_09885	group 1 glycosyl transferase	3.1	14.5	1.9	8.9
PF1743	PFC_09880	hypothetical protein	1.6	8.9	1.2	6.3
PF1750	PFC_09850	sugar-binding transport ATP-binding protein	-2.0	14.2	-1.2	9.0
PF1755	PFC_09825	fumarate hydratase	1.6	15.5	1.1	11.3
PF1783	PFC_09675	molybdenum cofactor biosynthesis protein MoeA	1.3	6.7	1.1	5.9
PF1802	PFC_09580	50S ribosomal protein L15P	2.9	12.4	1.5	6.1
PF1803	PFC_09575	50S ribosomal protein L30P	3.1	18.3	1.9	12.2
PF1804	PFC_09570	30S ribosomal protein S5P	3.6	18.8	1.7	10.1
PF1805	PFC_09565	50S ribosomal protein L18P	2.5	11.8	1.4	6.1

Table 8 Continued

Pfu ORF#	Pfu COM1 ID	Annotation	Average Log 2 fold change COM1 (2.5h)-(40h)	-log 10(p value)	Average Log 2 fold change MW76 (2.5h)-(40h)	-log 10(p value)
PF1807	PFC_09555	50S ribosomal protein L32e	3.2	17.7	1.9	11.6
PF1808	PFC_09550	50S ribosomal protein L6P	3.0	15.2	1.6	8.2
PF1811	PFC_09535	50S ribosomal protein L5P	3.1	18.9	2.0	13.8
PF1813	PFC_09525	50S ribosomal protein L24P	2.0	14.7	1.8	13.5
PF1815	PFC_09515	30S ribosomal protein S17P	1.6	14.6	1.3	12.5
PF1817	PFC_09505	translation initiation factor Sui1	3.6	14.3	1.6	5.9
PF1819	PFC_09495	30S ribosomal protein S3P	2.8	14.8	1.8	9.6
PF1822	PFC_09480	50S ribosomal protein L2P	2.5	17.1	1.8	13.2
PF1823	PFC_09475	50S ribosomal protein L23P	2.4	12.7	2.0	10.5
PF1824	PFC_09470	50S ribosomal protein L4P	2.6	15.6	1.4	8.8
PF1825	PFC_09465	50S ribosomal protein L3P	1.7	10.5	1.1	6.5
PF1831	PFC_09435	histone a1	1.9	22.4	2.0	23.2
PF1838	PFC_09400	hypothetical protein	-3.2	16.7	-2.2	12.2

Table 8 Continued

Pfu ORF#	Pfu COM1 ID	Annotation	Average Log 2 fold change COM1 (2.5h)-(40h)	-log 10(p value)	Average Log 2 fold change MW76 (2.5h)-(40h)	-log 10(p value)
PF1844	PFC_09370	hypothetical protein	-1.2	5.9	-1.2	6.1
PF1857	PFC_09305	cobalt transport ATP-binding protein	-1.0	5.6	-1.0	5.7
PF1886	PFC_09155	sugar kinase	1.2	6.0	1.6	8.1
PF1897	PFC_09100	hypothetical protein	-3.0	14.3	-2.1	10.3
PF1905	PFC_09060	protease	-1.6	10.2	-1.1	6.6
PF1907	PFC_09050	hypothetical protein	2.6	13.4	1.6	8.2
PF1935	PFC_08910	pullulanase	1.3	7.0	2.1	11.3
PF1936	PFC_08905	malg-like sugar transport inner membrane protein	1.6	5.8	1.6	6.2
PF1937	PFC_08900	malf-like sugar transport inner membrane protein	1.8	8.8	1.9	9.2
PF1938	PFC_08895	malE-like sugar binding protein	1.8	6.3	2.0	6.9

Table 8 Continued						
Pfu ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change COM1 (2.5h)-(40h)**	**-log 10(p value)**	**Average Log 2 fold change MW76 (2.5h)-(40h)**	**-log 10(p value)**
PF1940	PFC_08885	hypothetical protein	-1.4	7.7	-2.0	11.7
PF1947	PFC_08845	hypothetical protein	1.4	8.4	1.3	7.8
PF1962	PFC_08770	hypothetical protein	-2.0	12.7	-1.6	10.5
PF1993	PFC_08600	acidic ribosomal protein P0	3.5	21.9	1.4	11.2
PF1994	PFC_08595	50S ribosomal protein L12P	2.3	20.2	1.8	17.6
PF2000	PFC_08565	glycine dehydrogenase subunit 2	-1.7	8.1	-2.1	10.4
PF2005	PFC_08535	glycerol-3-phosphate dehydrogenase	1.1	8.4	1.1	8.7
PF2013	PFC_08495	hypothetical protein	1.9	15.8	1.2	10.4
PF2045	PFC_08340	hypothetical protein	-1.9	12.7	-1.9	12.4
PF2046	PFC_08335	hypothetical protein	-1.9	13.5	-2.0	14.1
PF2048	PFC_08320	peptidase	-1.3	8.6	-1.1	6.7

Table 8 Continued

Pfu ORF#	Pfu COM1 ID	Annotation	Average Log 2 fold change COM1 (2.5h)-(40h)	-log 10(p value)	Average Log 2 fold change MW76 (2.5h)-(40h)	-log 10(p value)
	PFC_09510	ribonuclease P protein component 1	3.4	14.6	1.7	7.3
	PFC_08380	hypothetical protein	-2.2	17.6	-1.2	10.7
	PFC_01065	hypothetical protein	1.1	5.8	1.1	5.6

FIGURES

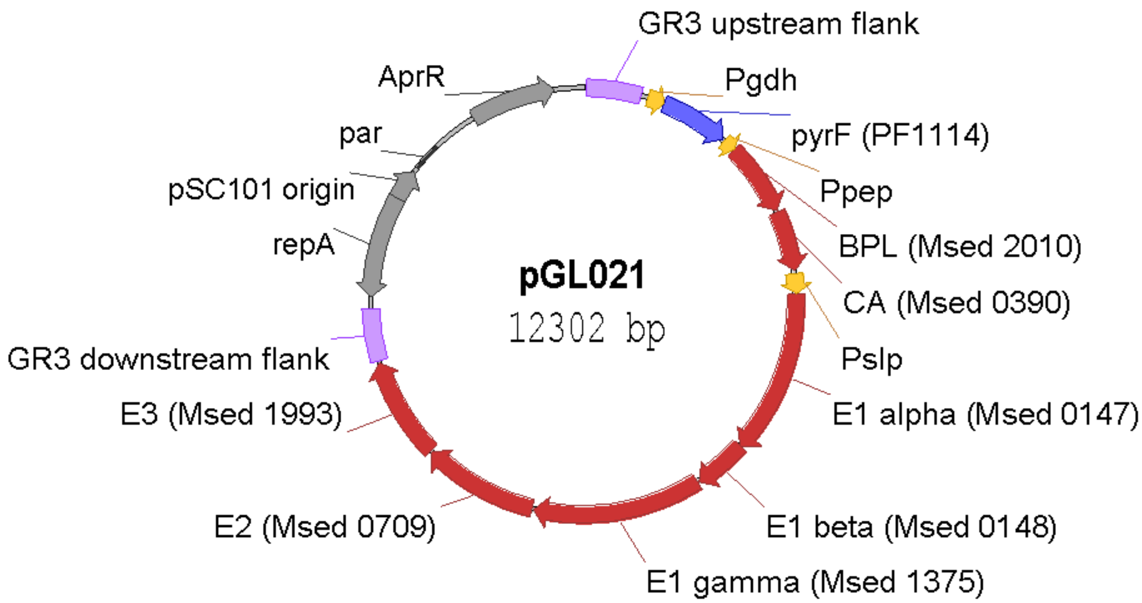


Figure 1. pGL021 used to generate strain MW76 for chromosomal insertion of synthetic operons into *P. furiosus* COM1. Plasmid elements are color-coded as follows: homologous recombination regions for *P. furiosus* chromosomal insertion, lavender; *P. furiosus* promoter regions, yellow; *P. furiosus pyrF* genetic marker, blue; *M. sedula* genes, red; his-tag, dark blue; *E. coli* plasmid maintenance elements, grey.

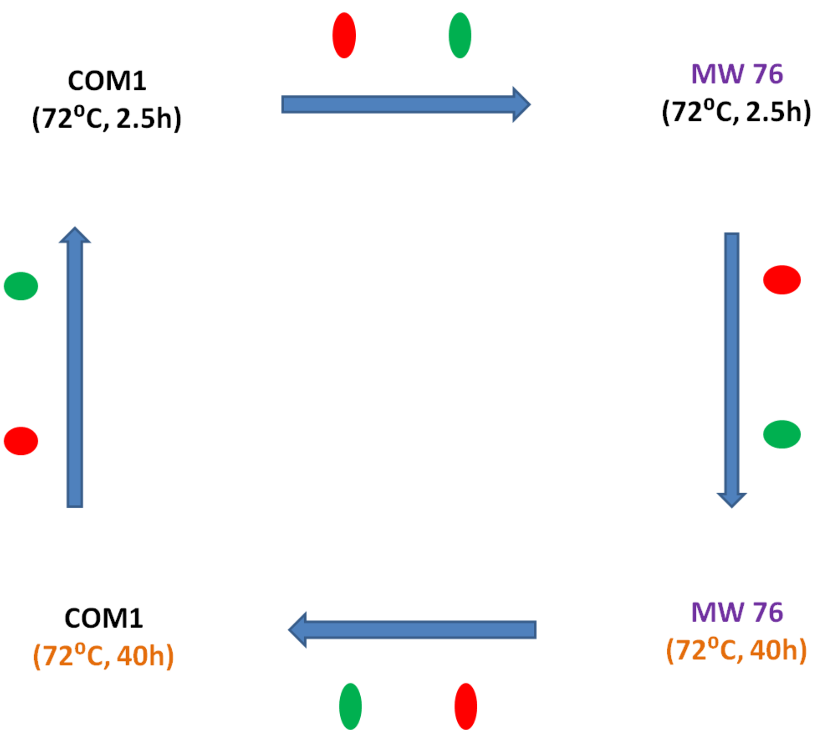


Figure 2. Experimental design for microarray hybridization for comparison of COM1 and MW76 grown with gas feed at 50 mL/min, agitation at 400 rpm, pH controlled at 5.5-6.5 by addition of 10 M NaOH. Samples were taken 2.5 h and 40 h after temperature shift (95°C→72°C) respectively. Green dots and Red dots represent cyanine 3- and cyanine 5-labeled samples, respectively.

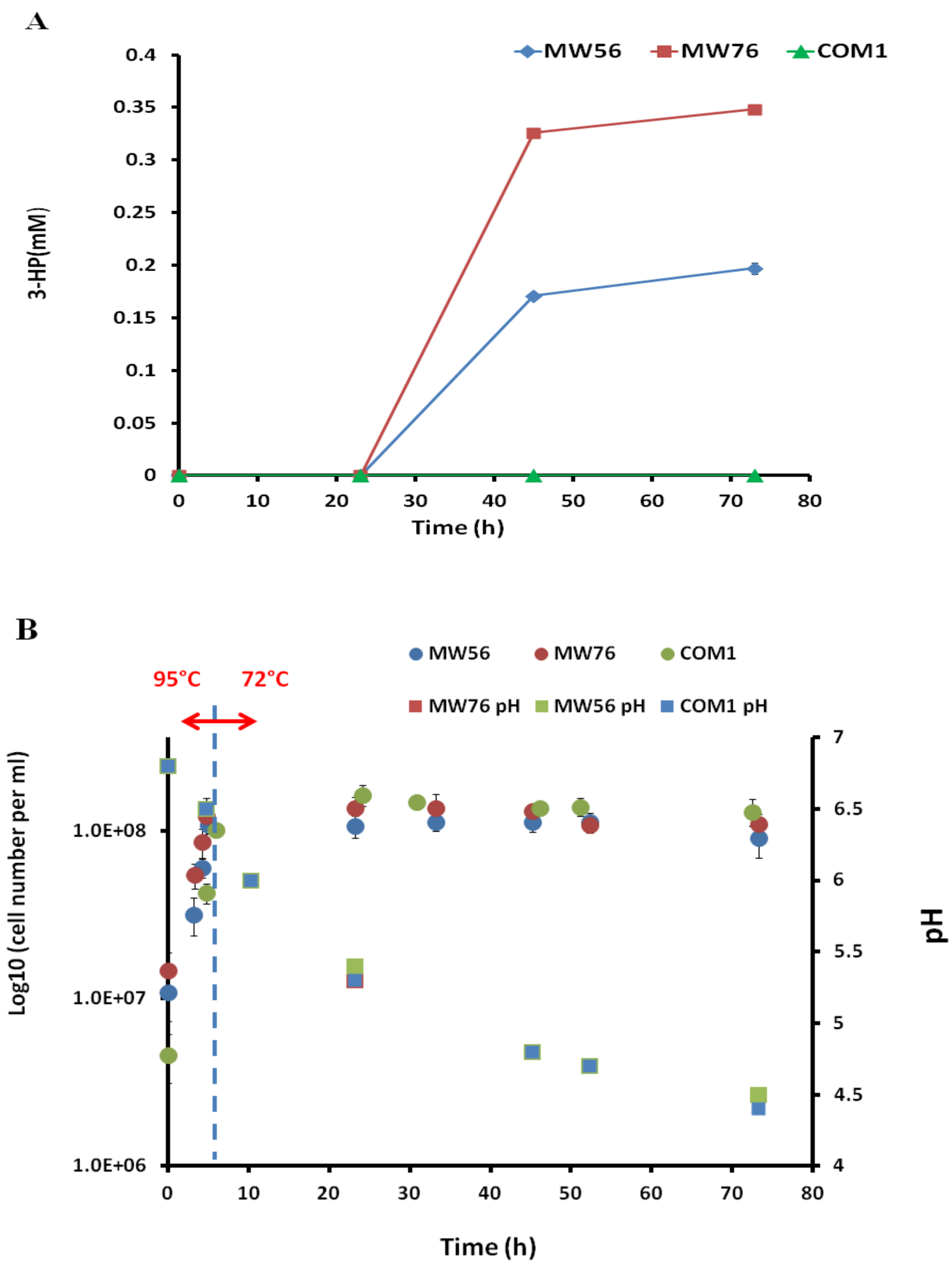


Figure 3. MW56, MW76 and COM1 for gas feed at 15 mL/min, agitation at 250 rpm, no pH control. (A) 3-HP production (B) Cell density and pH changes with time.

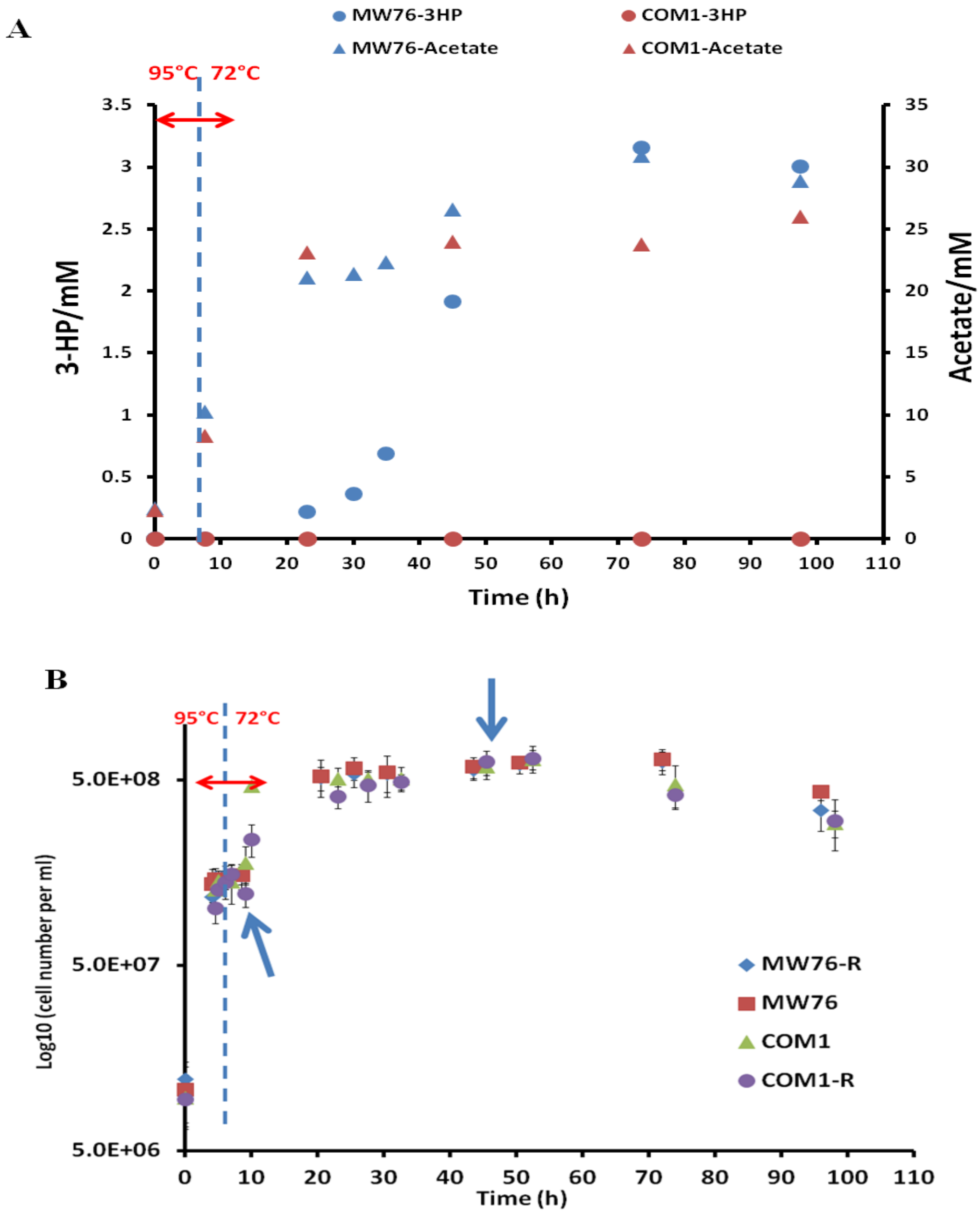


Figure 4. MW76 and COM1 grown with gas feed at 50 mL/min, agitation at 400 rpm, pH controlled at 5.5-6.5 by addition of 10 M NaOH. (A) 3-HP and acetate production; (B) Cell growth curve. MW76-R and COM1-R represents biological replicates. Blue arrow indicated time points of transcriptomic analysis.

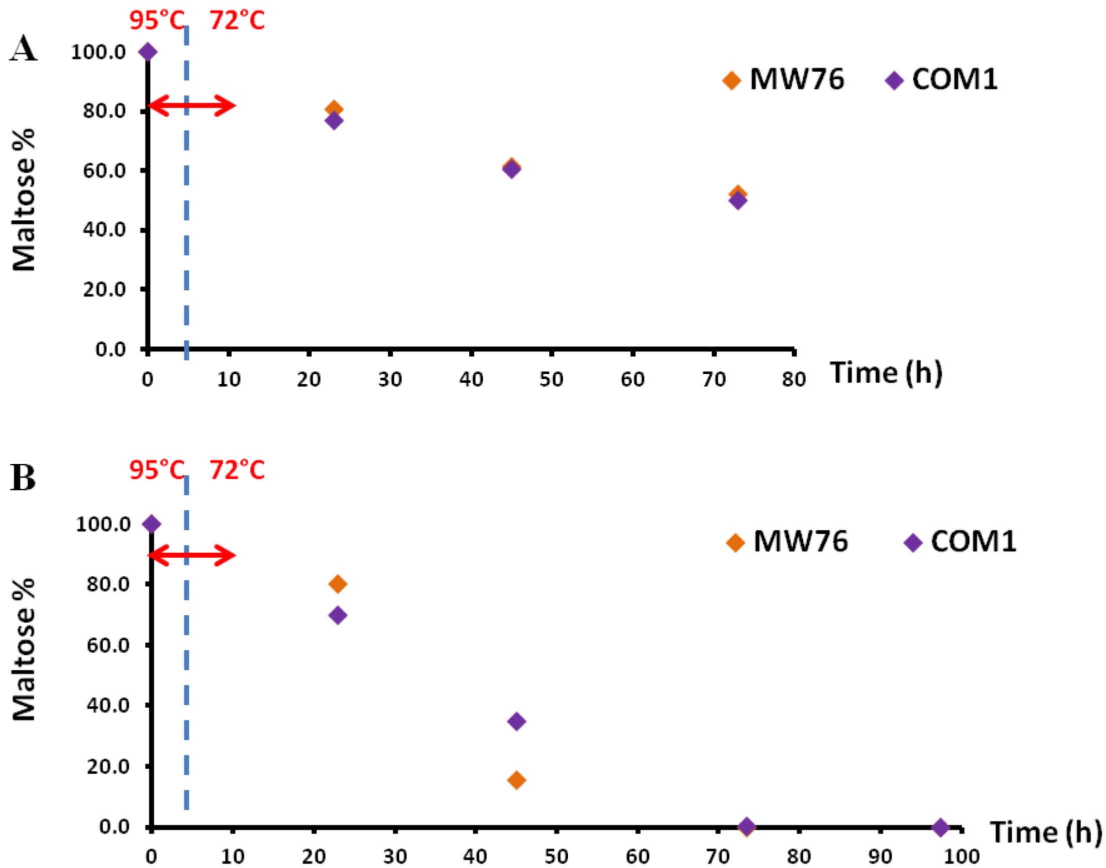


Figure 5. Maltose consumption for MW 56 and COM1: (A) gas feed at 15 mL/min), agitation at 250 rpm, no pH control; (B) gas feed at 50 mL/min, agitation at 400 rpm, pH controlled at 5.5-6.5 by addition of 10 M NaOH.

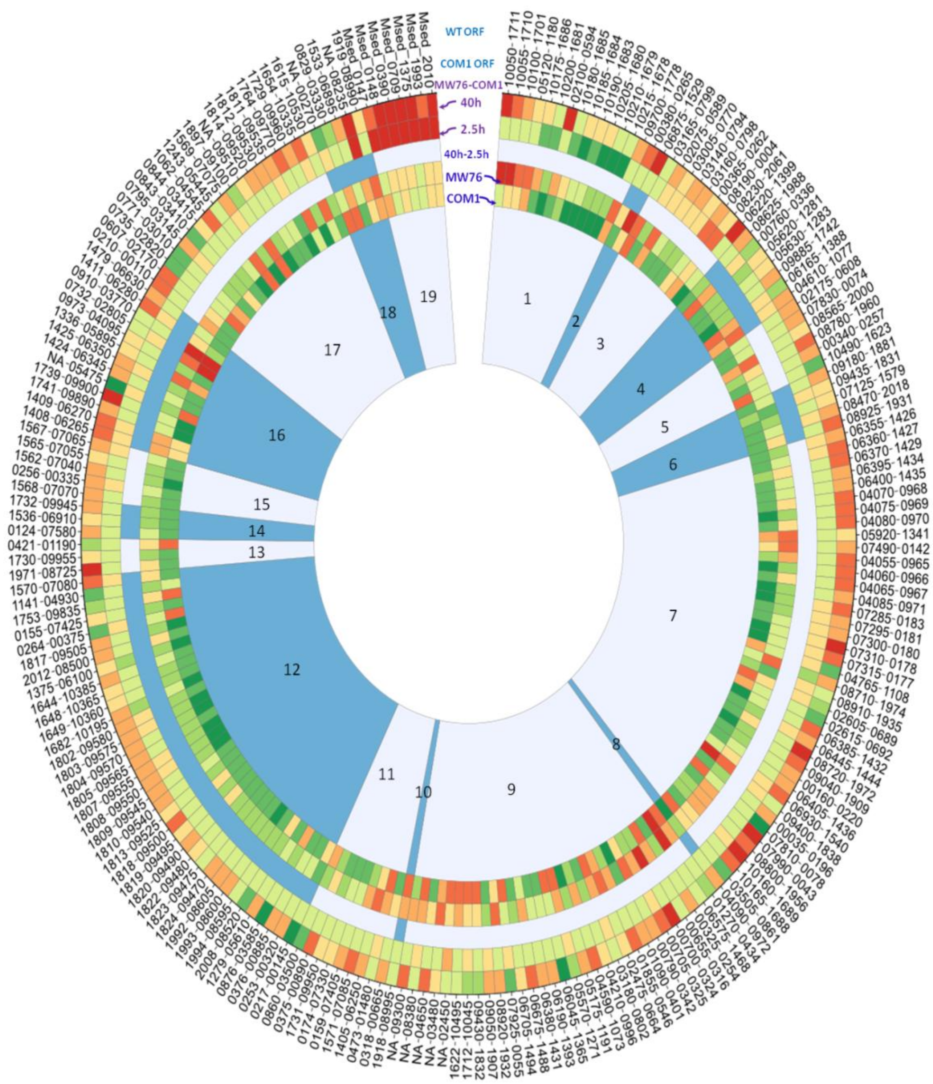


Figure 6. Circos plot generated from transcriptomic analysis for MW76 and COM1 grown with gas feed at 50 mL/min, agitation at 400 rpm, pH controlled at 5.5-6.5. ORFs were grouped as: 1. Amino acid biosynthesis 2. Biosynthesis of cofactors, prosthetic groups, and carriers 3. Cell envelope 4. Cellular processes 5. Central intermediary metabolism 6. DNA metabolism 7. Energy metabolism 8. Fatty acid and phospholipid metabolism 9. Hypothetical proteins 10. Mobile and extrachromosomal element functions 11. Protein fate 12. Protein synthesis 13. Purines, pyrimidines, nucleosides, and nucleotides 14. Regulatory functions 15. Transcription 16. Transport and binding proteins 17. Unclassified 18. Unknown function 19. Foreign genes

APPENDICES

Table S1: Pfu COM1 (72°C)2.5h vs 40h transcriptome				
Pfu ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change COM1 (2.5h)-(40h)**	**-log 10(p value)**
PF0011	PFC_08155	ABC transporter transmembrane component	-1.7	8.0
PF0012	PFC_08150	hypothetical protein	-1.4	6.9
PF0020	PFC_08115	hypothetical protein	1.2	9.4
PF0021	PFC_08110	DNA repair and recombination protein RadB	-1.3	8.9
PF0024	PFC_08090	glycerate kinase	1.1	7.6
PF0025	PFC_08085	hypothetical protein	-1.6	10.2
PF0028	PFC_08070	acetyl transferase	-1.4	10.2
PF0031	PFC_08055	threonine synthase	-1.6	5.5
PF0034	PFC_08035	hypothetical protein	-1.2	7.7
PF0036	PFC_08025	daunorubicin resistance membrane protein	-1.3	6.1
PF0043	PFC_07990	phosphoenolpyruvate synthase	2.7	24.4
PF0055	PFC_07925	HTH transcription regulator	-2.6	13.3
PF0062	PFC_07885	riboflavin-specific deaminase	-1.3	7.1
PF0072	PFC_07840	hypothetical protein	-1.6	6.1
PF0074	PFC_07830	alcohol dehydrogenase	2.1	11.2
PF0078	PFC_07810	2-phosphoglycerate kinase	-2.3	9.3
PF0086	PFC_07775	hypothetical protein	1.0	8.1
PF0090	PFC_07755	molybdenum cofactor biosynthesis protein A	-1.7	13.0
PF0096	PFC_07725	hypothetical protein	1.8	8.6
PF0098	PFC_07715	NAD synthetase	1.5	6.0
PF0101	PFC_07700	hypothetical protein	-1.1	7.8
PF0107	PFC_07670	hypothetical protein	-1.8	11.0
PF0114	PFC_07635	GTP-binding protein	-1.2	8.2
PF0116	PFC_07625	multiple sugar transport protein	-1.1	7.2
PF0118	PFC_07615	ABC transporter permease	-1.8	6.4
PF0124	PFC_07580	hypothetical protein	2.1	9.2
PF0126	PFC_07570	DNA repair protein rad25	1.5	6.8
PF0133	PFC_07535	hypothetical protein	1.5	7.3
PF0137	PFC_07515	methyltransferase	-1.1	5.5

Table S1 Continued				
Pfu ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change COM1 (2.5h)-(40h)**	**-log 10(p value)**
PF0145	PFC_07475	hypothetical protein	-1.5	11.0
PF0146	PFC_07470	potassium channel	-1.4	10.1
PF0155	PFC_07425	asparaginyl-tRNA ligase	2.2	6.1
PF0159	PFC_07405	proteasome subunit beta	2.1	11.1
PF0160	PFC_07400	TLDD-like protein	-1.1	6.2
PF0171	PFC_07345	hypothetical protein	-1.9	13.6
PF0173	PFC_07335	preprotein translocase subunit SecF	-1.1	8.0
PF0174	PFC_07330	preprotein translocase subunit SecD	2.2	12.3
PF0175	PFC_07325	trk system potassium uptake-like protein	1.1	7.0
PF0177	PFC_07315	V-type ATP synthase subunit I	3.0	16.3
PF0178	PFC_07310	V-type ATP synthase subunit K	3.6	13.8
PF0179	PFC_07305	V-type ATP synthase subunit E	1.6	6.6
PF0180	PFC_07300	V-type ATP synthase subunit C	2.6	16.6
PF0181	PFC_07295	V-type ATP synthase subunit F	2.4	15.7
PF0183	PFC_07285	V-type ATP synthase subunit B	3.3	12.6
PF0196	PFC_00035	glucose-6-phosphate isomerase	2.9	17.3
PF0207	PFC_00095	argininosuccinate synthase	1.8	8.9
PF0209	PFC_00105	ribosomal protein s6 modification protein	-1.5	6.1
PF0210	PFC_00110	oxidative cyclase	2.2	11.6
PF0211	PFC_00115	hypothetical protein	-1.9	10.3
PF0215	PFC_00135	enolase	1.7	14.9
PF0216	PFC_00140	Maf-like protein	-1.3	7.2
PF0220	PFC_00160	D-arabino 3-hexulose 6-phosphate formaldehyde lyase	2.7	14.7
PF0224	PFC_00185	hypothetical protein	-1.1	9.7
PF0231	PFC_00205	hypothetical protein	1.3	7.9
PF0239	PFC_00245	hypothetical protein	1.5	8.2
PF0241	PFC_00255	tryptophanyl-tRNA ligase	-1.2	8.5
PF0247	PFC_00285	HTH transcription regulator	-1.2	8.8
PF0250	PFC_00300	transcriptional regulator	-1.2	15.6
PF0252	PFC_00315	30S ribosomal protein S27ae	1.9	10.8

Table S1 Continued				
Pfu ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change COM1 (2.5h)-(40h)**	**-log 10(p value)**
PF0253	PFC_00320	30S ribosomal protein S24e	2.1	12.3
PF0254	PFC_00325	hypothetical protein	2.8	18.0
PF0255	PFC_00330	DNA-directed RNA polymerase subunit E"	2.0	9.6
PF0256	PFC_00335	DNA-directed RNA polymerase subunit E'	2.2	11.5
PF0257	PFC_00340	inorganic pyrophosphatase	2.5	12.5
PF0262	PFC_00365	MFS family transporter	2.1	12.5
PF0264	PFC_00375	histidyl-tRNA ligase	-2.2	9.9
PF0265	PFC_00380	2-amino-3-ketobutyrate coenzyme A ligase	2.2	14.2
PF0272	PFC_00425	alpha-amylase	1.4	6.1
PF0284	PFC_00485	hypothetical protein	-1.6	10.5
PF0290	PFC_00515	valyl-tRNA ligase	-1.2	6.3
PF0291	PFC_00520	prephenate dehydratase	-1.1	7.0
PF0293	PFC_00530	aminotransferase	1.0	9.0
PF0302	PFC_00585	co-factor modifying protein	2.0	13.4
PF0303	PFC_00590	hypothetical protein	-1.2	11.1
PF0310	PFC_00630	daunorubicin resistance ATP-binding protein drrA	1.5	5.9
PF0311	PFC_00635	hypothetical protein	1.2	6.2
PF0316	PFC_00655	hypothetical protein	-2.1	17.0
PF0318	PFC_00665	acylaminoacyl peptidase	2.0	16.9
PF0324	PFC_00700	hypothetical protein	1.1	8.8
PF0332	PFC_00740	flagellar accessory protein FlaH	-1.4	9.4
PF0335	PFC_00755	flagella-related protein D	-1.3	8.6
PF0336	PFC_00760	hypothetical protein	-2.8	23.3
PF0337	PFC_00765	flagellin	-1.5	10.9
PF0340	PFC_00780	HTH transcription regulator	-1.1	6.5
PF0341	PFC_00785	hypothetical protein	-1.2	9.6
PF0363	PFC_00950	beta-galactosidase	-1.0	6.4
PF0365	PFC_00940	hypothetical protein	-1.4	6.6
PF0374	PFC_00895	translation-associated GTPase	-1.3	6.4

Table S1 Continued				
Pfu ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change COM1 (2.5h)-(40h)**	**-log 10(p value)**
PF0375	PFC_00890	prefoldin subunit alpha	1.4	7.3
PF0376	PFC_00885	50S ribosomal protein LX	2.0	16.1
PF0386	PFC_00835	hypothetical protein	1.5	8.6
PF0393	PFC_01050	hypothetical protein	-1.5	9.2
PF0400	PFC_01085	hypothetical protein	-1.4	6.7
PF0401	PFC_01090	methyltransferase	4.1	19.7
PF0402	PFC_01100	UDP-or dTTP-glucose 4-epimerase or 4-6-dehydratase	-1.0	6.0
PF0412	PFC_01145	hypothetical protein	1.3	6.1
PF0413	PFC_01150	hypothetical protein	1.7	11.2
PF0419	PFC_01180	hypothetical protein	1.3	7.2
PF0421	PFC_01190	5-formaminoimidazole-4-carboxamide-1-(beta)-D-ribofuranosyl 5'-monophosphate synthetase-like protein	-2.1	12.5
PF0426	PFC_01215	phosphoribosylaminoimidazole carboxylase ATPase subunit	1.5	9.7
PF0430	PFC_01250	phosphoribosylglycinamide formyltransferase 2	-1.4	12.1
PF0433	PFC_01265	methyl-accepting chemotaxis-like protein	-1.1	9.4
PF0434	PFC_01270	hypothetical protein	-2.4	9.7
PF0437	PFC_01290	hypothetical protein	-1.1	8.5
PF0438	PFC_01300	hypothetical protein	-1.3	8.2
PF0439	PFC_01305	hypothetical protein	-1.1	7.2
PF0443	PFC_01330	membrane transport protein	-1.3	7.2
PF0445	PFC_01340	galactokinase	-1.9	9.5
PF0451	PFC_01370	hypothetical protein	-1.1	9.5
PF0456	PFC_01395	carboxypeptidase 1	1.3	9.9
PF0458	PFC_01405	nicotinamide-nucleotide adenylyltransferase	1.1	6.5
PF0460	PFC_01415	hypothetical protein	1.7	6.9
PF0467	PFC_01450	ATP-dependent protease LA	1.0	6.9
PF0472	PFC_01475	hypothetical protein	-1.2	6.0

Table S1 Continued				
Pfu ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change COM1 (2.5h)-(40h)**	**-log 10(p value)**
PF0473	PFC_01480	serine/threonine protein kinase	-2.0	15.5
PF0493	PFC_01580	hypothetical protein	-1.5	11.8
PF0496	PFC_01595	hypothetical protein	1.8	12.5
PF0508	PFC_01660	integral membrane glycosyltransferase	-1.9	13.8
PF0510	PFC_01670	nucleotide pyrophosphohydrolase	1.0	8.4
PF0511	PFC_01675	3-methyladenine DNA glycosylase	1.1	8.0
PF0515	PFC_01695	hypothetical protein	-1.1	7.3
PF0522	PFC_01730	aspartate aminotransferase	1.2	7.7
PF0523	PFC_01735	KEOPS complex Cgi121-like subunit	-1.2	8.9
PF0533	PFC_01785	indolepyruvate ferredoxin oxidoreductase subunit alpha	1.1	6.3
PF0536	PFC_01800	transposase	-1.2	5.9
PF0545	PFC_01850	hypothetical protein	1.1	7.0
PF0546	PFC_01855	hypothetical protein	1.0	7.1
PF0557	PFC_01915	3-ketoacyl-(acyl-carrier-protein) reductase	1.7	8.0
PF0560	PFC_01930	hypothetical protein	-1.7	9.5
PF0563	PFC_01945	hypothetical protein	-1.0	5.6
PF0565	PFC_01955	hypothetical protein	-1.4	6.3
PF0566	PFC_01960	hypothetical protein	-1.2	8.6
PF0567	PFC_01965	hypothetical protein	-1.6	16.2
PF0581	PFC_02035	hypothetical protein	-2.0	18.2
PF0585	PFC_02055	DEXX-box atpase	-1.1	5.6
PF0589	PFC_02075	mannose-6-phosphate isomerase	2.6	11.4
PF0591	PFC_02085	mannosyl-3-phosphoglycerate synthase	1.5	7.4
PF0594	PFC_02100	ornithine carbamoyltransferase	2.8	11.5
PF0596	PFC_02110	mRNA 3'-end processing factor	1.3	11.8
PF0607	PFC_02170	hypothetical protein	2.1	16.4
PF0608	PFC_02175	Fe-containing alcohol dehydrogenase	2.4	10.7
PF0613	PFC_02205	fructose-1,6-bisphosphatase	-1.3	6.5

Table S1 Continued				
Pfu ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change COM1 (2.5h)-(40h)**	**-log 10(p value)**
PF0615	PFC_02215	hydrogenase nickel incorporation protein	1.6	15.2
PF0631	PFC_02290	radical SAM domain-containing protein	-1.3	18.4
PF0641	PFC_02340	hypothetical protein	1.2	9.6
PF0646	PFC_02370	hypothetical protein	1.5	9.6
PF0651	PFC_02395	hypothetical protein	1.8	12.5
PF0655.1n	PFC_02420	hypothetical protein	-1.3	7.9
PF0663	PFC_02470	hypothetical protein	-1.0	5.5
PF0669	PFC_02500	hypothetical protein	1.3	6.0
PF0680	PFC_02560	hypothetical protein	-1.1	6.9
PF0687	PFC_02590	transcription initiation factor IIB chain b	-1.3	8.7
PF0689	PFC_02605	hypothetical protein	-2.3	15.5
PF0693	PFC_02620	hypothetical protein	-1.6	5.8
PF0696	PFC_02635	carbohydrate-binding protein	-1.1	7.0
PF0702	PFC_02655	x-pro dipeptidase	-1.1	6.4
PF0708	PFC_02690	hypothetical protein	1.1	6.2
PF0718	PFC_02735	hypothetical protein	1.8	8.9
PF0725	PFC_02770	hypothetical protein	1.4	6.6
PF0727	PFC_02780	hypothetical protein	1.1	8.1
PF0728	PFC_02785	hypothetical protein	-1.3	9.4
PF0731	PFC_02800	copper-transporting atpase, p-type	-1.2	9.7
PF0732	PFC_02805	daunorubicin resistance ATP-binding protein	-3.0	11.6
PF0733	PFC_02810	ABC transporter	-2.0	13.6
PF0735	PFC_02820	hypothetical protein	-2.1	13.0
PF0736	PFC_02825	hypothetical protein	1.3	6.4
PF0743	PFC_02875	hypothetical protein	-1.0	8.6
PF0744	PFC_02880	ABC transporter	-1.1	6.4
PF0751	PFC_02915	flavoprotein	2.0	12.6
PF0753	PFC_02925	2-oxoacid ferredoxin oxidoreductase subunit beta	1.9	11.6

Table S1 Continued				
Pfu ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change COM1 (2.5h)-(40h)**	**-log 10(p value)**
PF0754	PFC_02930	2-oxoacid:ferredoxin oxidoreductase subunit alpha	2.0	11.9
PF0759	PFC_02940	hypothetical protein	-1.3	6.9
PF0763	PFC_02965	hypothetical protein	1.5	12.3
PF0768	PFC_02995	acetyl / acyl transferase-like protein	1.5	5.9
PF0770	PFC_03005	glucose-1-phosphate uridylyltransferase	-1.8	10.7
PF0771	PFC_03010	NDP-sugar dehydrogenase	-1.3	10.4
PF0781	PFC_03060	hypothetical protein	-1.2	7.2
PF0794	PFC_03140	UDP-n-acetylglucosamine 2-epimerase	2.1	12.6
PF0795	PFC_03145	capsular polysaccharide biosynthesis protein	2.5	12.3
PF0796	PFC_03150	galactoside o-acetyltransferase	1.8	8.0
PF0798	PFC_03160	glycosyl transferase family protein	2.1	5.8
PF0799	PFC_03165	hypothetical protein	-2.5	8.8
PF0802	PFC_03180	hypothetical protein	-2.1	13.6
PF0812	PFC_03235	hypothetical protein	-1.6	5.9
PF0815	PFC_03250	hypothetical protein	-1.4	8.2
PF0823	PFC_03300	hypothetical protein	-1.7	9.1
PF0825	PFC_03310	prolyl endopeptidase	1.5	7.6
PF0827	PFC_03320	RNA methyltransferase	-1.0	6.6
PF0829	PFC_03330	hypothetical protein	-2.6	14.7
PF0831	PFC_03340	hypothetical protein	-1.2	9.0
PF0833	PFC_03350	low temperature requirement C-like protein	-1.0	6.0
PF0836	PFC_03365	cobalamin synthase	-1.5	5.6
PF0844	PFC_03415	hypothetical protein	2.2	17.3
PF0845	PFC_03420	indolepyruvate ferredoxin oxidoreductase subunit alpha	-1.7	17.3
PF0847	PFC_03430	hypothetical protein	-1.2	6.4
PF0854	PFC_03465	endo-1,4-beta-glucanase b	-1.1	7.4

Table S1 Continued				
Pfu ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change COM1 (2.5h)-(40h)**	**-log 10(p value)**
PF0855	PFC_03470	hypothetical protein	1.5	9.2
PF0855.1n	PFC_03480	hypothetical protein	-1.4	13.6
PF0856	PFC_03475	isopentenyl pyrophosphate isomerase	1.0	8.8
PF0860	PFC_03500	btpA family protein	-1.3	6.7
PF0861	PFC_03505	phosphoglucosamine mutase	-2.9	18.1
PF0866	PFC_03540	hypothetical protein	-1.1	8.4
PF0867	PFC_03545	hypothetical protein	-1.1	7.3
PF0871	PFC_03565	hypothetical protein	1.5	7.1
PF0876	PFC_03585	50S ribosomal protein L15e	-1.6	7.3
PF0880	PFC_03600	DEXX-box atpase	-2.0	14.2
PF0881	PFC_03605	ABC transporter	-1.4	8.5
PF0889	PFC_03655	hypothetical protein	-1.2	6.4
PF0894	PFC_03680	sulfhydrogenase subunit alpha	1.0	6.5
PF0900	PFC_03715	hypothetical protein	-1.0	9.7
PF0901	PFC_03725	hypothetical protein	-1.3	6.1
PF0903	PFC_03735	hypothetical protein	1.1	6.2
PF0904	PFC_03740	N-glycosylase/DNA lyase	-1.2	6.6
PF0905	PFC_03745	ABC transporter	-1.5	11.9
PF0909	PFC_03765	ferric enterobactin transport ATP-binding protein	-1.5	8.9
PF0910	PFC_03770	iron ABC transporter permease	-3.5	23.0
PF0911	PFC_03775	iron (III) ABC transporter	-1.3	10.7
PF0914	PFC_03790	hypothetical protein	-1.5	12.1
PF0918	PFC_03810	tRNA pseudouridine synthase A	-1.1	6.8
PF0925	PFC_03845	heme biosynthesis protein	-1.2	7.9
PF0927	PFC_03855	hypothetical protein	-1.1	6.9
PF0928	PFC_03860	hypothetical protein	-1.2	7.5
PF0929	PFC_03865	hypothetical protein	1.7	12.1
PF0930	PFC_03870	hypothetical protein	1.5	10.9
PF0937	PFC_03905	2-isopropylmalate synthase	1.1	6.8
PF0944	PFC_03940	hypothetical protein	-1.4	8.9
PF0950	PFC_03970	hypothetical protein	-1.8	7.7

Table S1 Continued				
Pfu ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change COM1 (2.5h)-(40h)**	**-log 10(p value)**
PF0953	PFC_03985	hypothetical protein	1.1	9.8
PF0957	PFC_04005	hypothetical protein	-1.5	10.0
PF0958	PFC_04010	hypothetical protein	-1.8	12.7
PF0962	PFC_04040	hypothetical protein	1.3	6.0
PF0965	PFC_04055	pyruvate ferredoxin oxidoreductase subunit beta	1.5	9.7
PF0966	PFC_04060	pyruvate ferredoxin oxidoreductase subunit alpha	4.0	24.0
PF0967	PFC_04065	pyuvate ferredoxin oxidoreductase subunit delta	2.5	11.3
PF0968	PFC_04070	2-ketoisovalerate ferredoxin oxidoreductase subunit beta	2.6	10.3
PF0969	PFC_04075	2-ketoisovalerate ferredoxin oxidoreductase subunit alpha	3.7	15.4
PF0970	PFC_04080	2-ketoisovalerate ferredoxin oxidoreductase subunit delta	2.2	10.8
PF0971	PFC_04085	pyruvate/ketoisovalerate ferredoxin oxidoreductase subunit gamma	3.2	15.9
PF0972	PFC_04090	hypothetical protein	3.2	16.8
PF0973	PFC_04095	acetyl-CoA acetyltransferase	2.2	10.5
PF0975	PFC_04105	hypothetical protein	-1.2	5.8
PF0979	PFC_04125	cobalamin adenosyltransferase	1.8	10.2
PF0984	PFC_04150	hypothetical protein	1.9	14.5
PF0994	PFC_04200	type II secretion system protein	1.1	7.1
PF0996	PFC_04210	hypothetical protein	2.1	9.1
PF1001	PFC_04235	oligopeptide transport system permease	-1.3	10.9
PF1005	PFC_04255	hypothetical protein	-1.2	7.4
PF1008	PFC_04270	phosphate ABC transporter ATP-binding protein	-1.1	7.8
PF1012	PFC_04290	ABC transporter	-1.3	12.7

Table S1 Continued				
Pfu ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change COM1 (2.5h)-(40h)**	**-log 10(p value)**
PF1015	PFC_04305	transposase b-like protein	-1.6	9.5
PF1032	PFC_04395	hypothetical protein	1.2	10.9
PF1033	PFC_04400	peroxiredoxin	2.0	12.0
PF1044	PFC_04455	hypothetical protein	-1.0	8.8
PF1049	PFC_04480	Pyruvate formate lyase activating like protein	-1.5	8.8
PF1053	PFC_04500	aspartate kinase	1.7	7.2
PF1059	PFC_04530	hypothetical protein	-1.1	6.5
PF1062	PFC_04545	hypothetical protein	3.4	13.4
PF1072	PFC_04585	hypothetical protein	-1.0	5.7
PF1075	PFC_04600	hypothetical protein	-1.3	11.5
PF1077	PFC_04610	hypothetical protein	-1.1	6.2
PF1080	PFC_04625	hypothetical protein	-1.3	7.0
PF1081	PFC_04630	hypothetical protein	-1.5	9.6
PF1083	PFC_04635	hypothetical protein	-1.2	6.2
PF1085.1n	PFC_04650	hypothetical protein	2.5	10.3
PF1107	PFC_04760	hit family protein	-1.3	6.6
PF1108	PFC_04765	alpha-dextrin endo-1, 6-alpha-glucosidase	-1.6	9.7
PF1115	PFC_04795	RNA-binding protein	1.0	7.1
PF1118	PFC_04810	CRISPR-associated protein Cas1	-1.1	6.6
PF1122	PFC_04830	hypothetical protein	-1.8	12.9
PF1123	PFC_04835	hypothetical protein	1.2	5.6
PF1128	PFC_04860	hypothetical protein	1.9	15.8
PF1130	PFC_04870	hypothetical protein	1.1	7.7
PF1144	PFC_04945	hypothetical protein	-1.0	7.3
PF1146	PFC_04955	hypothetical protein	-1.1	6.7
PF1178	PFC_05110	Na-dependent transporter	-1.0	6.6
PF1190	PFC_05170	hypothetical protein	-1.2	11.0
PF1191	PFC_05175	hypothetical protein	3.3	14.7
PF1206	PFC_05240	hypothetical protein	1.2	8.6
PF1209	PFC_05255	oligopeptide ABC transporter	1.1	8.2

Table S1 Continued				
Pfu ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change COM1 (2.5h)-(40h)**	**-log 10(p value)**
PF1214.2n	PFC_05285	hypothetical protein	-1.5	7.3
PF1218.1n	PFC_05315	hypothetical protein	-1.2	7.6
PF1219	PFC_05310	hypothetical protein	-1.2	8.5
PF1221	PFC_05325	dissimilatory sulfate adenylyltransferase	-1.4	10.4
PF1232	PFC_05380	4-aminobutyrate aminotransferase	-1.1	6.9
PF1239	PFC_05415	hypothetical protein	-1.2	7.4
PF1242	PFC_05440	dehydrogenase subunit alpha	-1.3	7.9
PF1243	PFC_05445	hypothetical protein	-2.4	6.3
PF1258	PFC_05495	ribose-5-phosphate isomerase A	1.3	6.1
PF1259	PFC_05500	hypothetical protein	1.4	8.4
PF1263	PFC_05530	hypothetical protein	1.1	5.5
PF1269	PFC_05560	methionine synthase	-1.7	13.7
PF1271	PFC_05570	hypothetical protein	-2.4	13.6
PF1279	PFC_05610	50S ribosomal protein L10e	2.6	16.2
PF1281	PFC_05620	superoxide reductase	3.0	25.0
PF1283	PFC_05630	rubrerythrin	2.4	7.9
PF1284	PFC_05635	hypothetical protein	1.4	10.5
PF1289	PFC_05660	molybdopterin biosynthesis protein moeb	1.3	7.9
PF1291	PFC_05670	hypothetical protein	1.2	8.3
PF1293	PFC_05680	prolyl-tRNA ligase	1.5	9.2
PF1295	PFC_05690	transcription factor	1.1	8.4
PF1299	PFC_05710	pyrrolidone-carboxylate peptidase	1.1	8.9
PF1310	PFC_05760	hypothetical protein	-1.8	11.6
PF1312	PFC_05770	hypothetical protein	-1.3	7.7
PF1313	PFC_05775	DNA processing SMF protein	-1.8	7.6
PF1314	PFC_05780	hypothetical protein	-1.2	7.2
PF1336	PFC_05895	cytosine permease	-1.4	9.6
PF1338	PFC_05905	transcriptional activator	-1.9	12.3
PF1341	PFC_05920	glycine cleavage system aminomethyltransferase T	-1.8	6.0
PF1344	PFC_05935	maleate cis-trans isomerase	1.2	6.1

Table S1 Continued				
Pfu ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change COM1 (2.5h)-(40h)**	**-log 10(p value)**
PF1350	PFC_05965	hypothetical protein	-1.3	11.7
PF1362	PFC_06030	hypothetical protein	-1.0	7.1
PF1365	PFC_06045	hypothetical protein	-2.2	12.3
PF1367	PFC_06055	50S ribosomal protein L7Ae	1.3	12.2
PF1368	PFC_06060	30S ribosomal protein S28e	1.9	18.9
PF1371	PFC_06075	hypothetical protein	-1.5	8.5
PF1375	PFC_06100	elongation factor 1-alpha	3.2	16.9
PF1376	PFC_06105	30S ribosomal protein S10P	1.2	8.1
PF1379	PFC_06120	cell division protein pelota	1.6	5.9
PF1383	PFC_06140	hypothetical protein	1.6	7.5
PF1388	PFC_06165	hypothetical protein	-2.9	18.5
PF1390	PFC_06175	hypothetical protein	-1.4	7.4
PF1393	PFC_06190	hypothetical protein	2.3	12.5
PF1399	PFC_06220	ATPase, vanadate-sensitive	4.3	33.9
PF1401	PFC_06230	peptidyl-prolyl cis-trans isomerase	1.5	9.9
PF1410	PFC_06275	binding-protein-dependent transport systems inner membrane component	1.5	8.3
PF1411	PFC_06280	dipeptide transport ATP-binding protein	1.8	10.7
PF1423	PFC_06340	monovalent cation/H+ antiporter subunit E	1.6	12.3
PF1424	PFC_06345	monovalent cation/H+ antiporter subunit F	2.3	8.3
PF1425	PFC_06350	monovalent cation/H+ antiporter subunit G	2.7	11.7
PF1426	PFC_06355	hypothetical protein	2.4	7.6
PF1427	PFC_06360	hypothetical protein	2.6	13.3
PF1429	PFC_06370	monovalent cation/H+ antiporter subunit C	1.9	13.9
PF1431	PFC_06380	hypothetical protein	2.7	12.6
PF1432	PFC_06385	mbh10 NADH dehydrogenase subunit	3.1	10.8
PF1434	PFC_06395	membrane bound hydrogenase alpha	2.3	19.2

Table S1 Continued				
Pfu ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change COM1 (2.5h)-(40h)**	**-log 10(p value)**
PF1435	PFC_06400	mbh13 NADH dehydrogenase subunit	2.1	7.9
PF1436	PFC_06405	NADH-plastoquinone oxidoreductase subunit	2.2	12.3
PF1444	PFC_06445	NADH dehydrogenase subunit B	2.5	11.2
PF1456	PFC_06515	hypothetical protein	-1.5	9.6
PF1460	PFC_06535	hypothetical protein	-1.1	5.7
PF1468	PFC_06575	hypothetical protein	-3.0	22.8
PF1478	PFC_06625	hypothetical protein	-1.1	7.3
PF1479	PFC_06630	oxidoreductase, Fe-S subunit	2.5	14.0
PF1480	PFC_06635	formaldehyde:ferredoxin oxidoreductase wor5	1.9	10.0
PF1484	PFC_06655	hypothetical protein	-1.2	7.6
PF1494	PFC_06705	hypothetical protein	2.3	8.6
PF1496	PFC_06715	hypothetical protein	-1.6	11.2
PF1505	PFC_06760	hypothetical protein	1.8	7.8
PF1519	PFC_06830	sugar-binding transport ATP-binding protein	-1.5	8.1
PF1521	PFC_06840	formate dehydrogenase subunit alpha	-1.1	8.2
PF1522	PFC_06845	ABC transporter	-1.2	6.9
PF1531	PFC_06885	phosphomethylpyrimidine synthase ThiC	-1.1	8.8
PF1533	PFC_06895	nodulation protein nfeD	2.3	8.0
PF1534	PFC_06900	stomatin	1.9	13.3
PF1536	PFC_06910	RNase III inhibitor	2.2	8.2
PF1539	PFC_06925	dihydroorotate dehydrogenase 1B	1.2	8.9
PF1541	PFC_06935	50S ribosomal protein L37e	-1.0	9.0
PF1546	PFC_06960	hypothetical protein	1.2	6.1
PF1547	PFC_06965	endoglucanase	1.7	6.5
PF1549	PFC_06975	RNA 3'-terminal-phosphate cyclase	1.0	9.5
PF1551	PFC_06985	hypothetical protein	-1.3	8.1

Table S1 Continued				
Pfu ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change COM1 (2.5h)-(40h)**	**-log 10(p value)**
PF1558	PFC_07020	30S ribosomal protein S7P	1.2	7.5
PF1561	PFC_07035	50S ribosomal protein L30e	1.9	9.7
PF1562	PFC_07040	DNA-directed RNA polymerase subunit A"	3.0	16.8
PF1565	PFC_07055	DNA-directed RNA polymerase subunit H	2.1	14.9
PF1567	PFC_07065	exosome complex RNA-binding protein Rrp42	2.1	14.3
PF1568	PFC_07070	exosome complex exonuclease Rrp41	2.9	18.6
PF1569	PFC_07075	exosome complex RNA-binding protein Rrp4	3.2	14.7
PF1570	PFC_07080	RNA-associated protein	2.1	11.8
PF1571	PFC_07085	proteasome subunit alpha	1.5	11.0
PF1579	PFC_07125	DNA topoisomerase VI subunit B	2.3	8.2
PF1580	PFC_07130	RNA-processing protein	1.8	8.1
PF1581	PFC_07135	rio1 protein	1.8	9.3
PF1583	PFC_07145	endopeptidase IV	1.2	10.2
PF1611	PFC_10555	cell division inhibitor minD	1.9	8.1
PF1615	PFC_10530	hypothetical protein	2.3	17.0
PF1619	PFC_10510	hypothetical protein	-1.3	9.5
PF1622	PFC_10495	n-type ATP pyrophosphatase superfamily protein	-2.2	11.2
PF1623	PFC_10490	pyruvoyl-dependent arginine decarboxylase	-2.2	11.5
PF1624	PFC_10485	hypothetical protein	-1.5	8.0
PF1627	PFC_10470	glycyl-tRNA ligase	1.8	8.5
PF1632	PFC_10445	hypothetical protein	-1.4	7.6
PF1641	PFC_10400	hypothetical protein	1.0	7.7
PF1642	PFC_10395	DNA-directed RNA polymerase subunit K	1.9	7.8
PF1644	PFC_10385	30S ribosomal protein S9P	2.3	17.0

Table S1 Continued				
Pfu ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change COM1 (2.5h)-(40h)**	**-log 10(p value)**
PF1645	PFC_10380	50S ribosomal protein L13P	1.5	6.3
PF1646	PFC_10375	50S ribosomal protein L18e	1.9	11.2
PF1647	PFC_10370	DNA-directed RNA polymerase subunit D	1.6	6.2
PF1648	PFC_10365	30S ribosomal protein S11P	2.8	17.0
PF1649	PFC_10360	30S ribosomal protein S4P	2.1	15.4
PF1650	PFC_10355	30S ribosomal protein S13P	1.1	11.8
PF1651	PFC_10350	hypothetical protein	-1.2	6.2
PF1656	PFC_10325	NADH oxidase	-1.8	11.1
PF1657	PFC_10320	histidyl-tRNA synthetase	1.4	11.4
PF1659	PFC_10310	bifunctional histidinal dehydrogenase/ histidinol dehydrogenase	1.6	5.7
PF1660	PFC_10305	imidazoleglycerol-phosphate dehydratase	1.4	12.0
PF1662	PFC_10295	1-(5-phosphoribosyl)-5-[(5-phosphoribosylamino)methylideneamino] imidazole-4-carboxamide isomerase	-1.5	5.6
PF1671	PFC_10250	hypothetical protein	-1.4	7.7
PF1672	PFC_10245	hypothetical protein	-1.2	6.0
PF1677	PFC_10220	Biotin biosynthesis protein BioY	-1.7	12.6
PF1678	PFC_10215	2-isopropylmalate synthase	4.8	17.6
PF1679	PFC_10210	3-isopropylmalate dehydratase large subunit	5.1	23.7
PF1680	PFC_10205	3-isopropylmalate dehydratase small subunit	4.3	25.0
PF1681	PFC_10200	lysine biosynthesis protein	4.3	16.5
PF1682	PFC_10195	ribosomal protein s6 modification protein	3.4	20.7
PF1683	PFC_10190	N-acetyl-gamma-glutamyl-phosphate reductase	3.2	12.4

Table S1 Continued				
Pfu ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change COM1 (2.5h)-(40h)**	**-log 10(p value)**
PF1684	PFC_10185	acetylglutamate/acetylaminoadipate kinase	3.5	19.8
PF1686	PFC_10175	acetyl-lysine deacetylase	2.1	9.6
PF1687	PFC_10170	hypothetical protein	1.4	6.3
PF1688	PFC_10165	transketolase	2.2	15.3
PF1690	PFC_10155	3-deoxy-7-phosphoheptulonate synthase	1.7	13.6
PF1691	PFC_10150	3-dehydroquinate synthase	2.0	13.4
PF1702	PFC_10095	aspartate aminotransferase	1.7	9.2
PF1704	PFC_10085	hypothetical protein	1.1	5.7
PF1706	PFC_10075	tryptophan synthase subunit beta	1.2	7.9
PF1708	PFC_10065	anthranilate synthase component II	1.2	7.0
PF1712	PFC_10045	hypothetical protein	-2.6	10.9
PF1717	PFC_10020	translation initiation factor IF-2 subunit gamma	1.6	6.9
PF1729	PFC_09960	phospho-sugar mutase	3.1	11.5
PF1730	PFC_09955	thymidylate kinase	2.4	10.4
PF1731	PFC_09950	signal recognition particle protein Srp54	2.3	12.2
PF1732	PFC_09945	AsnC family transcriptional regulator	2.3	10.1
PF1735.2n	PFC_09925	hypothetical protein	1.9	14.0
PF1737	PFC_09915	hypothetical protein	-1.2	6.4
PF1739	PFC_09900	trehalose/maltose binding protein	5.2	21.8
PF1741	PFC_09890	trehalose/maltose transport inner membrane protein	3.0	9.3
PF1742	PFC_09885	group 1 glycosyl transferase	3.1	14.5
PF1743	PFC_09880	hypothetical protein	1.6	8.9
PF1746	PFC_09865	hypothetical protein	1.1	6.9
PF1747	PFC_09860	hypothetical protein	1.5	7.5
PF1750	PFC_09850	sugar-binding transport ATP-binding protein	-2.0	14.2
PF1753	PFC_09835	glutamyl-tRNA ligase	-2.0	10.9

Table S1 Continued				
Pfu ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change COM1 (2.5h)-(40h)**	**-log 10(p value)**
PF1755	PFC_09825	fumarate hydratase	1.6	15.5
PF1770	PFC_09740	2-oxoglutarate ferredoxin oxidoreductase subunit gamma	1.3	7.2
PF1771	PFC_09735	2-oxoglutarate ferredoxin oxidoreductase subunit alpha	1.6	7.9
PF1772	PFC_09730	2-oxoglutarate ferredoxin oxidoreductase subunit beta	1.8	10.3
PF1773	PFC_09725	2-oxoglutarate ferredoxin oxidoreductase subunit gamma	1.9	8.6
PF1774	PFC_09720	iron ABC transporter ATP-binding protein	1.7	7.4
PF1783	PFC_09675	molybdenum cofactor biosynthesis protein MoeA	1.3	6.7
PF1787	PFC_09655	acetyl-CoA synthetase I subunit beta	1.8	6.0
PF1792	PFC_09630	5'-cyclic-nucleotide phosphodiesterase	-1.2	7.4
PF1796	PFC_09610	polyferredoxin	-1.8	16.7
PF1797	PFC_09605	d-nopaline dehydrogenase	1.1	7.7
PF1801	PFC_09585	preprotein translocase subunit SecY	1.3	6.2
PF1802	PFC_09580	50S ribosomal protein L15P	2.9	12.4
PF1803	PFC_09575	50S ribosomal protein L30P	3.1	18.3
PF1804	PFC_09570	30S ribosomal protein S5P	3.6	18.8
PF1805	PFC_09565	50S ribosomal protein L18P	2.5	11.8
PF1806	PFC_09560	50S ribosomal protein L19e	1.9	7.4
PF1807	PFC_09555	50S ribosomal protein L32e	3.2	17.7
PF1808	PFC_09550	50S ribosomal protein L6P	3.0	15.2
PF1809	PFC_09545	30S ribosomal protein S8P	3.2	14.4
PF1810	PFC_09540	30S ribosomal protein S14P	2.0	9.3
PF1811	PFC_09535	50S ribosomal protein L5P	3.1	18.9
PF1812	PFC_09530	30S ribosomal protein S4e	2.5	16.9
PF1813	PFC_09525	50S ribosomal protein L24P	2.0	14.7

Table S1 Continued				
Pfu ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change COM1 (2.5h)-(40h)**	**-log 10(p value)**
PF1814	PFC_09520	50S ribosomal protein L14P	2.4	20.2
PF1815	PFC_09515	30S ribosomal protein S17P	1.6	14.6
PF1817	PFC_09505	translation initiation factor Sui1	3.6	14.3
PF1818	PFC_09500	50S ribosomal protein L29P	2.3	17.1
PF1819	PFC_09495	30S ribosomal protein S3P	2.8	14.8
PF1820	PFC_09490	50S ribosomal protein L22P	2.1	11.6
PF1821	PFC_09485	30S ribosomal protein S19P	2.0	18.2
PF1822	PFC_09480	50S ribosomal protein L2P	2.5	17.1
PF1823	PFC_09475	50S ribosomal protein L23P	2.4	12.7
PF1824	PFC_09470	50S ribosomal protein L4P	2.6	15.6
PF1825	PFC_09465	50S ribosomal protein L3P	1.7	10.5
PF1830	PFC_09440	regulatory protein	-1.1	6.8
PF1831	PFC_09435	histone a1	1.9	22.4
PF1832	PFC_09430	GTP cyclohydrolase	-2.0	12.3
PF1838	PFC_09400	hypothetical protein	-3.2	16.7
PF1839	PFC_09395	CTP synthetase	1.7	7.3
PF1844	PFC_09370	hypothetical protein	-1.2	5.9
PF1846	PFC_09360	hypothetical protein	-1.4	10.6
PF1857	PFC_09305	cobalt transport ATP-binding protein	-1.0	5.6
PF1865	PFC_09260	regulatory protein	-1.3	5.9
PF1881	PFC_09180	DNA/RNA-binding protein albA	2.1	12.4
PF1886	PFC_09155	sugar kinase	1.2	6.0
PF1897	PFC_09100	hypothetical protein	-3.0	14.3
PF1899	PFC_09090	2-deoxy-d-gluconate 3-dehydrogenase	-1.1	5.5
PF1905	PFC_09060	protease	-1.6	10.2
PF1906	PFC_09055	adenosylmethionine-8-amino-7-oxononanoate aminotransferase	-1.0	6.2
PF1907	PFC_09050	hypothetical protein	2.6	13.4
PF1909	PFC_09040	ferredoxin	3.5	18.0
PF1913	PFC_09020	hypothetical protein	1.1	6.3

Table S1 Continued				
Pfu ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change COM1 (2.5h)-(40h)**	**-log 10(p value)**
PF1918	PFC_08995	hypothetical protein	-2.2	11.8
PF1926	PFC_08950	DNA repair and recombination protein RadA	1.0	6.4
PF1927	PFC_08945	hypothetical protein	-1.4	5.6
PF1931	PFC_08925	hypothetical protein	2.1	12.2
PF1935	PFC_08910	pullulanase	1.3	7.0
PF1936	PFC_08905	malg-like sugar transport inner membrane protein	1.6	5.8
PF1937	PFC_08900	malf-like sugar transport inner membrane protein	1.8	8.8
PF1938	PFC_08895	malE-like sugar binding protein	1.8	6.3
PF1940	PFC_08885	hypothetical protein	-1.4	7.7
PF1940.1n	PFC_08880	hypothetical protein	-1.2	14.7
PF1947	PFC_08845	hypothetical protein	1.4	8.4
PF1956	PFC_08800	fructose-bisphosphate aldolase	1.9	7.3
PF1960	PFC_08780	aldose reductase	1.9	15.5
PF1961	PFC_08775	aldehyde:ferredoxin oxidoreductase	1.3	10.2
PF1962	PFC_08770	hypothetical protein	-2.0	12.7
PF1970	PFC_08730	multiple sugar-binding transport ATP-binding protein	-1.3	9.8
PF1971	PFC_08725	anaerobic ribonucleoside triphosphate reductase	2.9	18.7
PF1972	PFC_08720	anaerobic ribonucleoside-triphosphate reductase activating protein	2.8	7.6
PF1975	PFC_08705	phosphoenolpyruvate carboxylase	1.3	7.7
PF1987	PFC_08630	hypothetical protein	1.8	12.7
PF1988	PFC_08625	cell division protein FtsZ	2.9	10.1
PF1990	PFC_08615	transcription antitermination protein NusG	1.3	7.8
PF1991	PFC_08610	50S ribosomal protein L11P	1.8	7.2

Table S1 Continued				
Pfu ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change COM1 (2.5h)-(40h)**	**-log 10(p value)**
PF1992	PFC_08605	50S ribosomal protein L1P	2.8	14.2
PF1993	PFC_08600	acidic ribosomal protein P0	3.5	21.9
PF1994	PFC_08595	50S ribosomal protein L12P	2.3	20.2
PF1997	PFC_08580	hypothetical protein	-1.6	5.9
PF1998	PFC_08575	HAD superfamily hydrolase	-1.4	6.4
PF2000	PFC_08565	glycine dehydrogenase subunit 2	-1.7	8.1
PF2005	PFC_08535	glycerol-3-phosphate dehydrogenase	1.1	8.4
PF2007	PFC_08525	hypothetical protein	-1.4	7.0
PF2012	PFC_08500	elongation factor EF-2	2.0	20.5
PF2013	PFC_08495	hypothetical protein	1.9	15.8
PF2018	PFC_08470	hypothetical protein	2.0	10.6
PF2024	PFC_08440	transposase	-1.1	5.8
PF2031	PFC_08410	iron-sulfur cluster binding protein	-1.2	9.5
PF2032	PFC_08405	hypothetical protein	-1.1	7.9
PF2045	PFC_08340	hypothetical protein	-1.9	12.7
PF2046	PFC_08335	hypothetical protein	-1.9	13.5
PF2048	PFC_08320	peptidase	-1.3	8.6
PF2058	PFC_08255	hypothetical protein	-1.4	6.2
PF2059	PFC_08250	aminopeptidase	-1.2	7.4
PF2061	PFC_08230	ABC transporter	-1.4	6.7
PF2064	PFC_08215	arylsulfatase	-1.7	10.3
	PFC_09510	ribonuclease P protein component 1	3.4	14.6
	PFC_01650	hypothetical protein	1.5	7.3
	PFC_01065	hypothetical protein	1.1	5.8
	PFC_00550	hypothetical protein	1.0	11.7
	PFC_04160	haloacid dehalogenase superfamily protein	-1.0	8.5
	PFC_04685	hypothetical protein	-1.1	7.2
	PFC_05515	hypothetical protein	-1.1	7.7
	PFC_08040	permease	-1.3	8.0
	PFC_00275	hypothetical protein	-1.3	7.5
	PFC_08415	ferredoxin-family protein	-1.4	8.6

Table S1 Continued				
Pfu ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change COM1 (2.5h)-(40h)**	**-log 10(p value)**
	PFC_03595	hypothetical protein	-1.4	8.0
	PFC_00270	HEPN domain-containing protein	-2.1	9.7
	PFC_08380	hypothetical protein	-2.2	17.6
	PFC_02450	hypothetical protein	-2.4	13.4
	PFC_05475	sodium dependent transporter	-2.4	15.1

Table S2: Pfu MW76 (72°C)2.5h vs 40h transcriptome				
Pfu ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change MW76 (2.5h)-(40h)**	**-log 10(p value)**
PF0007	PFC_08175	hypothetical protein	-1.4	8.1
PF0011	PFC_08155	ABC transporter transmembrane component	-2.0	9.5
PF0012	PFC_08150	hypothetical protein	-1.2	6.0
PF0014	PFC_08145	hypothetical protein	-1.4	8.0
PF0019	PFC_08120	DNA polymerase II large subunit	-1.5	6.0
PF0021	PFC_08110	DNA repair and recombination protein RadB	-1.6	11.2
PF0024	PFC_08090	glycerate kinase	1.3	9.3
PF0025	PFC_08085	hypothetical protein	-1.3	8.4
PF0028	PFC_08070	acetyl transferase	-1.1	8.0
PF0029	PFC_08065	hypothetical protein	-1.4	10.9
PF0030	PFC_08060	hypothetical protein	1.2	5.9
PF0032	PFC_08045	permease	1.4	6.6
PF0043	PFC_07990	phosphoenolpyruvate synthase	-1.7	18.1
PF0045	PFC_07980	hypothetical protein	-1.1	6.2
PF0071	PFC_07845	hypothetical protein	-1.2	5.7
PF0072	PFC_07840	hypothetical protein	-1.7	6.5
PF0085	PFC_07780	DNA helicase	1.3	6.7
PF0107	PFC_07670	hypothetical protein	-1.1	5.9
PF0130	PFC_07550	hypothetical protein	-1.2	7.7
PF0132	PFC_07540	phosphate transport system substrate-binding protein	-1.3	5.7
PF0142	PFC_07490	asparaginase	-2.1	9.5
PF0144	PFC_07480	hypothetical protein	-1.6	10.5
PF0152	PFC_07440	hypothetical protein	1.1	12.5
PF0162	PFC_07390	hypothetical protein	1.0	8.5
PF0173	PFC_07335	preprotein translocase subunit SecF	-1.1	7.9
PF0181	PFC_07295	V-type ATP synthase subunit F	1.1	8.2
PF0196	PFC_00035	glucose-6-phosphate isomerase	1.2	7.5
PF0207	PFC_00095	argininosuccinate synthase	1.6	8.2

Table S2 Continued				
Pfu ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change MW76 (2.5h)-(40h)**	**-log 10(p value)**
PF0209	PFC_00105	ribosomal protein s6 modification protein	-1.8	7.9
PF0210	PFC_00110	oxidative cyclase	-1.4	8.3
PF0211	PFC_00115	hypothetical protein	-1.8	10.1
PF0215	PFC_00135	enolase	1.2	10.6
PF0220	PFC_00160	D-arabino 3-hexulose 6-phosphate formaldehyde lyase	1.1	5.6
PF0222.1n	PFC_00175	hypothetical protein	-1.5	8.7
PF0248	PFC_00290	hypothetical protein	-1.2	10.2
PF0252	PFC_00315	30S ribosomal protein S27ae	1.2	6.4
PF0254	PFC_00325	hypothetical protein	1.1	7.3
PF0262	PFC_00365	MFS family transporter	2.0	12.1
PF0276	PFC_00445	oxidoreductase	-1.2	6.0
PF0280	PFC_00465	hypothetical protein	-1.8	9.5
PF0282	PFC_00475	hypothetical protein	-1.0	6.6
PF0283	PFC_00480	acylphosphatase	1.0	6.3
PF0291	PFC_00520	prephenate dehydratase	-1.5	10.1
PF0293	PFC_00530	aminotransferase	1.2	10.9
PF0299	PFC_00570	cobalamin synthase	-1.0	6.4
PF0302	PFC_00585	co-factor modifying protein	1.6	10.7
PF0303	PFC_00590	hypothetical protein	-1.0	10.0
PF0312	PFC_00640	ADP-dependent glucokinase	-1.6	7.4
PF0316	PFC_00655	hypothetical protein	-1.8	15.2
PF0321	PFC_00680	inosine monophosphate dehydrogenase-like protein	-1.3	9.7
PF0324	PFC_00700	hypothetical protein	-3.7	23.7
PF0325	PFC_00705	hypothetical protein	-1.7	8.1
PF0332	PFC_00740	flagellar accessory protein FlaH	-1.3	8.5
PF0336	PFC_00760	hypothetical protein	-1.3	13.8
PF0337	PFC_00765	flagellin	-1.0	6.9
PF0340	PFC_00780	HTH transcription regulator	-1.7	10.5
PF0342	PFC_00790	hypothetical protein	-2.0	6.3

Table S2 Continued				
Pfu ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change MW76 (2.5h)-(40h)**	**-log 10(p value)**
PF0349	PFC_01020	hypothetical protein	1.0	7.4
PF0356	PFC_00985	beta-galactosidase	-1.2	5.6
PF0357	PFC_00980	dipeptide-binding protein	-1.5	9.3
PF0363	PFC_00950	beta-galactosidase	-1.4	9.2
PF0366	PFC_00935	hypothetical protein	-1.1	7.0
PF0372	PFC_00905	molybdenum cofactor biosynthesis protein	-1.3	8.9
PF0379	PFC_00870	50S ribosomal protein L39e	1.7	12.1
PF0380	PFC_00865	hypothetical protein	-1.4	8.9
PF0386	PFC_00835	hypothetical protein	1.2	6.6
PF0399	PFC_01080	s1p family ribosomal protein	-1.8	9.2
PF0401	PFC_01090	methyltransferase	2.1	11.7
PF0402	PFC_01100	UDP-or dTTP-glucose 4-epimerase or 4-6-dehydratase	-1.3	8.3
PF0411	PFC_01140	protein-export membrane protein	-1.2	5.6
PF0413	PFC_01150	hypothetical protein	1.1	7.2
PF0426	PFC_01215	phosphoribosylaminoimidazole carboxylase ATPase subunit	1.3	7.8
PF0428	PFC_01225	hypothetical protein	-1.3	7.4
PF0429.1n	PFC_01245	transposase	1.1	11.4
PF0430	PFC_01250	phosphoribosylglycinamide formyltransferase 2	-1.0	8.7
PF0432.3n	PFC_01275	hypothetical protein	-1.2	6.0
PF0441	PFC_01320	galactose-1-phosphate uridylyltransferase	-1.9	6.5
PF0445	PFC_01340	galactokinase	-1.8	8.8
PF0450	PFC_01365	glutamine synthetase	-1.7	11.0
PF0460	PFC_01415	hypothetical protein	1.4	5.7
PF0470	PFC_01465	hypothetical protein	-1.3	7.7
PF0471	PFC_01470	hypothetical protein	-1.4	6.7
PF0473	PFC_01480	serine/threonine protein kinase	-1.5	11.6
PF0474	PFC_01485	DNA mismatch repair protein	-1.8	7.9

Table S2 Continued				
Pfu ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change MW76 (2.5h)-(40h)**	**-log 10(p value)**
PF0493	PFC_01580	hypothetical protein	-1.4	11.1
PF0501	PFC_01625	NTPase	-1.6	12.2
PF0508	PFC_01660	integral membrane glycosyltransferase	-1.4	10.5
PF0510	PFC_01670	nucleotide pyrophosphohydrolase	1.2	10.0
PF0512	PFC_01680	transcription factor	1.1	11.5
PF0513	PFC_01685	4-aminobutyrate aminotransferase	1.2	10.3
PF0519	PFC_01715	hypothetical protein	-1.3	7.3
PF0527	PFC_01755	hypothetical protein	-1.2	9.4
PF0529	PFC_01765	cobalt ABC transporter permease CbiQ	-1.2	8.7
PF0532	PFC_01780	hypothetical protein	1.2	8.1
PF0543	PFC_01840	molybdopterin converting factor subunit 1	1.3	16.2
PF0544	PFC_01845	hypothetical protein	-1.4	5.9
PF0546	PFC_01855	hypothetical protein	-1.2	8.3
PF0553	PFC_01890	arsenate reductase	-1.0	6.1
PF0554	PFC_01900	hypothetical protein	1.0	10.6
PF0559	PFC_01925	hydrogenase maturation protein HypF	1.1	7.5
PF0560	PFC_01930	hypothetical protein	-1.8	9.6
PF0564	PFC_01950	ATP-dependent RNA helicase hepa	1.3	12.1
PF0565	PFC_01955	hypothetical protein	-1.6	7.6
PF0566	PFC_01960	hypothetical protein	-1.9	13.7
PF0591	PFC_02085	mannosyl-3-phosphoglycerate synthase	1.2	5.6
PF0592	PFC_02090	ATP-dependent RNA helicase	-1.0	6.1
PF0598	PFC_02120	aspartate carbamoyltransferase regulatory subunit	1.1	6.5
PF0615	PFC_02215	hydrogenase nickel incorporation protein	1.0	10.0

Table S2 Continued				
Pfu ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change MW76 (2.5h)-(40h)**	**-log 10(p value)**
PF0617	PFC_02225	hydrogenase maturation protease HycI	1.0	6.9
PF0619	PFC_02235	hypothetical protein	1.1	6.1
PF0632	PFC_02295	hypothetical protein	-1.3	6.8
PF0634	PFC_02305	DEXX-box atpase	-1.2	9.4
PF0638	PFC_02325	hypothetical protein	-1.0	6.0
PF0641	PFC_02340	hypothetical protein	1.3	11.1
PF0642	PFC_02345	hypothetical protein	-1.7	8.5
PF0655.1n	PFC_02420	hypothetical protein	-1.4	8.4
PF0664	PFC_02475	hypothetical protein	-2.3	10.8
PF0668	PFC_02495	hypothetical protein	-2.0	10.6
PF0674	PFC_02530	ion gradient-generating decarboxylase subunit beta	-1.3	7.1
PF0677	PFC_02545	ski2-like helicase	-1.4	5.8
PF0691	PFC_02610	hypothetical protein	1.1	8.0
PF0693	PFC_02620	hypothetical protein	-1.6	5.7
PF0716	PFC_02725	3-hydroxyisobutyrate dehydrogenase	-1.6	7.9
PF0717	PFC_02730	methyltransferase	1.2	9.5
PF0721	PFC_02750	hypothetical protein	1.3	6.2
PF0728	PFC_02785	hypothetical protein	-1.1	7.7
PF0732	PFC_02805	daunorubicin resistance ATP-binding protein	-2.0	7.6
PF0743	PFC_02875	hypothetical protein	-1.3	10.9
PF0755	PFC_02935	non-phosphorylating glyceraldehyde-3-phosphate dehydrogenase	-1.5	9.5
PF0763	PFC_02965	hypothetical protein	1.1	8.8
PF0770	PFC_03005	glucose-1-phosphate uridylyltransferase	-2.3	13.8
PF0799	PFC_03165	hypothetical protein	-2.1	7.9
PF0802	PFC_03180	hypothetical protein	-1.7	11.0

Table S2 Continued				
Pfu ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change MW76 (2.5h)-(40h)**	**-log 10(p value)**
PF0808	PFC_03210	hypothetical protein	-1.4	7.4
PF0810	PFC_03220	hypothetical protein	-1.3	8.9
PF0812	PFC_03235	hypothetical protein	-1.8	6.9
PF0815	PFC_03250	hypothetical protein	-1.0	5.6
PF0818	PFC_03265	aminohydrolase	-1.3	7.9
PF0819	PFC_03270	50S ribosomal protein L14e	1.6	11.1
PF0825	PFC_03310	prolyl endopeptidase	1.2	5.7
PF0829	PFC_03330	hypothetical protein	-1.3	8.4
PF0830	PFC_03335	hypothetical protein	1.1	6.7
PF0831	PFC_03340	hypothetical protein	-1.0	7.3
PF0844	PFC_03415	hypothetical protein	1.1	9.1
PF0847	PFC_03430	hypothetical protein	-1.4	7.9
PF0851	PFC_03450	iron-dependent repressor	-1.5	8.3
PF0853	PFC_03460	5'-methylthioadenosine phosphorylase	-1.3	7.1
PF0855.1n	PFC_03480	hypothetical protein	-2.2	19.1
PF0860	PFC_03500	btpA family protein	1.1	5.5
PF0862	PFC_03520	tyrosyl-tRNA ligase	-1.9	8.5
PF0876	PFC_03585	50S ribosomal protein L15e	1.6	7.4
PF0880	PFC_03600	DEXX-box atpase	-1.6	11.8
PF0881	PFC_03605	ABC transporter	-1.2	7.1
PF0886	PFC_03640	hypothetical protein	-1.3	6.5
PF0889	PFC_03655	hypothetical protein	-1.6	9.8
PF0892	PFC_03670	cytochrome-c3 hydrogenase subunit gamma	-1.1	7.8
PF0902	PFC_03730	hypothetical protein	-1.1	7.6
PF0907	PFC_03755	hypothetical protein	-1.3	6.8
PF0910	PFC_03770	iron ABC transporter permease	-3.3	22.3
PF0915	PFC_03795	hypothetical protein	-1.3	8.6
PF0920	PFC_03820	nucleotidyltransferase	-1.2	7.5
PF0924	PFC_03840	hypothetical protein	-1.7	10.3

Table S2 Continued				
Pfu ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change MW76 (2.5h)-(40h)**	**-log 10(p value)**
PF0929	PFC_03865	hypothetical protein	1.2	8.8
PF0930	PFC_03870	hypothetical protein	1.2	8.2
PF0961	PFC_04035	hypothetical protein	-1.1	8.4
PF0965	PFC_04055	pyruvate ferredoxin oxidoreductase subunit beta	-1.6	10.4
PF0966	PFC_04060	pyruvate ferredoxin oxidoreductase subunit alpha	1.5	12.0
PF0972	PFC_04090	hypothetical protein	1.3	7.0
PF0973	PFC_04095	acetyl-CoA acetyltransferase	1.4	6.2
PF0975	PFC_04105	hypothetical protein	-1.2	5.6
PF0979	PFC_04125	cobalamin adenosyltransferase	1.3	7.0
PF0992	PFC_04190	hypothetical protein	-1.3	8.8
PF0994	PFC_04200	type II secretion system protein	1.1	7.0
PF0997	PFC_04215	hypothetical protein	1.2	8.0
PF1004	PFC_04250	alkaline phosphatase	-1.1	6.0
PF1016	PFC_04315	2-hydroxyhepta-2,4-diene-1,7-dioate isomerase	-1.2	6.2
PF1018	PFC_04325	50S ribosomal protein L41	-1.1	9.9
PF1032	PFC_04395	hypothetical protein	1.3	11.4
PF1045	PFC_04460	adenylate kinase	1.6	11.9
PF1048	PFC_04475	hypothetical protein	1.1	10.5
PF1049	PFC_04480	Pyruvate formate lyase activating like protein	-1.5	8.2
PF1059	PFC_04530	hypothetical protein	-1.4	9.0
PF1073	PFC_04590	hypothetical protein	-1.9	8.4
PF1074	PFC_04595	hypothetical protein	-1.0	5.8
PF1077	PFC_04610	hypothetical protein	-2.3	13.9
PF1080	PFC_04625	hypothetical protein	-1.2	6.4
PF1084	PFC_04640	phosphopantetheine adenylyltransferase	-1.4	8.6
PF1085	PFC_04645	hypothetical protein	-1.4	5.9

Table S2 Continued				
Pfu ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change MW76 (2.5h)-(40h)**	**-log 10(p value)**
PF1086	PFC_04655	hypothetical protein	1.0	10.4
PF1095	PFC_04700	hypothetical protein	-1.6	5.8
PF1096	PFC_04705	isoleucyl-tRNA ligase	-1.1	10.0
PF1100	PFC_04725	hypothetical protein	-1.6	8.3
PF1101	PFC_04730	hypothetical protein	-1.5	5.8
PF1105	PFC_04750	pyrazinamidase/nicotinamidase pxnc	-1.6	13.1
PF1107	PFC_04760	hit family protein	-1.2	6.5
PF1112	PFC_04780	hypothetical protein	-1.5	6.7
PF1128	PFC_04860	hypothetical protein	1.2	10.3
PF1130	PFC_04870	hypothetical protein	1.0	7.7
PF1136	PFC_04900	hypothetical protein	-1.2	8.3
PF1140	PFC_04925	translation initiation factor IF-2 subunit alpha	1.2	8.9
PF1141	PFC_04930	H/ACA RNA-protein complex component Nop10p	1.9	7.2
PF1143	PFC_04940	3-methyl-2-oxobutanoate hydroxymethyltransferase	-1.7	8.0
PF1153	PFC_04990	monovalent cation/H+ antiporter subunit E	-1.4	6.3
PF1159	PFC_05020	L-tyrosine decarboxylase	-1.5	6.8
PF1170	PFC_05070	hypothetical protein	1.1	9.0
PF1180	PFC_05120	threonine synthase	-2.2	11.0
PF1184	PFC_05140	hypothetical protein	-1.4	6.9
PF1188	PFC_05160	pyruvate kinase	1.1	6.6
PF1191	PFC_05175	hypothetical protein	1.6	7.1
PF1193	PFC_05180	DNA protection protein DPS	-1.7	13.4
PF1204	PFC_05230	seryl-tRNA ligase	-1.8	9.0
PF1220	PFC_05320	sulfate adenylyltransferase	-1.1	7.7
PF1239	PFC_05415	hypothetical protein	-1.1	6.9
PF1243	PFC_05445	hypothetical protein	-2.6	7.0
PF1247.1n	PFC_05470	transposase	1.2	5.9
PF1258	PFC_05495	ribose-5-phosphate isomerase A	1.3	5.9

Table S2 Continued				
Pfu ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change MW76 (2.5h)-(40h)**	**-log 10(p value)**
PF1271	PFC_05570	hypothetical protein	-1.3	6.9
PF1281	PFC_05620	superoxide reductase	1.5	16.0
PF1287	PFC_05650	ABC transporter	-1.3	11.5
PF1295	PFC_05690	transcription factor	1.1	8.3
PF1303	PFC_05730	hypothetical protein	1.0	7.2
PF1310	PFC_05760	hypothetical protein	-1.7	11.1
PF1313	PFC_05775	DNA processing SMF protein	-1.4	5.5
PF1314	PFC_05780	hypothetical protein	-1.5	9.0
PF1323	PFC_05830	hypothetical protein	1.3	6.5
PF1332	PFC_05875	hydrogenase subunit alpha	-1.1	6.9
PF1336	PFC_05895	cytosine permease	-2.3	15.4
PF1337	PFC_05900	transcriptional activator	1.1	5.9
PF1341	PFC_05920	glycine cleavage system aminomethyltransferase T	-2.1	7.2
PF1343	PFC_05930	X-Pro dipeptidase	-1.6	7.0
PF1348	PFC_05955	hypothetical protein	-2.0	12.3
PF1351	PFC_05970	threonyl-tRNA ligase	-1.7	14.2
PF1353	PFC_05980	hypothetical protein	1.3	8.5
PF1367	PFC_06055	50S ribosomal protein L7Ae	1.6	15.0
PF1368	PFC_06060	30S ribosomal protein S28e	1.7	17.3
PF1369	PFC_06065	50S ribosomal protein L24e	-1.2	9.0
PF1371	PFC_06075	hypothetical protein	-1.7	9.2
PF1374	PFC_06095	lipoate-protein ligase A	-1.9	10.8
PF1375	PFC_06100	elongation factor 1-alpha	1.5	8.2
PF1388	PFC_06165	hypothetical protein	-1.0	7.9
PF1405	PFC_06250	cleavage and polyadenylation specifity factor protein	-2.2	12.8
PF1413	PFC_06290	hypothetical protein	-1.5	7.7
PF1444	PFC_06445	NADH dehydrogenase subunit B	1.4	5.8
PF1460	PFC_06535	hypothetical protein	-1.2	6.3
PF1468	PFC_06575	hypothetical protein	-3.4	24.6
PF1483	PFC_06650	hypothetical protein	-1.4	7.2

Table S2 Continued				
Pfu ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change MW76 (2.5h)-(40h)**	**-log 10(p value)**
PF1488	PFC_06675	hypothetical protein	-1.7	8.4
PF1496	PFC_06715	hypothetical protein	-1.5	11.1
PF1499	PFC_06730	30S ribosomal protein S19e	1.5	8.4
PF1513	PFC_06800	hypothetical protein	-1.2	6.7
PF1527	PFC_06865	hypothetical protein	-1.8	15.1
PF1529	PFC_06875	pyridoxal biosynthesis lyase PdxS	-4.4	26.9
PF1535	PFC_06905	alpha-glucan phosphorylase	1.4	6.1
PF1536	PFC_06910	RNase III inhibitor	1.6	5.5
PF1541	PFC_06935	50S ribosomal protein L37e	-1.5	13.9
PF1548	PFC_06970	hypothetical protein	1.2	8.1
PF1549	PFC_06975	RNA 3'-terminal-phosphate cyclase	1.4	13.5
PF1558	PFC_07020	30S ribosomal protein S7P	1.2	7.6
PF1560	PFC_07030	transcription elongation factor NusA-like protein	1.8	6.3
PF1562	PFC_07040	DNA-directed RNA polymerase subunit A"	1.3	7.5
PF1565	PFC_07055	DNA-directed RNA polymerase subunit H	1.6	11.7
PF1566	PFC_07060	hypothetical protein	-1.3	8.8
PF1568	PFC_07070	exosome complex exonuclease Rrp41	1.5	10.5
PF1569	PFC_07075	exosome complex RNA-binding protein Rrp4	1.8	8.4
PF1571	PFC_07085	proteasome subunit alpha	-1.0	7.1
PF1572	PFC_07090	hypothetical protein	1.1	8.1
PF1585	PFC_07155	hypothetical protein	-1.7	8.7
PF1589	PFC_07175	hypothetical protein	-1.0	5.9
PF1600	PFC_07230	ferripyochelin binding protein	-1.7	10.6
PF1606	PFC_10585	dolichol-phosphate mannose synthase	1.4	6.7
PF1615	PFC_10530	hypothetical protein	1.4	10.8

Table S2 Continued				
Pfu ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change MW76 (2.5h)-(40h)**	**-log 10(p value)**
PF1619	PFC_10510	hypothetical protein	-1.2	8.3
PF1622	PFC_10495	n-type ATP pyrophosphatase superfamily protein	-1.8	9.5
PF1631	PFC_10450	UDP-N-acetyl-D-mannosaminuronate dehydrogenase	1.1	7.0
PF1632	PFC_10445	hypothetical protein	-1.5	8.4
PF1637	PFC_10420	mevalonate kinase	1.2	5.6
PF1642	PFC_10395	DNA-directed RNA polymerase subunit K	1.7	7.0
PF1646	PFC_10375	50S ribosomal protein L18e	1.6	9.4
PF1648	PFC_10365	30S ribosomal protein S11P	1.8	11.8
PF1649	PFC_10360	30S ribosomal protein S4P	1.4	10.8
PF1654	PFC_10335	ABC transporter	-2.3	8.1
PF1656	PFC_10325	NADH oxidase	-1.7	11.4
PF1657	PFC_10320	histidyl-tRNA synthetase	1.4	11.0
PF1662	PFC_10295	1-(5-phosphoribosyl)-5-[(5-phosphoribosylamino)methylideneamino] imidazole-4-carboxamide isomerase	-1.4	5.7
PF1679	PFC_10210	3-isopropylmalate dehydratase large subunit	1.5	9.9
PF1680	PFC_10205	3-isopropylmalate dehydratase small subunit	1.1	8.2
PF1681	PFC_10200	lysine biosynthesis protein	1.7	6.5
PF1686	PFC_10175	acetyl-lysine deacetylase	-1.3	5.8
PF1699	PFC_10110	3-phosphoshikimate 1-carboxyvinyltransferase	1.1	8.0
PF1701	PFC_10100	chorismate mutase	-2.3	13.8
PF1709	PFC_10060	anthranilate synthase component I	-1.4	7.1
PF1710	PFC_10055	anthranilate phosphoribosyltransferase	-3.4	13.1

Table S2 Continued				
Pfu ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change MW76 (2.5h)-(40h)**	**-log 10(p value)**
PF1711	PFC_10050	indole-3-glycerol-phosphate synthase	-3.3	10.1
PF1722	PFC_09995	histone a2	-1.2	11.3
PF1730	PFC_09955	thymidylate kinase	1.5	6.0
PF1731	PFC_09950	signal recognition particle protein Srp54	1.4	7.1
PF1735.2n	PFC_09925	hypothetical protein	1.3	9.5
PF1737	PFC_09915	hypothetical protein	-1.8	10.1
PF1742	PFC_09885	group 1 glycosyl transferase	1.9	8.9
PF1743	PFC_09880	hypothetical protein	1.2	6.3
PF1750	PFC_09850	sugar-binding transport ATP-binding protein	-1.2	9.0
PF1755	PFC_09825	fumarate hydratase	1.1	11.3
PF1756	PFC_09815	hypothetical protein	-1.6	9.6
PF1764	PFC_09770	metal-dependent hydrolase	-2.1	11.8
PF1778	PFC_09700	serine hydroxymethyltransferase	-2.2	7.0
PF1780	PFC_09690	ABC transporter	-1.6	8.1
PF1783	PFC_09675	molybdenum cofactor biosynthesis protein MoeA	1.1	5.9
PF1795	PFC_09615	sarcosine oxidase subunit alpha	-1.1	6.0
PF1802	PFC_09580	50S ribosomal protein L15P	1.5	6.1
PF1803	PFC_09575	50S ribosomal protein L30P	1.9	12.2
PF1804	PFC_09570	30S ribosomal protein S5P	1.7	10.1
PF1805	PFC_09565	50S ribosomal protein L18P	1.4	6.1
PF1807	PFC_09555	50S ribosomal protein L32e	1.9	11.6
PF1808	PFC_09550	50S ribosomal protein L6P	1.6	8.2
PF1811	PFC_09535	50S ribosomal protein L5P	2.0	13.8
PF1813	PFC_09525	50S ribosomal protein L24P	1.8	13.5
PF1815	PFC_09515	30S ribosomal protein S17P	1.3	12.5
PF1817	PFC_09505	translation initiation factor Sui1	1.6	5.9
PF1819	PFC_09495	30S ribosomal protein S3P	1.8	9.6
PF1822	PFC_09480	50S ribosomal protein L2P	1.8	13.2
PF1823	PFC_09475	50S ribosomal protein L23P	2.0	10.5

Table S2 Continued				
Pfu ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change MW76 (2.5h)-(40h)**	**-log 10(p value)**
PF1824	PFC_09470	50S ribosomal protein L4P	1.4	8.8
PF1825	PFC_09465	50S ribosomal protein L3P	1.1	6.5
PF1826	PFC_09460	hypothetical protein	1.5	7.0
PF1829	PFC_09445	nucleotidyltransferase	-1.8	6.3
PF1831	PFC_09435	histone a1	2.0	23.2
PF1838	PFC_09400	hypothetical protein	-2.2	12.2
PF1842	PFC_09380	hypothetical protein	-1.4	6.8
PF1844	PFC_09370	hypothetical protein	-1.2	6.1
PF1848	PFC_09350	3-hydroxy-3-methylglutaryl-CoA reductase	1.5	7.2
PF1857	PFC_09305	cobalt transport ATP-binding protein	-1.0	5.7
PF1867	PFC_09250	hypothetical protein	-1.5	7.2
PF1868	PFC_09245	integrase-recombinase protein	1.0	5.7
PF1883	PFC_09170	small heat shock protein	-1.8	10.4
PF1885	PFC_09160	glycosyl transferase family protein	-1.2	6.5
PF1886	PFC_09155	sugar kinase	1.6	8.1
PF1897	PFC_09100	hypothetical protein	-2.1	10.3
PF1905	PFC_09060	protease	-1.1	6.6
PF1907	PFC_09050	hypothetical protein	1.6	8.2
PF1908	PFC_09045	hypothetical protein	-1.7	17.2
PF1912	PFC_09025	2-methylthioadenine synthetase	-1.4	12.7
PF1914	PFC_09015	ribonuclease P protein component 3	-1.5	6.4
PF1918	PFC_08995	hypothetical protein	-1.7	9.5
PF1919	PFC_08990	dehydrogenase	-2.3	8.0
PF1932	PFC_08920	hypothetical protein	-2.5	12.1
PF1935	PFC_08910	pullulanase	2.1	11.3
PF1936	PFC_08905	malg-like sugar transport inner membrane protein	1.6	6.2
PF1937	PFC_08900	malf-like sugar transport inner membrane protein	1.9	9.2
PF1938	PFC_08895	malE-like sugar binding protein	2.0	6.9

Table S2 Continued				
Pfu ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change MW76 (2.5h)-(40h)**	**-log 10(p value)**
PF1940	PFC_08885	hypothetical protein	-2.0	11.7
PF1947	PFC_08845	hypothetical protein	1.3	7.8
PF1949	PFC_08835	hypothetical protein	-1.8	10.7
PF1954	PFC_08810	molybdopterin-guanine dinucleotide biosynthesis protein MobB	-1.9	9.4
PF1962	PFC_08770	hypothetical protein	-1.6	10.5
PF1966	PFC_08750	hypothetical protein	-1.6	6.5
PF1974	PFC_08710	thermosome, single subunit	-2.6	21.4
PF1981	PFC_08675	hypothetical protein	-1.1	7.2
PF1983	PFC_08665	hypothetical protein	-1.4	6.6
PF1993	PFC_08600	acidic ribosomal protein P0	1.4	11.2
PF1994	PFC_08595	50S ribosomal protein L12P	1.8	17.6
PF1999	PFC_08570	glycine dehydrogenase subunit 1	-1.3	6.0
PF2000	PFC_08565	glycine dehydrogenase subunit 2	-2.1	10.4
PF2002	PFC_08550	sucrose transport protein	1.0	14.1
PF2005	PFC_08535	glycerol-3-phosphate dehydrogenase	1.1	8.7
PF2008	PFC_08520	50S ribosomal protein L37Ae	1.3	6.6
PF2013	PFC_08495	hypothetical protein	1.2	10.4
PF2014	PFC_08490	bifunctional inositol-1 monophosphatase/fructose-1,6-bisphosphatase	1.5	8.0
PF2037	PFC_08385	hypothetical protein	-1.1	8.8
PF2045	PFC_08340	hypothetical protein	-1.9	12.4
PF2046	PFC_08335	hypothetical protein	-2.0	14.1
PF2048	PFC_08320	peptidase	-1.1	6.7
	PFC_09510	ribonuclease P protein component 1	1.7	7.3
	PFC_05475	sodium dependent transporter	1.1	7.0
	PFC_01065	hypothetical protein	1.1	5.6
	PFC_04695	hypothetical protein	1.0	10.0
	PFC_03230	hypothetical protein	-1.1	7.4

Table S2 Continued				
Pfu ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change MW76 (2.5h)-(40h)**	**-log 10(p value)**
	PFC_08380	hypothetical protein	-1.2	10.7
	PFC_02365	hypothetical protein	-1.3	9.9
	PFC_03055	hypothetical protein	-1.4	7.9
	PFC_00055	hypothetical protein	-1.4	6.2
	PFC_07605	hypothetical protein	-1.4	5.8
	PFC_04020	hypothetical protein	-1.5	6.3
	PFC_03570	hypothetical protein	-1.6	6.8
	PFC_00815	hypothetical protein	-1.6	6.1

Table S3: Pfu COM1 vs MW76 (72°C, 2.5h) transcriptome				
Pfu ORF# OR Msed ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change (COM1)-(MW76) 2.5h**	**-log 10(p value)**
PF0004	PFC_08190	hypothetical protein	-2.6	13.3
PF0005	PFC_08185	ABC transporter ATP-binding protein	-1.9	8.1
PF0044	PFC_07985	hypothetical protein	-1.0	7.4
PF0221	PFC_00165	hypothetical protein	1.0	6.0
PF0224	PFC_00185	hypothetical protein	-1.8	15.4
PF0532	PFC_01780	hypothetical protein	-1.2	8.6
PF1033	PFC_04400	peroxiredoxin	1.2	7.4
PF1209	PFC_05255	oligopeptide ABC transporter	2.0	14.0
PF1281	PFC_05620	superoxide reductase	1.0	11.4
PF1287	PFC_05650	ABC transporter	1.3	11.7
PF1678	PFC_10215	2-isopropylmalate synthase	3.5	13.9
PF1679	PFC_10210	3-isopropylmalate dehydratase large subunit	3.7	20.1
PF1680	PFC_10205	3-isopropylmalate dehydratase small subunit	3.2	21.0
PF1681	PFC_10200	lysine biosynthesis protein	3.0	12.3
PF1682	PFC_10195	ribosomal protein s6 modification protein	1.5	11.1
PF1683	PFC_10190	N-acetyl-gamma-glutamyl-phosphate reductase	2.7	10.5
PF1684	PFC_10185	acetylglutamate/acetylaminoadipate kinase	3.1	17.9
PF1685	PFC_10180	acetylornithine/acetyl-lysine aminotransferase	2.9	8.3
PF1686	PFC_10175	acetyl-lysine deacetylase	2.6	12.2
PF1974	PFC_08710	thermosome, single subunit	1.3	12.7
PF2060.1n	PFC_08235	arylsulfatase	-3.0	11.9
PF2061	PFC_08230	ABC transporter	-2.1	10.6
PF2064	PFC_08215	arylsulfatase	-1.0	6.1
	PFC_02450	hypothetical protein	-1.2	6.5

Table S3 Continued				
Pfu ORF# OR Msed ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change (COM1)-(MW76) 2.5h**	**-log 10(p value)**
Msed_1993	Msed_1993	3-hydroxypropionate dehydrogenase (NADP(+))	-3.0	12.4
Msed_0390	Msed_0390	carbonic anhydrase	-4.1	15.1
Msed_2010	Msed_2010	Biotin protein ligase	-4.3	24.4
Msed_0148	Msed_0148	biotin carboxyl carrier protein	-4.4	25.8
Msed_0709	Msed_0709	succinyl-CoA reductase (NADPH)	-4.7	22.2
Msed_0147	Msed_0147	biotin carboxylase	-4.8	24.5
Msed_1375	Msed_1375	acetyl-CoA carboxylase carboxyltransferase subunit alpha / acetyl-CoA carboxylase carboxyltransferase subunit beta / propionyl-CoA carboxylase carboxyltransferase subunit	-5.1	23.8

Table S4: Pfu COM1 vs MW76 (72°C, 40h) transcriptome				
Pfu ORF# OR Msed ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change (COM1)-(MW76) 40h**	**-log 10(p value)**
PF0014	PFC_08145	hypothetical protein	-1.6	8.8
PF0032	PFC_08045	permease	1.3	6.4
PF0034	PFC_08035	hypothetical protein	1.4	8.5
PF0038	PFC_08015	glyoxalase II	-1.0	9.0
PF0043	PFC_07990	phosphoenolpyruvate synthase	-4.1	30.1
PF0055	PFC_07925	HTH transcription regulator	2.7	13.8
PF0074	PFC_07830	alcohol dehydrogenase	-1.2	5.9
PF0078	PFC_07810	2-phosphoglycerate kinase	3.3	13.2
PF0081	PFC_07795	anion transport system permease	1.1	6.2
PF0090	PFC_07755	molybdenum cofactor biosynthesis protein A	1.2	9.4
PF0095	PFC_07730	hypothetical protein	-1.5	6.5
PF0105	PFC_07680	hypothetical protein	-1.2	7.2
PF0108	PFC_07665	L-fuculose phosphate aldolase	1.0	9.6
PF0114	PFC_07635	GTP-binding protein	1.2	8.6
PF0116	PFC_07625	multiple sugar transport protein	1.2	7.9
PF0124	PFC_07580	hypothetical protein	-1.8	7.7
PF0130	PFC_07550	hypothetical protein	-1.4	9.0
PF0132	PFC_07540	phosphate transport system substrate-binding protein	-1.8	7.9
PF0133	PFC_07535	hypothetical protein	-1.3	6.2
PF0144	PFC_07480	hypothetical protein	-2.0	13.0
PF0146	PFC_07470	potassium channel	1.1	8.0
PF0171	PFC_07345	hypothetical protein	1.3	9.4
PF0177	PFC_07315	V-type ATP synthase subunit I	-2.2	12.2
PF0178	PFC_07310	V-type ATP synthase subunit K	-3.2	12.5
PF0179	PFC_07305	V-type ATP synthase subunit E	-1.8	7.6
PF0180	PFC_07300	V-type ATP synthase subunit C	-1.8	12.2
PF0181	PFC_07295	V-type ATP synthase subunit F	-1.4	10.0
PF0183	PFC_07285	V-type ATP synthase subunit B	-2.2	8.4
PF0196	PFC_00035	glucose-6-phosphate isomerase	-1.4	8.7
PF0210	PFC_00110	oxidative cyclase	-2.9	15.7
PF0217	PFC_00145	50S ribosomal protein L44e	3.4	18.3
PF0218	PFC_00150	30S ribosomal protein S27e	1.4	7.6

Table S4 Continued				
Pfu ORF# OR Msed ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change (COM1)-(MW76) 40h**	**-log 10(p value)**
PF0220	PFC_00160	D-arabino 3-hexulose 6-phosphate formaldehyde lyase	-1.6	9.1
PF0231	PFC_00205	hypothetical protein	-1.2	7.4
PF0239	PFC_00245	hypothetical protein	-1.1	6.2
PF0242	PFC_00260	transposase	1.1	6.3
PF0250	PFC_00300	transcriptional regulator	1.6	19.3
PF0253	PFC_00320	30S ribosomal protein S24e	-1.2	7.5
PF0254	PFC_00325	hypothetical protein	-1.2	8.0
PF0256	PFC_00335	DNA-directed RNA polymerase subunit E'	-1.3	6.9
PF0264	PFC_00375	histidyl-tRNA ligase	2.6	11.8
PF0265	PFC_00380	2-amino-3-ketobutyrate coenzyme A ligase	-2.8	16.8
PF0296	PFC_00545	cobalamin biosynthesis protein	1.1	5.7
PF0311	PFC_00635	hypothetical protein	-1.1	5.6
PF0312	PFC_00640	ADP-dependent glucokinase	-1.5	6.8
PF0318	PFC_00665	acylaminoacyl peptidase	-1.9	15.9
PF0324	PFC_00700	hypothetical protein	-4.6	26.6
PF0325	PFC_00705	hypothetical protein	-2.8	13.4
PF0333	PFC_00745	flagella-related protein g	1.2	8.8
PF0335	PFC_00755	flagella-related protein D	1.0	6.8
PF0336	PFC_00760	hypothetical protein	1.2	12.6
PF0365	PFC_00940	hypothetical protein	2.0	9.5
PF0369	PFC_00920	endoglucanase	-1.5	8.2
PF0372	PFC_00905	molybdenum cofactor biosynthesis protein	-1.0	6.8
PF0374	PFC_00895	translation-associated GTPase	1.2	5.7
PF0375	PFC_00890	prefoldin subunit alpha	-2.1	11.8
PF0376	PFC_00885	50S ribosomal protein LX	-1.4	11.8
PF0378	PFC_00875	50S ribosomal protein L31e	1.1	7.7
PF0379	PFC_00870	50S ribosomal protein L39e	1.8	12.8

Table S4 Continued				
Pfu ORF# OR Msed ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change (COM1)-(MW76) 40h**	**-log 10(p value)**
PF0380	PFC_00865	hypothetical protein	-1.2	7.8
PF0383	PFC_00850	hypothetical protein	-1.7	10.3
PF0385	PFC_00840	hypothetical protein	-1.3	9.8
PF0393	PFC_01050	hypothetical protein	1.7	10.3
PF0399	PFC_01080	s1p family ribosomal protein	-1.9	9.5
PF0401	PFC_01090	methyltransferase	-1.6	8.6
PF0404	PFC_01110	hypothetical protein	1.0	10.0
PF0415	PFC_01160	dehydrogenase	1.2	6.0
PF0421	PFC_01190	5-formaminoimidazole-4-carboxamide-1-(beta)-D-ribofuranosyl 5'-monophosphate synthetase-like protein	1.5	9.0
PF0438	PFC_01300	hypothetical protein	1.4	8.6
PF0450	PFC_01365	glutamine synthetase	-1.9	12.3
PF0478	PFC_01505	alpha-amylase	-1.2	6.3
PF0483	PFC_01530	hypothetical protein	1.8	8.2
PF0487	PFC_01550	autolysin-like protein	1.4	6.4
PF0500	PFC_01620	met-10+ protein	-1.3	7.5
PF0505	PFC_01645	hypothetical protein	1.1	7.7
PF0512	PFC_01680	transcription factor	1.4	14.2
PF0514	PFC_01690	d-alanine glycine permease	1.2	9.5
PF0519	PFC_01715	hypothetical protein	-1.6	9.0
PF0536	PFC_01800	transposase	1.7	8.5
PF0546	PFC_01855	hypothetical protein	-2.0	14.0
PF0547	PFC_01860	hypothetical protein	-1.4	6.8
PF0548	PFC_01865	hydrogenase expression/formation protein	-1.1	7.8
PF0563	PFC_01945	hypothetical protein	1.2	6.6
PF0564	PFC_01950	ATP-dependent RNA helicase hepa	1.2	10.6
PF0567	PFC_01965	hypothetical protein	1.5	14.9

Table S4 Continued				
Pfu ORF# OR Msed ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change (COM1)-(MW76) 40h**	**-log 10(p value)**
PF0581	PFC_02035	hypothetical protein	1.2	12.1
PF0585	PFC_02055	DEXX-box atpase	1.4	8.0
PF0588	PFC_02070	bifunctional phosphomannomutase/phosphoglucomutase	-1.4	6.0
PF0589	PFC_02075	mannose-6-phosphate isomerase	-1.6	6.5
PF0593	PFC_02095	hypothetical protein	1.4	7.1
PF0594	PFC_02100	ornithine carbamoyltransferase	-4.0	15.8
PF0598	PFC_02120	aspartate carbamoyltransferase regulatory subunit	1.2	7.6
PF0606	PFC_02165	hypothetical protein	-2.0	7.5
PF0607	PFC_02170	hypothetical protein	-2.2	16.6
PF0609	PFC_02180	hypothetical protein	-1.1	6.1
PF0631	PFC_02290	radical SAM domain-containing protein	1.0	15.4
PF0637	PFC_02320	hypothetical protein	1.1	8.3
PF0642	PFC_02345	hypothetical protein	-1.4	6.5
PF0668	PFC_02495	hypothetical protein	-1.7	8.9
PF0674	PFC_02530	ion gradient-generating decarboxylase subunit beta	-1.3	7.1
PF0680	PFC_02560	hypothetical protein	1.5	9.2
PF0687	PFC_02590	transcription initiation factor IIB chain b	1.5	10.5
PF0689	PFC_02605	hypothetical protein	1.5	10.5
PF0691	PFC_02610	hypothetical protein	1.0	7.5
PF0692	PFC_02615	hydroxylamine reductase	2.2	11.0
PF0727	PFC_02780	hypothetical protein	-1.1	8.7
PF0730	PFC_02795	hypothetical protein	1.5	15.9
PF0731	PFC_02800	copper-transporting atpase, p-type	1.2	9.5
PF0733	PFC_02810	ABC transporter	1.7	12.1

Table S4 Continued				
Pfu ORF# OR Msed ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change (COM1)-(MW76) 40h**	**-log 10(p value)**
PF0735	PFC_02820	hypothetical protein	2.9	16.7
PF0737	PFC_02830	hypothetical protein	1.0	8.3
PF0753	PFC_02925	2-oxoacid ferredoxin oxidoreductase subunit beta	-1.8	11.2
PF0754	PFC_02930	2-oxoacid:ferredoxin oxidoreductase subunit alpha	-1.6	10.0
PF0755	PFC_02935	non-phosphorylating glyceraldehyde-3-phosphate dehydrogenase	-1.1	6.2
PF0771	PFC_03010	NDP-sugar dehydrogenase	2.3	16.9
PF0781	PFC_03060	hypothetical protein	1.1	6.4
PF0787	PFC_03105	hypothetical protein	1.1	6.9
PF0794	PFC_03140	UDP-n-acetylglucosamine 2-epimerase	-1.8	10.6
PF0795	PFC_03145	capsular polysaccharide biosynthesis protein	-1.5	7.4
PF0819	PFC_03270	50S ribosomal protein L14e	1.5	10.2
PF0821	PFC_03280	50S ribosomal protein L34e	1.4	14.1
PF0823	PFC_03300	hypothetical protein	1.4	7.4
PF0828	PFC_03325	n-type ATP pyrophosphatase superfamily protein	1.2	6.8
PF0829	PFC_03330	hypothetical protein	1.6	10.1
PF0832	PFC_03345	hypothetical protein	1.2	6.0
PF0843	PFC_03410	radical SAM domain-containing protein	2.3	9.6
PF0845	PFC_03420	indolepyruvate ferredoxin oxidoreductase subunit alpha	1.0	11.4
PF0851	PFC_03450	iron-dependent repressor	-1.2	6.4
PF0854	PFC_03465	endo-1,4-beta-glucanase b	1.4	9.7
PF0860	PFC_03500	btpA family protein	2.4	12.9
PF0861	PFC_03505	phosphoglucosamine mutase	2.6	16.6

Table S4 Continued				
Pfu ORF# OR Msed ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change (COM1)-(MW76) 40h**	**-log 10(p value)**
PF0876	PFC_03585	50S ribosomal protein L15e	3.6	16.3
PF0881	PFC_03605	ABC transporter	1.1	6.3
PF0884	PFC_03620	hypothetical protein	1.2	10.0
PF0892	PFC_03670	cytochrome-c3 hydrogenase subunit gamma	-1.4	10.1
PF0894	PFC_03680	sulfhydrogenase subunit alpha	-1.1	8.0
PF0897	PFC_03695	hypothetical protein	1.0	8.1
PF0909	PFC_03765	ferric enterobactin transport ATP-binding protein	1.1	5.8
PF0914	PFC_03790	hypothetical protein	1.3	10.4
PF0918	PFC_03810	tRNA pseudouridine synthase A	1.3	7.9
PF0921	PFC_03825	ABC transporter	1.0	7.3
PF0924	PFC_03840	hypothetical protein	-1.0	6.2
PF0953	PFC_03985	hypothetical protein	-1.6	13.9
PF0958	PFC_04010	hypothetical protein	1.6	11.5
PF0960	PFC_04030	hypothetical protein	1.7	9.9
PF0962	PFC_04040	hypothetical protein	-1.5	6.9
PF0965	PFC_04055	pyruvate ferredoxin oxidoreductase subunit beta	-2.9	17.7
PF0966	PFC_04060	pyruvate ferredoxin oxidoreductase subunit alpha	-2.1	15.5
PF0967	PFC_04065	pyuvate ferredoxin oxidoreductase subunit delta	-3.0	13.4
PF0968	PFC_04070	2-ketoisovalerate ferredoxin oxidoreductase subunit beta	-3.0	11.6
PF0969	PFC_04075	2-ketoisovalerate ferredoxin oxidoreductase subunit alpha	-2.2	9.6
PF0970	PFC_04080	2-ketoisovalerate ferredoxin oxidoreductase subunit delta	-2.6	14.0

Table S4 Continued				
Pfu ORF# OR Msed ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change (COM1)-(MW76) 40h**	**-log 10(p value)**
PF0971	PFC_04085	pyruvate/ketoisovalerate ferredoxin oxidoreductase subunit gamma	-1.5	7.7
PF0998	PFC_04220	GTP-binding protein	1.4	11.4
PF1012	PFC_04290	ABC transporter	1.3	12.1
PF1039	PFC_04430	geranylgeranylglyceryl phosphate synthase-like protein	-1.4	8.7
PF1044	PFC_04455	hypothetical protein	1.4	12.2
PF1062	PFC_04545	hypothetical protein	-2.3	9.1
PF1073	PFC_04590	hypothetical protein	-2.8	12.7
PF1077	PFC_04610	hypothetical protein	-1.3	8.1
PF1081	PFC_04630	hypothetical protein	1.1	6.8
PF1085.1n	PFC_04650	hypothetical protein	-2.2	9.2
PF1089	PFC_04670	hypothetical protein	1.1	8.0
PF1099	PFC_04720	formamidase (formamide amidohydrolase)	1.1	6.2
PF1104	PFC_04745	homoserine dehydrogenase	-1.4	6.8
PF1105	PFC_04750	pyrazinamidase/nicotinamidase pxnc	-1.3	10.9
PF1108	PFC_04765	alpha-dextrin endo-1, 6-alpha-glucosidase	2.2	13.2
PF1122	PFC_04830	hypothetical protein	1.3	9.6
PF1123	PFC_04835	hypothetical protein	-1.4	6.7
PF1140	PFC_04925	translation initiation factor IF-2 subunit alpha	1.8	13.3
PF1141	PFC_04930	H/ACA RNA-protein complex component Nop10p	2.2	8.4
PF1146	PFC_04955	hypothetical protein	1.8	11.1
PF1178	PFC_05110	Na-dependent transporter	1.1	6.8
PF1187	PFC_05155	hypothetical protein	1.1	8.4
PF1201	PFC_05220	hypothetical protein	1.5	10.0
PF1204	PFC_05230	seryl-tRNA ligase	-1.7	8.7
PF1209	PFC_05255	oligopeptide ABC transporter	1.7	12.4

Table S4 Continued				
Pfu ORF# OR Msed ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change (COM1)-(MW76) 40h**	**-log 10(p value)**
PF1213	PFC_05275	oligopeptide ABC transporter	1.3	8.2
PF1214.2n	PFC_05285	hypothetical protein	1.4	6.9
PF1218.1n	PFC_05315	hypothetical protein	1.5	10.1
PF1234	PFC_05390	chitinase	1.1	7.6
PF1255	PFC_05480	hypothetical protein	1.5	11.5
PF1259	PFC_05500	hypothetical protein	-1.5	9.1
PF1269	PFC_05560	methionine synthase	1.0	8.6
PF1270	PFC_05565	hypothetical protein	-1.5	6.5
PF1271	PFC_05570	hypothetical protein	1.1	5.9
PF1279	PFC_05610	50S ribosomal protein L10e	-1.5	9.7
PF1284	PFC_05635	hypothetical protein	-1.2	8.6
PF1285	PFC_05640	hypothetical protein	-1.1	6.4
PF1301	PFC_05720	hypothetical protein	1.1	7.7
PF1302	PFC_05725	hypothetical protein	1.2	7.9
PF1312	PFC_05770	hypothetical protein	1.4	8.3
PF1338	PFC_05905	transcriptional activator	1.7	10.8
PF1347	PFC_05950	hypothetical protein	1.4	7.5
PF1348	PFC_05955	hypothetical protein	-1.1	6.5
PF1349	PFC_05960	translation initiation factor IF-2B subunit beta	-1.4	9.7
PF1353	PFC_05980	hypothetical protein	1.3	8.7
PF1362	PFC_06030	hypothetical protein	1.0	7.0
PF1365	PFC_06045	hypothetical protein	3.5	17.8
PF1374	PFC_06095	lipoate-protein ligase A	-1.5	8.1
PF1388	PFC_06165	hypothetical protein	1.2	9.3
PF1390	PFC_06175	hypothetical protein	1.3	6.8
PF1391	PFC_06180	hypothetical protein	1.4	9.0
PF1393	PFC_06190	hypothetical protein	-1.8	9.4
PF1394	PFC_06195	phosphoglycerate dehydrogenase	-1.9	6.3
PF1399	PFC_06220	ATPase, vanadate-sensitive	-3.0	29.0

Table S4 Continued				
Pfu ORF# OR Msed ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change (COM1)-(MW76) 40h**	**-log 10(p value)**
PF1405	PFC_06250	cleavage and polyadenylation specifity factor protein	-1.1	6.0
PF1408	PFC_06265	dipeptide-binding protein	-2.3	7.9
PF1409	PFC_06270	dipeptide transport system permease	-2.3	5.5
PF1410	PFC_06275	binding-protein-dependent transport systems inner membrane component	-2.0	10.9
PF1411	PFC_06280	dipeptide transport ATP-binding protein	-2.6	14.5
PF1423	PFC_06340	monovalent cation/H+ antiporter subunit E	-1.4	11.1
PF1425	PFC_06350	monovalent cation/H+ antiporter subunit G	-1.6	6.9
PF1427	PFC_06360	hypothetical protein	-2.2	11.8
PF1429	PFC_06370	monovalent cation/H+ antiporter subunit C	-2.3	16.2
PF1431	PFC_06380	hypothetical protein	-1.6	7.9
PF1432	PFC_06385	mbh10 NADH dehydrogenase subunit	-2.1	7.2
PF1434	PFC_06395	membrane bound hydrogenase alpha	-1.9	16.8
PF1436	PFC_06405	NADH-plastoquinone oxidoreductase subunit	-1.7	9.6
PF1449	PFC_06480	monovalent cation/H+ antiporter subunit B	-1.4	8.3
PF1451	PFC_06490	monovalent cation/H+ antiporter subunit G	-1.5	6.9
PF1463	PFC_06555	hypothetical protein	1.0	8.6
PF1467	PFC_06570	ABC transporter	1.1	8.4

Table S4 Continued				
Pfu ORF# OR Msed ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change (COM1)-(MW76) 40h**	**-log 10(p value)**
PF1472	PFC_06595	aspartate/serine transaminase	-1.2	8.7
PF1474	PFC_06605	hypothetical protein	1.1	6.6
PF1479	PFC_06630	oxidoreductase, Fe-S subunit	-1.9	10.6
PF1480	PFC_06635	formaldehyde:ferredoxin oxidoreductase wor5	-1.9	10.1
PF1481	PFC_06640	hypothetical protein	-1.7	8.6
PF1485	PFC_06660	hypothetical protein	1.3	15.9
PF1488	PFC_06675	hypothetical protein	-2.4	11.7
PF1493	PFC_06700	hypothetical protein	1.5	8.1
PF1504	PFC_06755	large helicase-like protein	1.2	7.4
PF1522	PFC_06845	ABC transporter	1.0	5.6
PF1527	PFC_06865	hypothetical protein	-1.2	10.6
PF1529	PFC_06875	pyridoxal biosynthesis lyase PdxS	-3.6	24.3
PF1533	PFC_06895	nodulation protein nfeD	-1.8	6.1
PF1534	PFC_06900	stomatin	-1.6	11.7
PF1540	PFC_06930	ADP forming acetyl coenzyme A synthetase	-2.1	12.6
PF1547	PFC_06965	endoglucanase	-2.0	8.0
PF1550	PFC_06980	hypothetical protein	1.4	6.4
PF1561	PFC_07035	50S ribosomal protein L30e	-1.7	8.7
PF1562	PFC_07040	DNA-directed RNA polymerase subunit A"	-1.2	6.6
PF1563	PFC_07045	DNA-directed RNA polymerase subunit A'	-1.1	7.4
PF1567	PFC_07065	exosome complex RNA-binding protein Rrp42	-1.6	11.1
PF1568	PFC_07070	exosome complex exonuclease Rrp41	-1.2	8.6
PF1569	PFC_07075	exosome complex RNA-binding protein Rrp4	-1.6	7.2
PF1570	PFC_07080	RNA-associated protein	-2.8	15.4

Table S4 Continued				
Pfu ORF# OR Msed ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change (COM1)-(MW76) 40h**	**-log 10(p value)**
PF1571	PFC_07085	proteasome subunit alpha	-2.3	16.0
PF1586	PFC_07160	hydrogenase expression/formation protein	1.0	8.6
PF1601	PFC_07235	AsnC family transcriptional regulator	-1.8	8.8
PF1609	PFC_10565	hypothetical protein	1.3	7.0
PF1610	PFC_10560	hypothetical protein	1.4	9.5
PF1617	PFC_10520	hypothetical protein	1.5	7.0
PF1623	PFC_10490	pyruvoyl-dependent arginine decarboxylase	2.9	14.7
PF1644	PFC_10385	30S ribosomal protein S9P	-1.8	13.9
PF1651	PFC_10350	hypothetical protein	1.6	8.5
PF1654	PFC_10335	ABC transporter	-2.7	9.5
PF1670	PFC_10255	alkaline serine protease	1.8	16.4
PF1671	PFC_10250	hypothetical protein	1.1	6.0
PF1672	PFC_10245	hypothetical protein	1.1	5.6
PF1682	PFC_10195	ribosomal protein s6 modification protein	-1.8	12.6
PF1688	PFC_10165	transketolase	-2.2	15.1
PF1689	PFC_10160	transketolase	-3.0	9.1
PF1690	PFC_10155	3-deoxy-7-phosphoheptulonate synthase	-1.1	9.5
PF1691	PFC_10150	3-dehydroquinate synthase	-1.2	8.2
PF1693	PFC_10140	shikimate 5-dehydrogenase	-1.1	6.4
PF1694	PFC_10135	shikimate kinase	-1.2	8.5
PF1701	PFC_10100	chorismate mutase	-1.4	8.4
PF1702	PFC_10095	aspartate aminotransferase	-1.3	6.6
PF1706	PFC_10075	tryptophan synthase subunit beta	-1.2	7.6
PF1707	PFC_10070	N-(5'-phosphoribosyl)anthranilate isomerase	-1.6	10.1

Table S4 Continued				
Pfu ORF# OR Msed ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change (COM1)-(MW76) 40h**	**-log 10(p value)**
PF1708	PFC_10065	anthranilate synthase component II	-1.3	7.7
PF1710	PFC_10055	anthranilate phosphoribosyltransferase	-2.8	11.0
PF1711	PFC_10050	indole-3-glycerol-phosphate synthase	-3.3	10.0
PF1722	PFC_09995	histone a2	-1.8	15.8
PF1723	PFC_09990	hypothetical protein	1.1	9.2
PF1724	PFC_09985	hypothetical protein	1.1	5.9
PF1729	PFC_09960	phospho-sugar mutase	-1.8	6.3
PF1738	PFC_09910	sugar kinase	1.0	6.4
PF1739	PFC_09900	trehalose/maltose binding protein	-3.3	16.2
PF1745	PFC_09870	hypothetical protein	-1.4	10.0
PF1753	PFC_09835	glutamyl-tRNA ligase	1.3	7.0
PF1756	PFC_09815	hypothetical protein	-1.1	6.2
PF1764	PFC_09770	metal-dependent hydrolase	-2.2	12.5
PF1770	PFC_09740	2-oxoglutarate ferredoxin oxidoreductase subunit gamma	-2.0	11.3
PF1771	PFC_09735	2-oxoglutarate ferredoxin oxidoreductase subunit alpha	-1.2	5.6
PF1772	PFC_09730	2-oxoglutarate ferredoxin oxidoreductase subunit beta	-1.5	8.9
PF1773	PFC_09725	2-oxoglutarate ferredoxin oxidoreductase subunit gamma	-1.8	8.5
PF1792	PFC_09630	5'-cyclic-nucleotide phosphodiesterase	1.1	6.9
PF1802	PFC_09580	50S ribosomal protein L15P	-1.5	6.3
PF1803	PFC_09575	50S ribosomal protein L30P	-1.4	8.8
PF1804	PFC_09570	30S ribosomal protein S5P	-1.2	6.6
PF1806	PFC_09560	50S ribosomal protein L19e	-1.8	7.1
PF1807	PFC_09555	50S ribosomal protein L32e	-1.5	8.9
PF1808	PFC_09550	50S ribosomal protein L6P	-1.6	8.3
PF1809	PFC_09545	30S ribosomal protein S8P	-1.8	8.1

Table S4 Continued				
Pfu ORF# OR Msed ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change (COM1)-(MW76) 40h**	**-log 10(p value)**
PF1810	PFC_09540	30S ribosomal protein S14P	-1.5	6.6
PF1811	PFC_09535	50S ribosomal protein L5P	-1.2	7.8
PF1812	PFC_09530	30S ribosomal protein S4e	-1.6	12.0
PF1814	PFC_09520	50S ribosomal protein L14P	-1.0	9.9
PF1817	PFC_09505	translation initiation factor Sui1	-1.6	6.0
PF1818	PFC_09500	50S ribosomal protein L29P	-2.2	16.5
PF1820	PFC_09490	50S ribosomal protein L22P	-1.8	10.3
PF1821	PFC_09485	30S ribosomal protein S19P	-1.1	10.9
PF1824	PFC_09470	50S ribosomal protein L4P	-1.2	7.3
PF1826	PFC_09460	hypothetical protein	1.8	8.8
PF1832	PFC_09430	GTP cyclohydrolase	1.7	10.2
PF1833	PFC_09425	hypothetical protein	1.1	6.7
PF1858	PFC_09295	cysteine synthase	-1.4	10.0
PF1861	PFC_09280	endo-1,4-beta-glucanase-like protein	-1.1	6.7
PF1866	PFC_09255	S-adenosylmethionine synthetase	-1.1	6.0
PF1881	PFC_09180	DNA/RNA-binding protein albA	-1.2	6.9
PF1892	PFC_09125	hypothetical protein	-1.3	6.3
PF1899	PFC_09090	2-deoxy-d-gluconate 3-dehydrogenase	1.5	7.8
PF1901	PFC_09080	histone acetyltransferase	1.2	9.3
PF1907	PFC_09050	hypothetical protein	-1.2	5.7
PF1909	PFC_09040	ferredoxin	-2.5	13.6
PF1912	PFC_09025	2-methylthioadenine synthetase	-1.4	13.0
PF1921	PFC_08980	phosphoserine phosphatase	1.3	7.3
PF1931	PFC_08925	hypothetical protein	-2.9	16.1
PF1932	PFC_08920	hypothetical protein	-1.5	7.1
PF1940.1n	PFC_08880	hypothetical protein	1.3	15.4
PF1951	PFC_08825	asparagine ligase A	-1.8	7.6
PF1956	PFC_08800	fructose-bisphosphate aldolase	-2.4	9.9

Table S4 Continued				
Pfu ORF# OR Msed ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change (COM1)-(MW76) 40h**	**-log 10(p value)**
PF1960	PFC_08780	aldose reductase	-2.2	17.8
PF1971	PFC_08725	anaerobic ribonucleoside triphosphate reductase	-3.2	20.1
PF1972	PFC_08720	anaerobic ribonucleoside-triphosphate reductase activating protein	-3.2	8.8
PF1974	PFC_08710	thermosome, single subunit	-1.6	14.8
PF1977	PFC_08695	quinolinate synthetase	-1.6	8.8
PF1987	PFC_08630	hypothetical protein	-1.4	10.0
PF1990	PFC_08615	transcription antitermination protein NusG	-1.1	6.3
PF1992	PFC_08605	50S ribosomal protein L1P	-1.4	7.1
PF1993	PFC_08600	acidic ribosomal protein P0	-1.6	12.2
PF1998	PFC_08575	HAD superfamily hydrolase	1.8	9.1
PF2002	PFC_08550	sucrose transport protein	1.2	16.0
PF2004	PFC_08540	glycerol kinase	-1.1	6.1
PF2006	PFC_08530	NADH oxidase	-1.0	6.6
PF2007	PFC_08525	hypothetical protein	1.4	7.2
PF2008	PFC_08520	50S ribosomal protein L37Ae	2.2	11.2
PF2012	PFC_08500	elongation factor EF-2	-1.7	18.3
PF2018	PFC_08470	hypothetical protein	-1.3	6.2
PF2022	PFC_08450	hypothetical protein	1.1	6.4
PF2023	PFC_08445	DNA invertase	1.0	5.6
PF2024	PFC_08440	transposase	1.2	6.9
PF2031	PFC_08410	iron-sulfur cluster binding protein	1.1	9.3
PF2047	PFC_08330	l-asparaginase	1.0	6.5
PF2047.1n	PFC_08325	hypothetical protein	-1.2	6.9
PF2053	PFC_08290	AsnC family transcriptional regulator	1.4	7.8
PF2060.1n	PFC_08235	arylsulfatase	-3.2	12.7
	PFC_05475	sodium dependent transporter	3.4	19.5
	PFC_00270	HEPN domain-containing protein	2.1	10.0

Table S4 Continued				
Pfu ORF# OR Msed ORF#	**Pfu COM1 ID**	**Annotation**	**Average Log 2 fold change (COM1)-(MW76) 40h**	**-log 10(p value)**
	PFC_08040	permease	1.7	11.1
	PFC_02695	Regulatory protein, MarR	1.5	6.5
	PFC_03085	hypothetical protein	1.4	6.4
	PFC_00645	signal peptidase I	1.3	8.2
	PFC_02645	PepQ-1 X-pro dipeptidase	1.3	7.8
	PFC_05515	hypothetical protein	1.2	8.1
	PFC_03625	hypothetical protein	1.2	7.0
	PFC_04160	haloacid dehalogenase superfamily protein	1.1	9.3
	PFC_03630	hypothetical protein	1.1	6.0
	PFC_01895	hypothetical protein	1.1	7.0
	PFC_01230	proline permease	1.1	7.9
	PFC_03570	hypothetical protein	-1.9	8.7
	PFC_09300	hypothetical protein	-2.3	6.3
	Msed_1993	3-hydroxypropionate dehydrogenase (NADP(+))	-2.5	10.4
	Msed_0147	biotin carboxylase	-2.9	18.0
	Msed_2010	Biotin protein ligase	-4.0	23.3
	Msed_0390	carbonic anhydrase	-4.4	16.1
	Msed_1375	acetyl-CoA carboxylase carboxyltransferase subunit alpha / acetyl-CoA carboxylase carboxyltransferase subunit beta / propionyl-CoA carboxylase carboxyltransferase subunit	-4.9	23.2
	Msed_0148	biotin carboxyl carrier protein	-5.1	27.9
	Msed_0709	succinyl-CoA reductase (NADPH)	-5.5	24.3