Journal of
Proteomics & Bioinformatics

Research Article **Open Access**

Genome-wide Transcriptional Response during the Shift to N_2-fixing Conditions in *Heliobacterium modesticaldum*

Daniel Sheehy[1,2,3,4], Yih-Kuang Lu[2], Fawsia Osman[1], Zana Alattar[1], Catalina Flores[1], Hallie Sussman[1], Sahba Zaare[1], Maria Dooling[1], Amir Meraban[1], Patricia Baker[1,3], Jeffrey W. Touchman[2] and Kevin E. Redding[1,3]*

[1]School of Molecular Sciences, Arizona State University, Tempe, AZ, USA
[2]School of Life Sciences, Arizona State University, Tempe, AZ, USA
[3]Center for Bioenergy & Photosynthesis, Arizona State University, Tempe, AZ, USA
[4]School of Arts and Sciences, Boston University, Boston, MA, USA

Abstract

Heliobacteria are the only known phototrophic Firmicute; all known members of this group appear to be capable of N_2 fixation but incapable of CO_2 fixation. They are anoxygenic and possess the simplest photosynthetic apparatus known. The sequence of the 3.1-Mb genome of *Heliobacterium modesticaldum,* a moderate thermophile within the family Heliobacteriaceae, is publicly available. The focus of this study is to understand how this organism operates at a fundamental level by examining changes in its transcriptome during a shift from ammonium-containing medium to N_2-fixing conditions. RNA was purified from cells grown with pyruvate as the carbon source and ammonia or N_2 as the nitrogen source. After rRNA depletion, the RNA pool was sequenced using the Ion Torrent PGM platform. We found that the nitrogenase gene cluster was only expressed under N_2-fixing conditions, concomitant with increased expression of the high-affinity ammonium transporter. Most genes were down-regulated in N_2-fixing conditions by a factor of at least three. A drastic down-regulation of the highly expressed genes encoding proteins involved in the cyclic electron transport chain also occurred. The photosynthetic *pshA* transcript also decreased more than 100-fold but subsequent photochemical analysis demonstrated no large drop in the concentration of the reaction center protein complex. This indicates that there is a role for substantial translational regulation in some genes. The transcriptomic analyses revealed a network of differentially expressed genes in *H. modesticaldum*. This study represents the first step in the creation of a quantitative genome-scale metabolic model establishing *H. modesticaldum* as a model organism for the Heliobacteriaceae family.

Keywords: Heliobacteria; Next generation sequencing; Transcriptome

Introduction

Originally discovered in 1983 [1], heliobacteria have been a source of great interest, as they are the only known phototrophic group within the Firmicutes [2]. Heliobacteria also possess the simplest known photosynthetic apparatus. Comprised of a type I homodimeric reaction center (RC) in the cytoplasmic membrane, this photosynthetic apparatus uses a unique pigment, bacteriochlorophyll *g* (BChl *g*) [3]. Unlike other phototrophs, heliobacteria lack a peripheral antenna system. They are not autotrophic and must obtain carbon from organic molecules [4,5], but are capable of robust nitrogen fixation [6]. Like many other bacteria in the order Clostridiales, heliobacteria can grow fermentatively in the dark using simple organic acids, such as pyruvate [7]. In the light, they use the same carbon sources and grow at a significantly faster rate ("photoheterotrophic growth"). These characteristics identify heliobacteria as distinct among the seven photosynthetic bacterial lineages. Several heliobacterial species have been discovered, but only the genome sequence of *H. modesticaldum* has been reported [8]. *H. modesticaldum* is an obligatory anaerobe and is moderately thermophilic with an optimal growth temperature of 50-52°C [9]. The genome of *H. modesticaldum* is a single 3.1-Mb circular chromosome possessing 3,138 putative open reading frames (ORFs) [8]. The photosynthetic gene cluster is contained in a single operon, encoding the HbRC core polypeptide (PshA) and pigment biosynthetic enzymes (e.g. BchBNL). The order of these genes is the same as in the cluster in *Heliobacillus mobilis* [10]. While many enzymes involved in the central pathway of carbohydrate metabolism were predicted, it was recently reported that *H. modesticaldum* has an incomplete reductive TCA cycle lacking the key enzyme ATP-citrate lyase [11], as was also observed through ^{13}C-labeling [7]. Like all heliobacteria, *H. modesticaldum* is capable of fixing nitrogen, even at the elevated temperatures under which it grows [9]. Nitrogen fixation is catalyzed by the nitrogenase enzyme complex, and results in the reduction of atmospheric dinitrogen (N_2) to ammonium (NH_4^+) and the production of molecular hydrogen [12]. This process requires large amounts of chemical energy (16 ATP) and reducing power (8 Fd^{red}) to convert one N_2 to two molecules of NH_4^+ [13], which is then assimilated into many biomolecules. Sequence similarity predicts the use of a Mo-Fe group I nitrogenase consisting of a homodimer of NifD/K polypeptides [14]. The primary pathway for NH_4^+ assimilation in heliobacteria is the glutamine synthetase/glutamate synthase pathway [15]. This pathway is essential for growth, because glutamine is the primary intracellular nitrogen donor for purine and pyrimidine synthesis. Both ATP and reducing power are required in carbon metabolism, nitrogen assimilation, and hydrogen production, inextricably linking these pathways. In *H. modesticaldum*, the high-energy demand required for nitrogen fixation during diazotrophic growth has resulted in strict regulation of the *nif* genes encoding for nitrogenase. Thus, the addition of NH_4^+ to cultures

results in repression of nitrogenase activity [15,16]. However, the effects of N$_2$-fixing versus non-fixing conditions on genome-wide expression and metabolism in *Heliobacterium* remained to be determined. The sequencing of the genome of *H. modesticaldum*, along with recent proteomic studies, have provided insights into the energy metabolism of this phototroph. However, metabolic studies have been correlated to only a few genes related to energy and carbon metabolism at the transcriptomic level [17]. To enhance our understanding of the energy metabolism of *H. modesticaldum*, it is necessary to explore the entire mRNA link between genome and proteome. We report here a single-nucleotide resolution map of the *H. modesticaldum* Ice1 transcriptome under N$_2$-fixing and non-fixing conditions. In general, we observed low-level repression of transcription genome-wide upon a shift to N$_2$-fixing conditions. On the contrary, a few genes, such as the genes involved in N$_2$-fixation and ammonium scavenging, were upregulated. There were also several cases of even more drastic down-regulation, including the core genes of cyclic photophosphorylation.

Materials and Methods

Growth of *H. modesticaldum*

Isolated colonies of *H. modesticaldum* strain Ice1 were cultured in gel-rite media modified by Lin and Casida (1984) for thermophilic application [18] inside an anaerobic Coy glove box at 52°C under infrared lights at 780 nm. Cells were inoculated in Pyruvate-Yeast Extract (PYE) growth media [16], which contains 1 g/L of NH$_4$SO$_4$ as the source of nitrogen and "vitamin levels" of yeast extract (0.02%). The PYE-NH$_4^+$ medium was made by elimination of NH$_4$SO$_4$ and increasing the amount of Na$_2$S$_2$O$_3$•5H$_2$O from 0.2 g to 0.4 g. The pH was adjusted to 6.8 with H$_2$SO$_4$ prior to autoclaving. Growth was monitored spectroscopically at an Optical Density (OD) at 625-nm, as minimal photosynthetic pigments absorb at this wavelength in heliobacteria [3,17]. Cells were grown under anaerobic conditions at 52°C. From PYE or PYE-NH$_4^+$ conditions, 3 ml of cells in late exponential growth phase were inoculated into 300 ml of similar media. This ensured no traces of ammonia were carried over to the next generation of cells. Extraction was performed when cells were in mid-log phase of growth at an OD of 0.303 (PYE-NH$_4^+$) and 0.410 (PYE). A biological replicate for each condition was also prepared.

Isolation and purification of mRNA

Total RNA was isolated using the Purelink total RNA isolation kit (Invitrogen, USA). Cells were anaerobically extracted and lysed using the needle homogenization method offered in the kit protocol. Once total RNA was extracted the solutions were treated with DNaseI via the Ambion RiboPure Bacteria procedure to remove all genomic DNA. Depletion of the 16S and 23S rRNAs was performed via the Ambion MICROBE*Express* kit through subtractive hybridization with capture oligonucleotides. The resulting solution contained tRNAs, 5S rRNAs, and enriched mRNA.

Library preparation and Ion Torrent sequencing

Library preparation was performed using the Ion Total RNA-Seq Kit (Ambion, USA). Both sample and WT control RNAs (1 μg/μL HeLa total RNA) were fragmented using RNase III at 37°C for 10 minutes. The fragmented RNA was purified via Ambion's RiboMinus™ Concentration Module. The resulting mRNA yield was measured on the Bioanalyzer (Agilent, USA). These enriched mRNA samples were hybridized and ligated with Ion Torrent adaptors. Strand specificity was retained using the Ion Adaptor Mix containing oligonucleotides with a single-stranded degenerate sequence at the 3' end and a defined

sequence at the 5' end. This effectively constrains the RNA orientation, with sequencing only performed from the 5' end of the sense strand. Reverse transcription was performed and the cDNA was purified and size selected using Agencourt's AMPure® XP reagent for an optimal fragment size between 30-200 bp. The cDNA samples were amplified using the provided Ion 5' and 3' PCR Primers for 16 cycles. Fragment size was verified on the Bioanalyzer (Figure S1). The prepared cDNA libraries were loaded on the chip and sequenced using the Ion PGM sequencer per the manufacturer's instructions (Life Technologies, USA).

Bioinformatic analysis of RNA-seq data

The sequenced reads captured on the Ion Torrent platform were analyzed by following procedure described here. Initially, the sequencing reads were subject to filtering out of those with the poor quality and more than 50% of Ns in reads using PRINSEQ [19]. Bowtie2 [20] with default parameters was used to align filtered RNA-seq reads against *H. modesticaldum* chromosome as the reference genome. The resulting sequence alignment files were imported into Partek Genomics Suite (Partek Inc., St. Louis, MO) to compute raw and fragments per kilobase of exon model per million mapped (RPKM) reads for the normalized expression values of each transcript. A stringent filtering criterion with RPKM value of 1.0 [21] was used to obtain expressed transcripts. The RPKM values of filtered transcripts were log-transformed using log2 (RPKM + offset) with an offset value of 1.0, and fold changes in transcript expression, differential expression, and p-values were generated from these using the Partek software with default settings. Metabolic pathways that were significantly enriched between +/-NH$_4^+$ were identified using Cytoscape [22] with the reference pathways of *H. modesticaldum* from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database [23].

Real-time qPCR

A 1 μg concentration of non-fragmented mRNA was added to an iScript cDNA Synthesis Kit (Bio-Rad, USA) and reverse transcription was performed. Primers were designed and synthesized for qPCR using the SsoFast EvaGreen Supermix kit (Bio-Rad, USA). Enzyme activation was performed at 95°C for 1 minute, followed by 40 cycles of denaturation at 95C for 15 seconds and annealing/extension at 57°C for 30 seconds. The C$_t$ values were determined based on the PCR cycle number that crossed an arbitrarily defined threshold. Gene expression was determined using the 2$^{-\Delta\Delta Ct}$ method [24]. Three technical replicates were performed and the mean values were reported (Table S1). Primers used are listed in Table S2.

Reaction center measurements

Cells were grown in PYE and PYE-NH$_4^+$ to mid-log phase of growth. Aliquots from both cultures were diluted to OD$_{650}$ = 0.30 to equalize cell density between the two conditions. Cells were transferred into a 1 cm anaerobic cuvette and brought to room temperature before analysis with a JTS-10 LED spectrometer (Bio-Logic; Claix, France). Absorption from P$_{800}$ was monitored using 10 μs pulses of light from a 810 nm LED; 780 nm high-pass filters were placed in front of the detectors to block actinic light provided by a saturating flash (31 mJ) from a frequency-doubled Nd/YAG laser [25]. The bleaching values were measured six times per sample with a 90-s dark adaptation and cuvette inversion between measurements. After the photobleaching measurements, an acetone extraction was performed on both samples anaerobically; the supernatant was measured for BChl *g* concentration and the pellet was resuspended in 50 μl of 50 mM MOPS (pH 7). A bicinchoninic acid assay was performed on the pellet and the

resulting protein concentration values were used to normalize both the photobleaching kinetic data and BChl *g* concentrations (Thermo Scientific, USA).

Results

The baseline transcriptome

In order to be able to correlate the transcriptomic data with previous proteomic studies, it was important to use comparable growth conditions. Therefore, we chose phototrophic growth in Pyruvate Yeast Extract (PYE) medium as the baseline condition, as it is the typical growth mode used in most studies on this organism. The presence of yeast extract in this media was shown to have no significant effect on the expression levels of genes involved in metabolism and other cellular functions [17]. Following cellular extraction during mid-log phase growth, depletion of 16S and 23S rRNA was performed. This procedure reduces the sequencing bias towards highly expressed genes, as it was shown that 98% of the transcriptome consisted of rRNA in three different bacterial species when total RNA was used for transcriptome analysis [26]. These highly expressed transcripts reduce the coverage of mRNA reads and undermine transcriptome analysis accuracy. Electrophoretic analysis of the RNA pools indicated that the depletion was largely successful (Figure S1 of Supplemental Material). After rRNA depletion, cDNA libraries of three biological replicates were generated and sequenced on the Ion Torrent. The sequencing reads of the PYE culture were pooled together, resulting in 3.75 million reads and 256 Mbp (> 83-fold sequencing depth of the genome). Following the same procedure as described above, the sequenced reads of the PYE-NH$_4^+$ culture resulted in 3.71 million reads and 345 Mbp (>112-fold genome coverage). Further, we filtered the sequencing reads with poor quality or more than 50% ambiguous bases, and the remaining reads were subject to the alignment. Reads totaling 74 and 69 Mbps from the ammonium-replete and ammonium-deplete cultures, respectively, were mapped onto the genome of *H. modesticaldum*. Both data sets resulted in a single-nucleotide resolution transcriptome with an average coverage of 24- and 22-fold per nucleotide, respectively. The inherent strand specificity of the Ion Torrent methodology revealed several interesting features regarding *H. modesticaldum's* transcriptome. First, one of the genome strands encodes two thirds of the protein coding sequences while the opposite strand encodes the remaining coding sequences (Figure S2A). This observation is consistent with the location of open reading frames on the sequenced genome [8].

Additionally, the strand-specific sequencing revealed non-coding RNAs, which appear to include *cis*-encoded antisense RNAs. These are RNAs that both overlap and are complementary to the sense strand. Examples of antisense RNAs in *H. modesticaldum's* transcriptome can be seen throughout the NiFe hydrogenase cluster as both cis (red) and complementary to the sense strand (green) (Figure S3). The function of antisense RNAs in bacteria is increasingly being recognized as an important regulator of metabolic and physiological processes [27]. In the cyanobacterium *Synechocystis* sp. PCC 6803, there are antisense RNAs for approximately 10% of the total protein-coding genes, indicating that regulation by antisense RNA may be a significant form of gene regulation in phototrophic bacteria [28]. We have annotated 775 potential antisense transcripts using an arbitrary cut-off of ≥ 10 in coverage (Data Set S1), indicating that the use of antisense RNA could be as high (or even higher) in heliobacteria.

Another feature of the baseline transcriptome is several regions of high intergenic transcription, which may indicate non-annotated genes. For instance, the region between 1352.5-1358.5 Kbps contains a high level of intergenic transcription, and nBLAST attributes a 99 percent identity to 16S rRNA for both the non-annotated transcripts and the intervening "hypothetical gene" HM1_3154 (Figure S2B). Table S1 lists 53 transcripts that map to intergenic regions (i.e. not overlapping with genes on either strand) and have coverage ≥ 10 and minimal length of 50 bases. BLAST searches were performed on all peptides potentially encoded by this region, but the vast majority did not have significant matches. Three of them are potential pseudogenes, with significant matches to peptides 50-55 residues long. One transcript, however, has 58% identity (73% similarity) to the ParA chromosome partitioning protein of *Paenibacillus polymyxa* over the entire length of 282 residues, and thus very likely represents a gene that was missed in the primary genome annotation (Table S1).

The ammonium-deplete transcriptome

With the general features of *H. modesticaldum's* transcriptome established, it becomes possible to quantify the expression of the genome at the RNA level and how this changes in response to an environmental change. We repeated the RNA-seq experiment with cells deprived of ammonium, thus requiring them to fix N$_2$ present in the headspace gas (~97% N$_2$). Cells were diluted 100-fold into PYE medium lacking ammonium (PYE-NH$_4^+$) and grown to saturation. This culture was then used to inoculate a new PYE-NH$_4^+$ culture. The growth curves of the cultures grown in ammonium replete (PYE) and ammonium deplete (PYE-NH$_4^+$) media are shown in Figure 1, and cells were harvested for RNA extraction at an optical density (OD) of 625 nm at 0.410 (PYE) and 0.303 (PYE-NH$_4^+$). In both cases, cells were grown photoheterotrophically using pyruvate as the carbon source. We determined that the cells were dependent upon the N$_2$ in the headspace for their nitrogen needs by also inoculating a PYE-NH$_4^+$ culture in a bottle that had been purged with argon to remove all N$_2$. This culture stopped growing very soon after inoculation and did not attain a density more than 25% that of the PYE-NH$_4^+$ culture that had a headspace of 97% N$_2$ (data not shown).
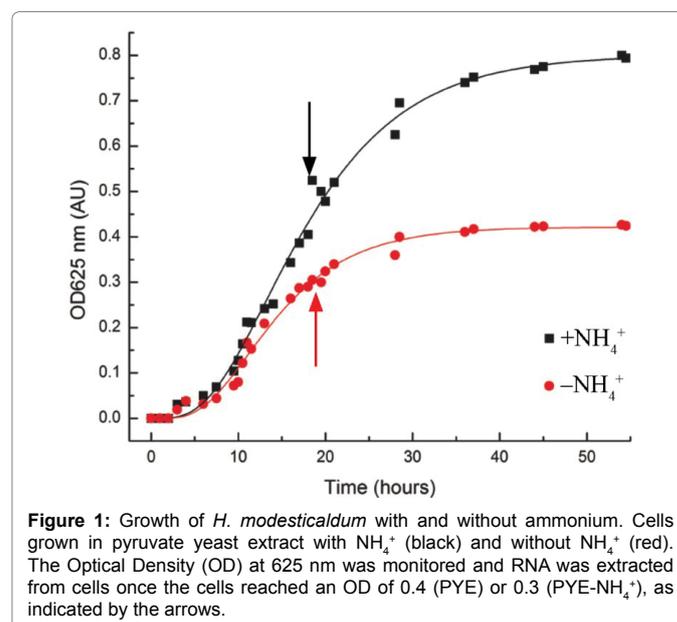

Figure 1: Growth of *H. modesticaldum* with and without ammonium. Cells grown in pyruvate yeast extract with NH$_4^+$ (black) and without NH$_4^+$ (red). The Optical Density (OD) at 625 nm was monitored and RNA was extracted from cells once the cells reached an OD of 0.4 (PYE) or 0.3 (PYE-NH$_4^+$), as indicated by the arrows.

With the high energetic requirements of N$_2$ fixation, we predicted a significant change in metabolism from expression levels of genes in the *nif* operon, which encodes the nitrogenase enzyme and ancillary functions, which was observed. Although the genes encoding nitrogenase (*nifDHK*) are hardly expressed in ammonium-replete medium, they are strongly expressed in the medium without ammonium (Data Set S3, hmo00910b). We also observed a decreased rate of growth (Figure 1), consistent with previous observations [17] and indicative of a global change in metabolism. Our primary findings are detailed below, and all genes are shown in Data Set S2 along with the corresponding expression values.

If one removes all genes for which no reads were found in either condition (and for which a change in expression level is therefore impossible to calculate), one is left with 2430 genes, representing 73% of the 3268 genes annotated in the genome. This tends to remove genes that are either very poorly expressed in general and/or for which the change in expression is very large. The average fold change in expression upon removal of ammonium for these 2430 genes is -3.93± 8.05; i.e. a 4-fold decrease in expression, with a large spread in values from a 12-fold decrease to a 4-fold increase (within one standard deviation). The overall change in expression across the genome appears to be significant (P-value = 0.0124, based on a two-tailed, paired t-test of the data; Data Set S2). The number of genes displaying different fold change values is tabulated in Table 1. There is a strong bias towards down-regulation. For example, there are over 20 times more genes with a >10-fold decrease in expression than there are genes with a >10-fold increase in expression (224 *vs.* 10). It is important to note that some of the most highly expressed genes (e.g. ribosomal RNA and protein) exhibited almost no change in message level, which explains why most genes could exhibit a decrease in expression. It is thus likely that the total cellular RNA level hardly decreased at all, as one might expect.

Discussion

In order to put into useful context, the changes in expression of the 2757 genes for which we possess data, we have made use of the annotated metabolic pathways in the Kyoto Encyclopedia of Genes and Genomes (KEGG). All of the pathways in the KEGG for the *H. modesticaldum* genome were analyzed and a table was constructed for each pathway by incorporating the expression data for all genes encoding products involved in the pathway. These tables are named according to the KEGG numbering system (e.g. hmo00010 is the glycolysis/gluconeogenesis pathway). Each can be found as a pathway in one of the Data Sets described in the Supplemental Materials. Each Data Set contains a set of pathways that are similar in nature, and the organization of the Data Sets mirrors the discussion that follows. The results from all of these pathways are summarized in Table 2. The statistical significance of the overall expression changes of the genes in each pathway were assessed by paired Student's t-test (two-tailed) of the RPKM values for each gene in the two conditions (plus and minus ammonium). Overall expression changes with P-values lower than 0.05 were deemed insignificant. By this measure, expression changes in 39 of the 62 analyzable KEGG pathways were significant. To correct for the multiple comparisons problem, we applied the Bonferroni correction (i.e. α=0.05/62). After this procedure, only 7 of the KEGG pathways were deemed statistically significant (Table 2).

Nitrogen fixation, uptake, and incorporation

Providing nitrogen in the form of ammonium eliminates the need to fix N$_2$ via the energy-expensive nitrogenase system (encoded by the *nif* operon), which requires 16 ATP and 8 Fd^{red} for each N$_2$ molecule fixed [13]. Not only is nitrogenase activity absent in cells grown in medium with ammonium, but the activity is lost very quickly after addition of ammonium to N$_2$-fixing cultures, in this species [9] and
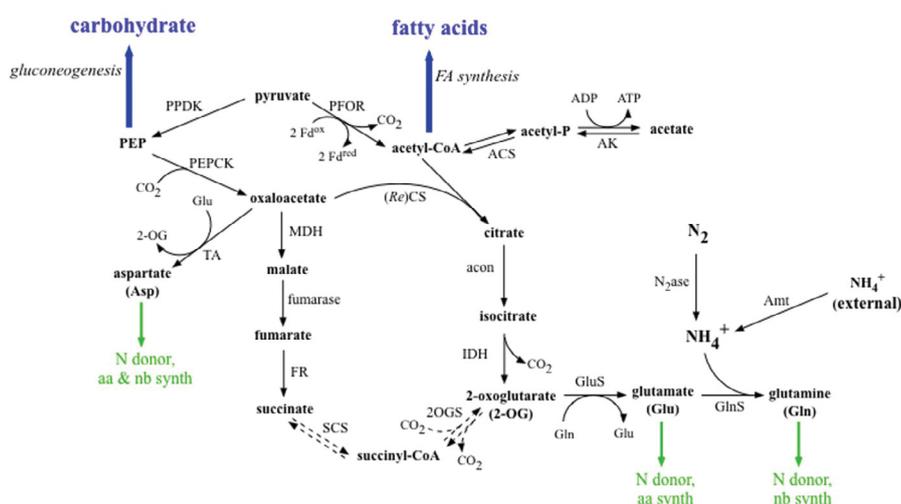

Figure 2: Proposed connections between fundamental carbon metabolism and nitrogen assimilation. The major pathways of carbon metabolism when cells are grown with pyruvate as the sole carbon and electron source are shown, along with the major nitrogen incorporations reactions. Ammonia is transported into the cell when available, or is made by nitrogenase by reduction of N$_2$ when unavailable. The ammonia is first incorporated into an organic molecule by Gln synthetase. Gln is used to make Glu, which can be used to make Asp and the rest of the amino acids via transaminases. Gln, Glu, and Asp are the nitrogen donors for almost every other N-containing biomolecule, including amino acids (aa) and nucleobases (nb). For simplicity, the consumption/generation of reducing equivalents and ATP is largely left off this figure. Since the direction of the TCA cycle between succinate and 2-OG is unclear at this time, the arrows of the SCS and 2OGS reactions are dotted. Abbreviations: 2OGS, 2-oxoglutarate synthase; Acon, aconitase; Amt, ammonium transporter; CS, citrate synthase; FR, fumarate reductase; GlnS, glutamine synthetase; GluS, glutamate synthase; IDH, isocitrate dehydrogenase; MDH, malate dehydrogenase; N$_2$ase, nitrogenase; PEPCK, phosphoenolpyruvate carboxykinase; PEP, phosphoenolpyruvate; PFOR, pyruvate:ferredoxin oxidoreductase; PPDK, pyruvate phosphate dikinase; SCS, succinyl-CoA synthetase.

Fold change[1]	# genes[2]
<20	71
<10	224
<5	629
<2.5	1278
-1.25 to -2.5	557
-1.25 to +1.25	252
+1.25 to +2.5	233
>2.5	110
>5	27
>10	10
>20	9

Table 1: Description: Summary of effect of ammonium depletion upon gene expression.

[1]Fold change is defined as in Table 1 (e.g. -20 is a 20-fold decrease, -1.25 is a 20% decrease, +5 is a 5-fold increase, etc.)

[2]The number of genes with the indicated fold change is from the 2430 genes with reads found in both conditions.

KEGG pathway (common name)	Pathway Description	Average fold change (n)[1] P-value of RPKM differences[2] Summary of findings[3]
hmo00010 Glycolysis/ Gluconeogenesis	While *H. modesticaldum* does not metabolize glucose, the enzymes of the glycolysis pathway are likely used for gluconeogenesis, which generates glucose from non-carbohydrate sources (e.g. pyruvate), and for central carbon metabolism.	-8.0 ± 4.9 (14) P-value = 0.00284 The genes involved in the gluconeogenesis pathway were all downregulated in the absence of ammonium by a factor of ~8.
hmo00020 Tricarboxylic acid (TCA) Cycle	*H. modesticaldum* cannot use oxygen as an electron acceptor, since it is a strict anaerobe, so this pathway is not used for respiration. The TCA cycle is likely used to produce precursors for other biosynthetic pathways (e.g. succinate, 2-oxoglutarate, and citrate).	-3.6 ± 3.5 (17) P-value = 0.00644 Expression of the genes in the TCA pathway was reduced by a factor of ~4 on average, except for the *pckA* gene encoding phosphoenolpyruvate carboxykinase (PEPCK), which was downregulated by approximately 16-fold.
hmo00030 Pentose Phosphate Pathway	This pathway generates NADPH and 5-carbon sugars through the oxidation of glucose and subsequent isomerization of pentoses. Primarily anabolic, the pentose phosphate pathway serves as an alternative to glycolysis and as a mechanism to generate pentoses and NADPH. The lack of the oxidative enzymes (glucose-6-phosphate dehydrogenase and gluconate-6-phosphate dehydrogenase) indicates that this pathway does not exist, and the pentose isomerases are used to interconvert carbohydrates. (A reductive pentose phosphate pathway also does not exist, due to the lack of Rubisco.)	-7.5 ± 4.2 (13) P-value = 0.0171 The genes involved in the pentose phosphate pathway were all downregulated in the absence of ammonium by approximately 7.5-fold, on average, with individual fold changes ranging from 2.8-fold to 14.7-fold reduction.
hmo00051 Fructose/Mannose metabolism	The purpose of the enzymes in the fructose and mannose pathways is to convert fructose and mannose to glycolysis precursors. Mannose is converted to fructose, which is then eventually converted to glyceraldehyde, where it can then enter the gluconeogenesis pathway.	-5.4 ± 5.3 (12) P-value = 0.0772 Expression of genes involved in the fructose and mannose metabolism was not significantly affected by the absence of ammonium.
hmo00052 Galactose Metabolism	This is part of the Leloir pathway, UDP-galactose is first converted to UDP-glucose, then to glucose-1 phosphate. The major enzymes are hexokinase, UTP-glucose-1-phosphate uridylyltransferase, and UDP-glucose 4-epimerase.	-2.5 ± 0.8 (4) P-value = 0.25 Expression of genes involved in galactose metabolism was not significantly affected by the absence of ammonium.
hmo00061 Fatty Acid Biosynthesis	Fatty acid biosynthesis relies on the fatty acid synthases and acetyl CoA carboxylases. Fatty acid synthases are large multi-enzyme complexes that create fatty acids from acetyl-CoA and malonyl-CoA. Fatty acids are long aliphatic chains with a terminal carboxylate, and are used in many applications, from energy storage to membrane structure. Acetyl-CoA carboxylase plays an essential role in regulating fatty acid synthesis; it catalyzes the committed step in fatty acid synthesis, the production of malonyl-CoA	-1.7 ± 2.8 (10) P-value = 0.35 Expression of genes involved in fatty acid biosynthesis was not significantly affected by the absence of ammonium.
hmo00071 Fatty Acid Degradation	*H. modesticaldum* has never been reported to use fatty acids as carbon sources or energy sources; its three defined carbon sources are pyruvate, lactate and acetate. Inspection of related bacteria (*Desulfitobacterium hafniense* Y51 and *Desulfotomaculum reducens*) revealed that the same enzymes were missing (dehydrogenase, hydratase, CPT2 and CPT1). Therefore, it is unlikely that β-oxidation of fatty acids takes place in *H. modesticalum*.	-6.4 ± 7.9 (6) P-value = 0.065 The only two genes with significant expression in the fatty acid degradation pathway were HM1_0073 and *aas*, which were downregulated in the absence of ammonium by about 6 fold and 2 fold, respectively. HM1_0073 codes for 3-hydroxyacyl-CoA dehydrogenase while aas codes for 2-acylglycerophospho-ethanolamine acyltransferase, both of which are used in fatty acid biosynthesis.

hmo00130 Quinone Biosynthesis	This organism appears to make use of the alternate pathway in Gram-positive bacteria that proceeds via futalosine. Some genes (*mqnB* and *mqnD*) appear to be missing, but these activities may be supplied by homologs (i.e. MenC/D) and 2 other genes in the cluster containing other menaquinone synthesis genes. Methyl group addition is catalyzed by MenG, while decarboxylation and isoprenylation appear to be catalyzed by UbiD/X and UbiA, respectively.	-6.3 ± 7.4 (7) P-value = 0.00585 Expression of genes involved in menaquinone biosynthesis was not significantly affected by the absence of ammonium.
hmo00190 Oxidative Phosphorylation	Because *H. modesticaldum* is strictly anaerobic, it does not use oxygen as an electron acceptor. However, four of the five key enzyme complexes of the general oxidative phosphorylation pathway are retained (Complexes I to III and ATP synthase), which generate most of the organism's ATP during phototrophic growth. The cytochrome *bd* complex is also present, most likely as a way to scavenge O$_2$.	-2.2 ± 4.3 (31) P-value = 0.000692 Expression of all genes involved in oxidative phosphorylation decreased by an average of 2.5-fold. (Changes are deemed statistically significant after Bonferroni correction.)
hmo00190b Light-driven electron flow	This table was created to include all of the components lacking in hmo00190 in order to complete the presumed cyclic electron flow pathway. This includes the cyt c_{553}, the HbRC subunits, and ferredoxins, as well as FNR (ferredoxin-NADP+ oxidoreductase).	-73.5 ± 42.3 (5) P-value = 0.0244 The reaction center, as well as its electron donor (cyt c_{553}) and acceptors (ferredoxins PshB1 and PshB2) are highly expressed, but their expression is strongly reduced in the absence of ammonium. Expression of FNR is much lower and is only reduced by a factor of 4-5.
hmo00190c Hydrogenases	The [NiFe] hydrogenase is likely used to oxidize H$_2$ and reduce the quinone pool. There are 2 genes for [FeFe] hydrogenases, which typically oxidize ferredoxin and reduce protons to H$_2$, or the reverse, but it is unclear if they are active, given the apparent absence of maturation genes.	-6.4 ± 5.2 (15) P-value = 0.0000316 The expression of genes related to production of the hydorgenase enzymes appears to be reduced by a factor of 6-7. This finding does not support their use in recycling the hydrogen produced by the nitrogenase during nitrogen fixation. (Changes are deemed statistically significant after Bonferroni correction.)
hmo00230 Purine Synthesis	Purine metabolism is an anabolic process that begins with ribose-5-phosphate, resulting in the synthesis of nucleic acids and nucleotides. Specifically, ribose-5-phosphate is the precursor of IMP, which can then be used to synthesize ATP and GTP. GTP and ATP can then be used for RNA synthesis or converted to dATP and dGTP and used in DNA synthesis. Ribsose-5-phosphate may also be converted to 5-aminoimidazole ribonucleotide and feed into thiamine metabolism	-3.7 ± 3.2 (24) P-value = 0.0229 Expression of all genes involved in the synthesis of purine decreased by an average of about four-fold.
hmo00240 Pyrimidine Synthesis	The purpose of this pathway is to synthesize the pyrimidines uracil, cytosine, thymine, and their derivatives. The reactions leading up to the end products are anabolic in nature. Some of the key enzymes of the main branch (from Asp and carbamoyl phosphate to UMP) are aspartate carbamoyltransferase, and orotate phosphoribosyl-transferase. The pyrimidine metabolic pathway also connects with metabolic pathways of pentose phosphate and several amino acids.	-3.6 ± 3.9 (15) P-value = 0.0116 Expression of all genes involved in the synthesis of pyrimidine decreased by an average of about four-fold.
hmo00250 Alanine, Aspartate, & Glutamate Metabolism	This pathway focuses on the synthesis of the amino acids alanine, aspartate, and glutamate. The reactions involved are generally anabolic as they synthesize the compounds from smaller precursors. The primary branch utilizes the enzymes glutamate synthase, glutamine synthetase, glucosamine-fructose-6-phosphate aminotransferase, amidophosphoribosyltransferase, and carbamoyl-phosphate synthase. The pathway connects to glycolysis/gluconeogenesis and a number of side/feeder pathways.	-2.4 ± 2.7 (11) P-value = 0.103 Expression of the genes involved in metabolism of alanine, aspartate and glutamate was not significantly affected by the absence of ammonium.
hmo00260 Glycine, Serine, & Threonine Metabolism	The purpose of this pathway is to synthesize the amino acids glycine, serine, and threonine. The reactions are generally catabolic as more complex molecules yield simpler ones. More branches yield glycine and serine than threonine, since they can generally be reused in a broader range of molecules; some of the key enzymes are serine hydroxymethyl-transferase, phosphor-glyceromutase, and phosphoserine aminotransferase. The pathway leads into a number of other pathways, such as glycolysis and glyoxylate metabolism.	-2.8 ± 3.5 (12) P-value = 0.0273 Expression of all genes involved in the metabolism of serine, glycine and threonine decreased by an average of 3-fold.
hmo00270 Cysteine & Methionine Metabolism	This pathway focuses on synthesizing the amino acids cysteine and methionine and their derivatives. The reactions are primarily anabolic in nature. Many of the branches lead to cysteine and homocysteine; important enzymes in the pathway include serine O-acetyltransferase, cysteine synthase A, and those of the S-adenosyl methionine cycle. The pathway connects with the synthesis of some other amino acids and their derivatives, as well as sulfur metabolism and methionine salvage.	-2.9 ± 3.5 (17) P-value = 0.0978 Expression of the genes involved in metabolism of cysteine and methionine was not significantly affected by the absence of ammonium. (Note that most of these genes decreased by a factor of ~3.4, while *metC* showed an increase in expression of 4.4-fold.)

hmo00280 Valine, Leucine, & Isoleucine Degradation	The purpose of this pathway is to degrade the amino acids valine, leucine, and isoleucine into intermediates that can then be used in the citric acid cycle, pyrimidine metabolism, propanoate metabolism, and for the biosynthesis of more complex molecules. The lack of a branched chain □-keto acid dehydrogenase and most of the subsequent enzymes specific to this pathway, however, indicates that it does not exist in *H. modesticaldum*. The aminotransferase is used in the biosynthetic pathway, and the other enzymes are used in other metabolic pathways. Furthermore, this organism cannot use branched chain amino acids as a carbon source.	-7.0 ± 7.7 (7) P-value = 0.0169 Expression of genes assigned to this pathway underwent a 7-fold reduction, but this pathway does not exist in this organism.
hmo00290 Valine, Leucine and Isoleucine Biosynthesis	The main purpose for this pathway is to synthesize valine, leucine, and isoleucine, which contain branched aliphatic side chains. The starting point is pyruvate and 2-oxo-butanoate, which is derived from threonine or pyruvate. Acetolactate synthase is a key enzyme that catalyzes the reaction between threonine and pyruvate. Isoleucine and valine synthesis utilize the same enzymes throughout the pathway. During the synthesis of valine, 2-oxoisovalerate can be converted to 2-isopropylmelate to initiate the synthesis of leucine.	-7.4 ± 1.6 (8) P-value = 0.00124 Expression of genes involved in valine, leucine and isoleucine biosynthesis decreased by a factor of 7.5.
hmo00300 Lysine Biosynthesis	The purpose of this pathway is to synthesize lysine from aspartate or homoserine. Some important enzymes in this reaction are aspartate kinase, homoserine dehydrogenase, and diaminopimelate decarboxylase. This pathway also feeds into the pathway for peptidoglycan biosynthesis.	-2.5 ± 1.1 (9) P-value = 0.0328 On average, genes involved in the biosynthesis of lysine decreased by 2.5-fold.
hmo00330 Arginine and Proline Metabolism	The purpose of this pathway is to synthesize the amino acids arginine and proline. The reaction of glutamate to ornithine is catalyzed by the enzyme acetyltransferase. The reaction of ornithine to arginine is catalyzed by ornithine carbamoyltransferase. The synthesis of proline from glutamate is catalyzed by □-glutamyl kinase.	-2.4 ± 2.0 (18) P-value = 0.0463 On average, the genes involved in the metabolism of arginine and proline decreased by 2.5-fold.
hmo00340 Histidine Metabolism	The purpose of this pathway is to synthesize histidine, originating from the pentose phosphate pathway. The initial step is the reaction of PRPP with ATP phosphoribosyltransferase. Phosphoribulosyl-formimino-AICAR-phosphate is a branchpoint metabolite; conversion to AICAR leads to purine metabolism. Conversion to imidazole-glycerol-3-phosphate by imidazoleglycerol-phosphate dehydratase will produce histidine.	-12.1 ± 7.4 (8) P-value = 0.0000175 Expression of genes involved in histidine metabolism were down-regulated by 12-fold on average. (Changes are deemed statistically significant after Bonferroni correction.)
hmo00400 Phenylalanine, Tyrosine, & Tryptophan Biosynthesis	The purpose of this pathway is to synthesize the aromatic amino acids: phenylalanine, tyrosine and tryptophan. D-Erythrose-4-phophate is converted to chorismate by the shikimate pathway. Thereafter, each aromatic amino acid is synthesized by distinct pathways. If anthranilate synthase is used, the product (anthranilate) will go through a series of reactions to produce Tryptophan. Chorimate can also be converted to prephenate by chorismate mutase. This can then lead to the production of phenylalanine with the initial enzymes chorismate mutase/prephenate dehydratase present or tyrosine if only prephenate dehydrogenase is used.	-11.4 ± 7.1 (14) P-value = 0.0162 The expression of genes involved in phenylalanine, tryptophan and tyrosine biosynthesis decreased by an average of 14.5-fold. A strange exception was the two *aroF* genes: the highly expressed one decreased by 59-fold, while the lower expressed one slightly increased (1.4-fold) in the absence of ammonium.
hmo00450 Selenocompound Metabolism	This metabolism represents the process by which Selenium-based species are incorporated into biological compounds, which in turn can be incorporated into proteins. The pathways involved constitute an anabolic process. The principle pathways for this metabolism used by *H. modesticaldum* involve the conversion of selenate into selenocysteinyl-tRNA^{Sec} and the conversion of Selenocysteine into Seleno-methionyl-tRNA^{Met}.	-4.25 ± 3.95 (8) P-value = 0.0277 Expression of genes involved in selenocompound metabolism were reduced an average of 5-fold. A notable exception to this trend – *metC* – showed a 4.5-fold increase in expression.
hmo00473 D-Alanine Metabolism	This pathway interconverts D-Alanine and L-Alanine and converts D-Alanine to D-Alanyl-D-Alanine, which is used in peptidoglycan metabolism.	-1.76 ± 0.57 (2) P-value = 0.33 (not significant) The highly expressed alanine racemase is reduced by ~20-fold.
hmo00500 Starch & Glycogen metabolism	This pathway is used to synthesize □(1,4)-linked glucose polymer using UDP-glucose as a monomer donor, which is synthesized by UTP-glucose-1-phosphate uridylyltransferase from glucose-1-phosphate. The presence of branching and debranching enzymes suggest that the polymer present in this organism has occasional □(1,6)-linked branchpoints and thus would resemble glycogen. Glycogen phosphorylase is used to liberate glucose monomers as glucose-1-phosphate when the cell has need of the carbon source.	-2.9 ± 1.3 (8) P-value = 0.0214 On average, expression of the genes in this pathway decreases ~3-fold in the absence of ammonium.
hmo00520 Amino/Sugar Nucleotide Synthesis	This pathway involves the attachment of sugars to nucleotides and the modification of these nucleotide sugars. *H. modesticaldum* has mechanisms to create and modify UDP and GDP sugars. These pathways interconnect with a variety of other metabolic processes. Glycolysis, gluconeogenesis, and the fructose and mannose metabolisms provide sugars to start the anabolic process.	-4.4 ± 3.9 (16) P-value = 0.000959 The expression of genes involved in amino/sugar nucleotide synthesis decreased by an average of four-fold.

hmo00540 Lipopolysaccharide Biosynthesis	*H. modesticaldum* only possess the part of this pathway that produces the S-layer glycoprotein from sedoheptulose-7-phosphate. This is an anabolic process. Note that enzyme 2.7.7.71 (D-glycero-alpha-D-manno-heptose 1-phosphate guanylyltransferase), which is required for assembly of S-layer glycoprotein in some Gram-positive bacteria, seems to be missing.	-1.4 ± 2.3 (5) P-value = 0.23 The expression of genes in lipopolysaccharide biosynthesis was not significantly affected by the absence of ammonium.
hmo00550 Peptidoglycan Synthesis	Peptidoglycan is a macromolecule made of long amino-sugar polymers, cross-linked by short peptides. It forms the cell wall that encases the cell membrane, providing the cell with structural support and protection. The glycan strands are typically comprised of repeating N-acetylglucosamine (GlcNAc) and N-acetylmuramic acid (MurNAc) disaccharides. Each MurNAc is linked to a peptide of three to five amino acid residues. Disaccharide subunits are first assembled on the cytoplasmic side of the bacterial membrane on a polyisoprenoid anchor (lipid I and II). Polymerization of disaccharide subunits by transglycosylases and cross-linking of glycan strands by transpeptidases occur on the other side of the membrane.	-5.0 ± 4.7 (20) P-value = 0.000970 The expression of genes involved in peptidoglycan synthesis decreased by an average of five-fold.
hmo00561 Glycerolipid metabolism	Glycerolipid metabolism is an anabolic process that results in the production of phosphatidic acids. Phosphatidic acids are composed of a glycerol backbone, a fatty acid bonded to carbon-1, a fatty acid bonded to carbon-2, and a phosphate group bonded to carbon-3. Glycerol is first phosphorylated by glycerol kinase to glycerol 3-phosphate. Glycerol-3-phosphate is then acylated at positions 1 and 2. The final product is 1,2-diacyl- glycerol-3-phosphate (phophatidate), which serves as the biosynthetic precursor for the formation of all glycerolipids.	-8.4 ± 0.3 (4) P-value = 0.0582 The expression of genes involved in glycerolipid metabolism was not significantly affected by the absence of ammonium. (Note that expression of most genes in the glycerolipid metabolism pathway was reduced by a factor of ~8 on average, except for glycerol kinase, which showed a 3.4-fold increase in expression.)
hmo00564 Glycerophospholipid metabolism	Glycerophospholipid metabolism is an anabolic process resulting in the formation of glycerol-based phospholipids, which are a structural component of the cell membrane. Phosphatidate is the precursor. Various enzymes attach different alcohol headgroups in a phosphoester linkage. In Heliobacteria, these headgroups include serine, ethanolamine, and glycerol. (They lack choline.)	-5.9 ± 3.0 (7) P-value = 0.00458 Expression of all the genes in the glycerophospholipid metabolism pathway was reduced by a factor of approximately 6 on average.
hmo00620 Pyruvate Metabolism	Pyruvate plays a vital role in the metabolism of *H. modesticaldum*, and serves as one of its major sources of carbon. Pyruvate can be carboxylated to OAA by pyruvate carboxylase. Alternatively, it can also be converted to PEP by pyruvate phosphate dikinase. Incorporation of carbon into biomolecules then can be achieved through PEPCK, which converts PEP to OAA. In either case, OAA can then be converted to fumarate and succinate by the "left arm" of the TCA cycle.	-5.3 ± 4.9 (13) P-value = 0.0427 Expression of the genes in the pyruvate metabolism pathway was reduced by a factor of approximately 6 on average.
hmo00630 Glyoxylate Metabolism	The glyoxylate cycle does not appear to exist in *H. modesticaldum*, as it lacks the key enzymes, isocitrate lyase and malate synthase. The role of the *glcD* gene product (glycolate oxidase) is unclear at this time.	-4.45 (1) P-value = undefined Expression of *glcD* is downregulated by a factor of 4-5.
hmo00640 Propanoate Metabolism	Propanoate metabolism results in the production of propionyl-CoA, which is converted to succinyl-CoA, an intermediate in the TCA cycle.	-2.6 ± 3.3 (15) P-value = 0.00305 Expression of genes in the propanoate metabolism pathway was reduced by a factor of approximately 3 on average.
hmo00650 Butanoate Metabolism	Butanoate is a four-carbon fatty acid formed by bacterial fermentation of carbohydrates. It is often used for ketone body production and fed into the citrate cycle, glycolysis, or glutamate synthesis. It does not appear that butanoate is metabolized in *H. modesticaldum* due to the absence of many key enzymes.	-3.4 ± 3.5 (7) P-value = 0.0381 Expression of genes in the butanoate metabolism pathway was reduced by a factor of approximately 3 on average.
hmo00670 1-Carbon Folate Metabolism	This pathway covers one-carbon molecule metabolism utilizing folate. Reduction of dihydrofolate (DHF) yields tetrahydrofolate (THF). THF can serve as an acceptor/donor of 1-carbon units in three possible oxidation states: N^5-methyl-THF; N^5,N^{10}-methylene-THF, and N^{10}-formyl-THF (and others in the same oxidation state).	-4.2 ± 1.9 (11) P-value = 0.000829 Expression of genes in the carbon folate pathway was reduced by a factor of approximately 4 on average.
hmo00730 Thiamine Metabolism	Thiamine, also known as vitamin B1, is composed of pyrimidine and thiazole structures, which combine to form thiamine phosphate and are made via separate pathways. The thiamine phosphate product is phosphorylated to yield the active form of the cofactor, thiamine pyrophosphate. In this anabolic pathway, thiamine is made through several parallel pathways, utilizing glycine and tyrosine as starting points.	-6.7 ± 7.2 (10) P-value = 0.00170 The expression of genes involved in thiamine metabolism decreased by a factor of 6-7 on average.
hmo00740 Riboflavin Metabolism	Riboflavin, also known as vitamin B2, is an important element of cofactors flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD). These cofactors function as redox agents for a variety of processes (i.e. the electron transport chain) due to their ability to accept/donate a hydride plus a proton (i.e. 2 electrons and 2 protons) in 1-electron or 2-electron steps. Riboflavin is synthesized from the GTP product of metabolized purines and then converted to FMN and FAD by the enzymes riboflavin kinase and FMN adenylytransferase, respectively.	0.05 ± 1.94 (6) P-value = 0.59 The expression of genes involved in riboflavin metabolism was unaffected by the presence/absence of ammonium.

hmo00750 Vitamin B6 Metabolism	Vitamin B6 is a water-soluble vitamin with active form pyridoxal phosphate (PLP), an essential cofactor in amino acid metabolism. Glyceraldehyde 3-phosphate (G3P) and ribulose 5'-phosphate (Ru5P) produced from glycolysis and the pentose phosphate pathway are used to form the active pyridoxal 5'-phosphate.	-10.9 ± 9.3 (4) P-value = 0.0645 The expression of genes involved in vitamin B6 metabolism was not significantly affected by the absence of ammonium.
hmo00760 Nicotinate and Nicotinamide Metabolism	Nicotinate (vitamin B3) and nicotinamide are precursors to NAD^+ and $NADP^+$, which are used for redox reactions. Quinolinate is generated from aspartate, and is converted to nicotinate D-ribonucleotide and then to deamino-NAD^+. NAD^+ is then synthesized by NAD^+ synthase. Some of this is phosphorylated into $NADP^+$ by NAD^+ kinase.	-2.2 ± 2.4 (6) P-value = 0.26 The expression of genes involved in nicotinate and nicotinamide metabolism was not significantly affected by the absence of ammonium.
hmo00770 Pantothenate and CoA Biosynthesis	Pantothenate, also known as vitamin B5, is the precursor of coenzyme A, which functions to help transfer fatty acids from the cytoplasm to the mitochondria. Pyruvate is first converted to 2-acetolactate and then to 2,3-dihydroxy-3-methylbutanoate. The hydroxyl group on the beta-carbon is then removed and a hydroxyl methyl group is added to form 2-dehydropantoate, which is reduced and ligated with beta-alanine to create (R)-Pantothenate. This is then combined with cysteine to form (R)-4'-phosphopantothenoyl-L-cysteine which is decarboxylated and linked to an adenosine 3'-5'-monophosphate to make dephospho-CoA. Dephospho-CoA kinase then synthesizes CoA from this product.	-4.0 ± 5.1 (14) P-value = 0.0184 The expression of genes involved in pantothenate and CoA biosynthesis decreased by an average of four-fold.
hmo00780 Biotin Metabolism	Biotin, also known as vitamin H or B7, is a water-soluble B vitamin that is a coenzyme for carboxylases, which are involved in the synthesis of fatty acids and the amino acids isoleucine and valine. Biotin is synthesized from pimelate thioester through a two-stage biosynthetic pathway and is made from two precursors, alanine and pimeloyl-CoA via three enzymes. Through the final four steps of the biotin bicyclic ring assembly, pimeloyl-ACP is converted to biotin.	-1.3 ± 0.2 (4) P-value = 0.0719 The expression of genes involved in biotin metabolism was unaffected by the absence of ammonium.
hmo00790 Folate Synthesis	Folate, also known as vitamin B9, is necessary for 1-carbon metabolism (see hmo01200). Folate biosynthesis follows purine metabolism, which yields GTP. GTP is subjected to a series of reactions catalyzed by GTP cyclohydrolase I and is eventually released in the form of 7,8-dihydroneopterin 3'-triphosphate. Alkaline phosphatase then dephosphorylates this into dihydroneopterin. This product is eventually converted to 7,8-Dihydrofolate, which is then used to synthesize tetrahydrofolate for folate metabolism (see hmo00670).	-4.7 ± 2.3 (8) P-value = 0.0115 The expression of genes involved in folate synthesis decreased by an average of four-fold.
hmo00860 Porphyrin and Chl Synthesis	The purpose of this pathway is to synthesize heme, (bacterio) chlorophyll, and corrin. This anaerobic pathway originates with the production of 5-aminolevulinate from Glu-tRNA. 5-aminolevulinic acid dehydratase catalyzes the reaction of 5-aminolevulinate into uropophyrinogen III. At uropophyrinogen III, the pathway splits into two distinct paths; one leading to Vitamin B12 coenzyme by uroporphyrinogen III synthase/methyltransferase and the other to porphyrin and chlorophyll by coproporphyrinogen III oxidase. The conversion of threonine to Vitamin B12 is catalyzed by GHMP-kinase. Vitamin B12 is also synthesized from dimethylbenzimidazole by nicotinate-nucleotide--dimethylbenzimidazole phosphoribosyltransferase.	-8.6 ± 6.7 (37) P-value = 0.0000036 The expression of genes involved in porphyrin and chlorophyll synthesis decreased by an average of 8.5-fold. (Changes are deemed statistically significant after Bonferroni correction.)
hmo00900 Terpenoid Backbone Biosynthesis	The terpenoid backbone biosynthesis pathway results in the formation of natural products consisting of isoprene units. The melanovate and non-melanovate pathways are used to make terpenoid building blocks, which can be condensed to form sterols and carotenoids. *H. modesticaldum* seems to utilize only the non-melanovate pathway, and the products made also feed into menaquinone and other terpenoid-quinone biosynthesis pathways.	-8.4 ± 5.8 (9) P-value = 0.0584 The expression of genes involved in the terpenoid backbone biosynthesis pathway was not significantly affected by the absence of ammonium.
hmo00906 Carotenoid Biosynthesis	*H. modesticaldum* produces only one carotene, 4,4'-diaponeurosporene. The typical pathway for such a C30 carotene requires only two gene products: CrtM and CrtN. Only the latter has been annotated in the genome of this organism.	-11.4 (1) P-value = undefined Expression of the *crtN* gene was reduced by a factor of approximately 11 in the absence of ammonium.
hmo00910a Nitrogen Metabolism (Non-Nitrogen Fixation Genes)	The purpose of this pathway is to incorporate ammonia into amino acids. First, glutamine synthetase catalyzes the reaction between glutamate and ammonium to produce glutamine. Although the organism appears to have components of the enzyme nitrite reductase (which reduces nitrite to ammonia), *H. modesticaldum* has not been reported to use nitrite as a nitrogen source.	-4.7 ± 3.7 (5) P-value = 0.25 The expression of genes involved in nitrogen metabolism (but not nitrogen fixation) was not significantly affected by the absence of ammonium.
hmo00910b Nitrogen Metabolism (Nitrogen Fixation Genes)	*H. modesticaldum* possesses the ability to fix N_2 during both chemotrophic and phototrophic growth. Nitrogen fixation is a metabolic pathway that reduces nitrogen to ammonia. This energy-intensive process is carried out by the enzyme nitrogenase. 2 ATPs must be utilized in order for the Fe protein to pump an electron into dinitrogenase. A total of 8 electrons are used per dinitrogen reduced to 2 ammonia; H_2 is produced as a byproduct. The *nifE, nifN, nifX, fdxB, nifB, nifV, nifS* and HM1_1865 gene products are all involved in the process of nitrogenase assembly and maturation.	24.9 ± 19.4 (9) 22.8 ± 16.4 (14) P-value = 0.0281 (0.0291) Expression of genes involved in nitrogen fixation were found to be up-regulated by an average of ~25-fold in the absence of ammonium. If one assumes an RPKM of 1 in the ammonium-replete case for the 5 genes with no reads in that condition, then the increase in expression can be estimated to be ~23-fold (i.e. the estimation of up-regulation by the absence of ammonium is fairly robust).

hmo00920 Sulfur Metabolism	Sulfur metabolism takes place in *H. modesticaldum* via both oxidative and reductive pathways. It is present in nucleic acids, proteins, polysaccharides, phenols, etc. This organism does not seem to reduce sulfate, but rather reduces thiosulfate to sulfide, which is then used to make cysteine. Pathways that involve the incorporation of sulfur in this organism requires the input of energy.	1.4 ± 3.1 (10) P-value = 0.59 Expression of genes involved in sulfur metabolism was unaffected by the absence of ammonium. (A notable exception was one of the *cysD* genes, which was down-regulated about 9-fold.)
hmo00970 Aminoacyl-tRNA Biosynthesis	Aminoacyl-tRNAs deliver amino acids to the ribosome during translation for integration into the polypeptide chain that is being produced. During aminoacyl-tRNA biosynthesis, an aminoacyl tRNA synthetase catalyzes the esterification of the amino acid to its cognate tRNA (a.k.a. "charging" the tRNA).	-3.4 ± 2.6 (44) P-value = 0.0148 On average, the expression of genes involved in the synthesis of aminoacyl-tRNAs was reduced by a factor of 3.5 in the absence of ammonium.
hmo02010 ABC Transporters	ABC (ATP Binding Cassettes) Transporters couple ATP hydrolysis to active transport of a wide variety of molecules. *H. modesticaldum* has ABC transporters that move mineral and organic ions, monosaccharides, phosphate and amino acids, and metal cations/ iron siderophores/vitamins.	-2.6 ± 3.8 (44) P-value = 0.00444 Expression of genes involved in the ABC transporters pathway decreased by an average of 2.5-fold.
hmo02020 Two-Component System	The two-component signal transduction system allows the bacteria to sense and react to changes in its internal and external environment. Each two-component system consists of a sensor and regulator, allowing the cell to respond by inducing transcriptional changes. Typically, the sensor protein histidine kinase phosphorylates itself in response to environmental signals. *H. modesticaldum* contains at least seven such systems, presumably responding to diverse signals such as changes in temperature, pH, osmolarity, etc.	-3.5 ± 3.0 (28) P-value = 0.0448 Expression of genes involved in the two-component signal transduction system was reduced by a factor of ~3.5. A notable exception to this trend, *csrA*, displayed a 7-fold increase in expression.
hmo02030 Bacterial Chemotaxis	The chemotaxis pathway of *H. modesticaldum* allow the bacterium to detect chemical gradients through its receptors and through a signal cascade regulate its flagella. Direction and speed of rotation of the flagella can also be modulated.	-5.7 ± 2.9 (14) P-value = 0.000266 Expression of all genes involved in the chemotaxis pathway was reduced by an average of 5- to 6-fold. (Changes are deemed statistically significant after Bonferroni correction.)
hmo02040 Flagellar Assembly	This pathway is responsible for the assembly of proteins that constitute the flagella of the bacterium, granting it motility. It consists of a series of proteins responsible for the export across the membrane and assembly of the flagellum, as well as the structural components of the flagellum itself.	-6.8 ± 6.3 (19) P-value = 0.31 (0.0000161) The expression of genes involved in flagellar assembly was not significantly affected by the absence of ammonium if all genes are included. If one excludes the 6 genes that increase in expression, however, the overall drop in expression was ~7-fold (significant after Bonferroni correction).
hmo02060 Phosphotransferase System (PTS)	The phosphotransferase system (PTS) is used by bacteria for the uptake of carbohydrates, especially hexoses and disaccharides. PTS also catalyzes their conversion into phosphoesters to facilitate transportation. PTS separated into two components: enzyme I along with histidine phosphocarrier protein, and membrane-bound sugar specific permeases. However, in *H. modesticaldum*, only the PEP-protein phosphotransferase seems to be present.	-6.8 (1) P-value = undefined The *ptsI* gene exhibited a 7-fold decrease in expression. (It is unclear if the PTS system actually exists in this organism.)
hmo03010 Ribosome	The ribosome is the site at which translation of mRNA into an amino acid polypeptide chain takes place. For *H. modesticaldum*, ribosomal RNAs that are involved in the construction of the ribosome include 23S/5S and 16S rRNAs to make the large 50S and small 30S subunit, respectively.	-1.0 ± 3.0 (78) P-value = 0.0654 The expression of genes involved in the construction of ribosomes was unaffected by the absence of ammonium.
hmo03018 RNA Degradation	This pathway involves the enzymes used in the process of RNA degradation. This is a catalytic process that is used to routinely degrade messenger RNA into its component nucleotides.	-1.1 ± 2.9 (9) P-value = 0.0511 Expression of genes for enzymes and proteins involved in the RNA degradation pathway is unaffected by the absence of ammonium.
hmo03020 RNA Polymerase	RNA Polymerase is responsible for synthesis of all RNAs in the bacterial cell, using the genomic DNA as a template. Bacterial RNA Polymerase is made up of five subunits: β', β, two α subunits, and the ω subunit. β' is the largest subunit and encompasses some of the active site that performs RNA synthesis. The rest of the active site is located in the β subunit. The two α subunits contain determinants for interactions with DNA promoters and regulatory factors. The ω subunit is the smallest and stabilizes the formation of the protein.	-1.3 ± 2.9 (5) P-value = 0.78 The expression of genes involved in RNA polymerase was unaffected by the absence of ammonium.
hmo03030 DNA Replication	DNA Replication is the process by which the chromosomal DNA is copied. Proteins involved includes the subunits of the DNA polymerase III holoenzyme, and a variety of other enzymes including helicase and primase. (Heliobacteria appear to missing subunits Θ, ψ, and χ in the DNA polymerase III holoenzyme, but these subunits are typically not found in the Firmicutes.)	-6.7 ± 7.6 (12) P-value = 0.407 The expression of genes involved in the DNA replication pathway is reduced by a factor of ~7 on average.

hmo03060 Protein export	*H. modesticaldum* possesses both the Sec-SRP pathway and the Twin arginine targeting (Tat) pathway. In the former, translating or newly translated proteins with an N-terminal signal peptide (or internal integral transmembrane helix) are bound by SRP and presented to the SecYEG translocase. Working with the SecA ATPase, the unfolded polypeptide is translocated from the cytosol to the exterior, where it will fold. The Tat pathway recognizes proteins with a characteristic N-terminal signal containing 2 consecutive Arg residues and is able to transport folded proteins across the membrane.	-3.5 ± 7.6 (13) P-value = 0.0944 The expression of genes involved in the bacterial secretion system was not significantly affected by the absence of ammonium.
hmo03070 Bacterial Secretion System	Heliobacteria are Gram-positive bacteria and thus lack an outer membrane. Unsurprisingly, the type I-V secretion systems are not required. Secreted proteins are transported across the membrane by both the Sec-SRP pathway and the Twin arginine targeting (Tat) pathway. Thus, all complete systems overlap entirely with the hmo03060 pathway.	(Since all genes in this table are also on hmo03060, we do not consider them separately here.)
hmo03410 Base excision repair	The process of base excision repair involves the recognition and removal of damaged or mismatched bases from DNA. After excision of the damaged base by a DNA glycosylase, the AP endonuclease cleaves the phosphodiester backbone just 5' of the apurinic/apyrimidinic site, allowing DNA pol I to replace the damaged strand with a newly synthesized DNA strand.	-4.7 ± 5.1 (6) P-value = 0.0217 The expression of genes involved in the base excision repair pathway is reduced by a factor of ~5 on average
hmo03420 Nucleotide excision repair	Nucleotide excision repair is performed by the UvrABC endonuclease. A dimer of the UvrA ATPas/GTPase binds to UvrB. The UvrA dimer functions to detect damaged sites on the DNA double helix, and UvrB attaches to the damaged site. The UvrA dimer then disassociates and UvrC attaches to UvrB, resulting in cleavage of DNA upstream and downstream of the lesion. The damaged segment is removed by PcrA/UvrD DNA helicase. UvrB prevents the reannealing of the excised segment until displaced by DNA polymerase I, which fills in the correct nucleotide sequence.	-4.5 ± 4.6 (7) P-value = 0.0214 The expression of genes involved in the nucleotide excision repair pathway is reduced by a factor of about 4.5 on average.
hmo03430 Mismatch repair	DNA mismatch repair involves the recognition and excision mismatched base pairs, and the re-synthesis of the excised DNA. Note that enzyme MutH appears to be missing from this pathway, as is generally true for all Firmicutes, which use a MutH-independent pathway.	-3.6 ± 3.0 (9) P-value = 0.0206 Expression of these two genes is relatively low even in the presence ammonium, but seems to decrease about 3.6-fold in the absence of ammonium.
hm03440 Homologous recombination	The RecFOR pathway, also called the RecF pathway, repairs DNA in bacteria via homologous recombination. The pathway typically repairs single-strand gaps, but is also capable of repairing double-strand breaks when the RecBCD pathway is inactivated or absent, as may be the case in *H. modesticaldum*. The pathway produces a Holliday junction intermediate and concludes with the branch migration and resolution of the Holliday junction.	-5.8 ± 3.9 (19) P-value = 0.000453 The expression of genes involved in the homologous recombination pathway is reduced by a factor of about 6 on average. (Changes are deemed statistically significant after Bonferroni correction.)
hmo03450 Non-homologous end joining	Non-homologous end-joining is a form of double strand break repair. It is initiated when the Ku protein recognizes and binds to the broken ends of double stranded DNA. Ku then forms a bridge between the two ends, aligning them as well as preventing their degradation. DNA LigD is then recruited, which has polymerase activity that fills the gap. Finally, the ligase activity of LigD ligates the two ends together, thus sealing the break.	P-value = 0.61 The expression of genes involved in non-homologous end joining was extremely low under both conditions; no significant differences were detected.
hmo04122 Sulfur Relay System	The sulfur relay system is used to pass sulfur atoms from ubiquitin-like proteins to Molybdenum cofactors (Moco) and thiamine via separate pathways. The sulfur-containing cofactors Moco and thiamine exist in this organism and require sulfur carrier proteins to be synthesized.	-2.50 ± 3.42 (6) P-value = 0.83 The expression of genes involved in sulfur transport was not significantly affected by the absence of ammonium.

Table 2: For each pathway in the KEGG for *H. modesticaldum*, a table was created (present in the Supplemental Information named after the KEGG pathway; e.g. Table hmo00010 is for pathway hmo00010, which is glycolysis). Each row in Table 2 summarizes one of these tables, providing the KEGG designation, name and brief description of the pathway, as well as a summary of the RNA-seq data.
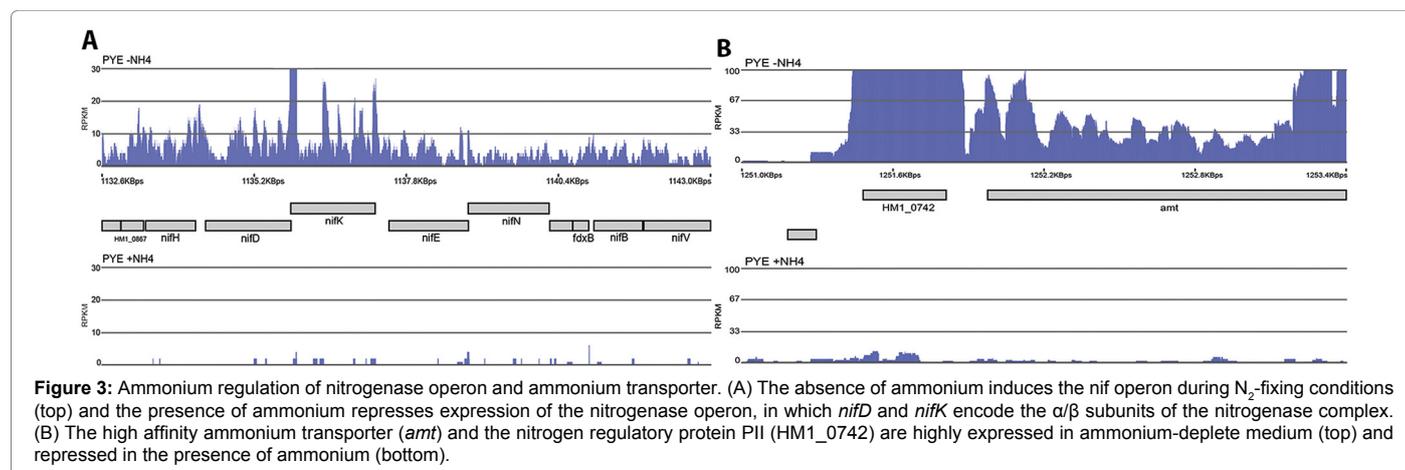
[1]Fold change is defined as in Table 1 (e.g. -20 is a 20-fold decrease, -1.25 is a 20% decrease, +5 is a 5-fold increase, etc.)

[2]The reported P-value is the result of a paired Student's t-test (2-tailed) comparing the RPKM values of genes in the KEGG pathway in the plus and minus ammonium conditions.

[3]A summary of the overall changes in expression of genes involved in each KEGG pathway (with some exceptions noted) is provided. Expression changes in those pathways for which the P-value < 0.05 were not deemed significant. The six pathways for which the P-value was less than 0.000806 (i.e. 0.05/62), reflecting the Bonferroni correction, were also noted.

other members of the family [6,29]. The RNA-sequencing data for all genes involved in nitrogen metabolism can be found in Data Set S3. However, due to the very different regulation seen for different sets of genes, we pulled out the genes involved in nitrogen fixation and put them into a separate table (pathway hmo00910b), leaving the rest of the genes involved in basic nitrogen metabolism in the first table (pathway hmo00910a).

When the bacterium was placed in N$_2$-fixing conditions, expression of genes for nitrogen fixation was up-regulated (Figure 3A and Data Set S3, pathway hmo00910b). In particular, the *nifD* and *nifK* transcripts, encoding the core subunits of nitrogenase, increased in abundance by a factor of 25-50. A similar increase was seen for *nifH*, which encodes the Fe protein that provides electrons to the nitrogenase enzyme, but it was difficult to estimate, as no reads were detected aligning to this gene

Figure 3: Ammonium regulation of nitrogenase operon and ammonium transporter. (A) The absence of ammonium induces the nif operon during N₂-fixing conditions (top) and the presence of ammonium represses expression of the nitrogenase operon, in which *nifD* and *nifK* encode the α/β subunits of the nitrogenase complex. (B) The high affinity ammonium transporter (*amt*) and the nitrogen regulatory protein PII (HM1_0742) are highly expressed in ammonium-deplete medium (top) and repressed in the presence of ammonium (bottom).

in the baseline transcriptome. In cases such as this, a value of 1 RPKM in was used for the baseline value in order to estimate a minimum expression difference. If one excludes genes without reads in the baseline transcriptome, the average up-regulation of nitrogen fixation genes was ~25-fold (n=9). If one includes such genes (for *nifH, nifN, fdxB,* HM1_0868 and HM1_0867), using an assumed baseline value of 1 RPKM, then an average increase of 23-fold (n=14) was estimated, indicating that this is a fairly robust value. This up-regulation in structural genes, maturation factors, and regulatory protein gene transcription was expected, since all of these components are needed for the reduction of molecular N₂ and are in general not required for other functions.

The exception to this may be *nifS* and HM1_1865, which encode a cysteine desulfurase and FeS cluster assembly protein, respectively, and are not found in the *nif* operon. Thus, the corresponding gene products likely play a role in functions required during normal growth (e.g. FeS cluster assembly) as well as assembly of the FeMoCo cofactor in dinitrogenase. Expression of these two genes is quite high in the baseline transcriptome, and only increases about 3-fold under N₂-fixing conditions (Data Set S3, pathway hmo00910b). The other exception is *nifX*, which is found in the *nif* operon, and is only up-regulated by ~3-fold as well. Its expression is quite low in both transcriptomes (3.4 and 10 RPKM), so this change may not be as reliable as for the other members of the operon.

The *nif* gene cluster had been previously cloned and sequenced from *Heliobacterium chlorum*, and the gene order is completely conserved in *H. modesticaldum*. In both species, there are three genes at the 5' end of the *nif* operon that appear to be involved in regulation of N₂ fixation. Immediately upstream of *nifH* are the *nifI₁* and *nifI₂* genes (HM1_0868 and HM1_0867, respectively), which are members of the GlnB family of PII proteins. These PII-like proteins are typically regulated by glutamine (as a signal for nitrogen), 2-oxoglutarate (as a signal for carbon), and ATP (as a signal for energy) in a wide variety of Bacteria and Achaea [30-32]. The *nifI₁/nifI₂* gene pair diverges from this family and is found in the *nif* operons of several anaerobic diazotrophs, including *Clostridium acetobutylicum* and *Methanococcus maripaldus* [33]. Just upstream of this gene pair is a conserved protein that appears to be unique to heliobacteria, called "*orf1*" [34]; it was proposed to alter nitrogenase expression by controlling termination of transcription after the *orf1* sequence, thus attenuating expression of the downstream *nif* genes [33]. Under N₂-fixing conditions, expression of all 3 genes increases (Figure 3A and Data Set S3, pathway hmo00910b), as was seen previously in *Heliobacterium chlorum* [34].

Growth in ammonium-deplete medium also resulted in a large up-regulation of the *amt* gene (~59-fold; Figure 3B), a gene similar to the bacterial *amtB* gene encoding a high-affinity ammonium transporter [35]. In fact, this represents the largest increase in expression that can be reliably measured (i.e. where the expression in the baseline transcriptome was significant). It is likely that the cells in low-ammonium conditions up-regulate *amt* in an attempt to scavenge any available NH₄⁺, which would represent a lower energetic cost than fixing N₂. Conversely, NH₄⁺ is relatively abundant in the baseline condition, which implies that there is adequate NH₄⁺ influx via low-affinity transporters present during mid-log phase growth. Therefore, *amt* expression is likely down-regulated through a feedback mechanism in that situation. The *AmtB* protein has also been implicated in regulation of nitrogen metabolism in general and of nitrogenase in particular in multiple bacterial species [36,37]. It is interesting that only 13 genes display an increase in expression of 15-fold or higher in the absence of ammonium, and all of them are either in the *nif* operon or in a three-gene cluster containing the *amt* gene. In the latter, *amt* is preceded by HM1_0742, a PII homolog, which exhibits a 28-fold increase in expression. It is followed by HM1_0740, which displays similarity to the sporulation-specific YabG peptidase of *Bacillus subtilis* and whose expression increases 48-fold; it is unclear at this time what the role of this gene product would be in heliobacteria. All other genes in the genome exhibit fold changes of 11 or less in the absence of ammonium.

In contrast to the genes involved in N₂ fixation, the rest of the genes involved in basic nitrogen metabolism were not significantly affected by the absence of ammonium (Data Set S3, pathway hmo00910a). Once NH₄⁺ is either synthesized or transported into the cell, it is incorporated into glutamine via glutamine synthetase (*glnA*). Glutamate is synthesized by amination of 2-OG using the amide N of glutamine and NADPH as a reductant, catalyzed by the enzyme glutamate synthase. Glutamate is used as a nitrogen donor for synthesis of many amino acids and other biomolecules. As this organism appears to lack a glutamate dehydrogenase, glutamine synthetase is thus the only entry point for ammonium into biomolecules, and glutamate synthase is the only way to make glutamate. The *glnA* gene is highly expressed under both conditions (306 *vs.* 102 RPKM), being reduced by only 3-fold upon transfer to ammonium-deplete conditions. (This level of decrease of mRNA abundance is typical of most genes, as discussed below.) This is expected, given that glutamine synthetase will be the major entry point for ammonium under both conditions. The genes encoding glutamate synthase (*gltB* and *gltD*) were expressed at a lower level. Most of the genes for incorporation of ammonia were down-regulated

in ammonium-deplete media by a factor of 4-5 on average. As will be seen, this behavior is typical for most metabolic pathways. Thus, expression of genes for fundamental nitrogen metabolism behaved much as expected upon a shift from ammonium-replete to N$_2$-fixing conditions.

Bioenergetics and electron transfer

Photosynthesis is an efficient mechanism for ATP production that does not compete for the carbon of pyruvate, much of which is wasted as CO$_2$ and small molecules (*e.g.* acetate, ethanol) during fermentation [7]. The heliobacterial reaction center (HbRC) uses absorbed photons to oxidize periplasmic cyt c_{553} and reduce cytoplasmic ferredoxin (Fd). The electrons from Fd^{red} are sent back to cyt c_{553} via complex I and the cyt b_6c complex in a cyclic electron transport (cET) process that produces a proton gradient that can be used by ATP synthase to phosphorylate ATP [38-40]. Because the KEGG does not have the heliobacterial cET pathway, we made use of the oxidative phosphorylation pathway (hmo00190) as the baseline pathway, as it contains complexes I (NADH dehydrogenase), II (succinate dehydrogenase), III (cyt b_6c), and V (ATP synthase), despite the fact that fact that heliobacteria lack complex IV (cytochrome oxidase), being strict anaerobes. Other genes important in bioenergetics not represented in the KEGG pathways were added to separate tabs; hmo00190b contains the genes encoding additional proteins involved in the cET pathway (e.g. the HbRC, its electron donors and acceptors), while hmo00190c lists the genes for hydrogenases.

Given the importance of the HbRC to cellular metabolism, we expected the *pshA* gene encoding the HbRC core polypeptide to be highly expressed during photoheterotrophic growth. In fact, *pshA* is the twelfth highest protein-encoding gene expressed (1446 RPKM; Data Set S4, pathway hmo00190b, Figure 4B). During the elucidation of the HbRC structure to 2.2 Å, we discovered a previously unknown subunit called PshX [41]. This small polypeptide consists of a single transmembrane alpha-helix with a few residues on either side. The resolution of the structure was sufficient to allow identification of sidechains and subsequent identification of the gene, which was confirmed by mass spectrometry. The gene encoding this subunit (HM1_0821), which had not been previously annotated, is not in the photosynthetic gene cluster nor does it appear to be in an operon. However, its normalized expression level (1685 RPKM; Data Set S4, pathway hmo00190b, Figure 4C) is almost exactly the same as *pshA*. This makes sense as the stoichiometry of the subunits is 2:2 in the HbRC.

The electron donor to the HbRC is a membrane-anchored cyt c_{553} encoded by the *cyhA* gene; it is highly expressed (490 RPKM; Data Set S4, pathway hmo00190b). A high expression level was also found for genes *pshB2* and *pshB1* (Figure S4), which encode the [4Fe-4S] dicluster ferredoxins, now thought to be the primary electron acceptors of the HbRC [42]. The *pshB1* and *pshB2* genes are adjacent in the genome and are likely co-transcribed, but for unknown reasons the second mRNA (*pshB1*) is about five-fold more abundant than the first one (*pshB2*), and is in fact the fifth most highly expressed protein-encoding gene (2507 RPKM; Data Set S4, Figure 4B and Figure S4). The ferredoxins are expected to interact with Fd-NADP$^+$ oxidoreductase (FNR) to reduce NAD(P)$^+$ to NAD(P)H and generate reducing power for metabolic reactions as well as for cyclic electron flow. While the FNR enzyme was not originally annotated in the genome, the enzyme and its gene (HM1_0289) were subsequently identified and found to be

well expressed [17]; it has, however, nowhere near as highly expressed as the other genes in this pathway (29 RPKM; Data Set S4, pathway hmo00190b).

The genes involved directly in photosynthetic electron transfer are strongly regulated by ammonium abundance. In the absence of ammonium, there is a 44-fold decrease in the *pshA* transcript, a 78-fold decrease in the *cyhA* transcript, a 57-fold decrease in the *pshB1* transcript, and a ~140-fold decrease in the *pshB2* transcript (Figures 4B and S4, Data Set S4, pathway hmo0019b). This result is surprising given the distinct advantage provided by the HbRC for growth, and one would not expect the abundance of the HbRC, or its electron donors and acceptors, to be drastically lowered in N$_2$-fixing conditions. In contrast, the mRNA levels of the other proteins in the proposed cyclic electron transfer pathway (complex I and III) as well as ATP synthase (complex V) do not follow the same pattern (Data Set S4, pathway hmo00190). They are not as highly expressed, but they are also not nearly as strongly repressed by the absence of ammonium (average repression of 2- to 3-fold). The same is true of the genes encoding the cytochrome *bd* menaquinol oxidase (*cydA/B*), whose primary role is likely to scavenge O$_2$ rather than contribute to energy production.

Such a large drop in transcript abundance after a shift to ammonium-deplete medium is not a common feature of the other abundant mRNAs. For example, *rpoE* (σ^{24} factor) only decreases by 1.6-fold and *cstA* (carbon starvation protein) increases by 1.4-fold. In addition, the four most highly expressed genes in ammonium-replete conditions (HM1_3146, HM1_3147, HM1_3149, HM1_3151; 6200-19000 RPKM), all of which are conserved hypothetical proteins of unknown function located next to each other in the genome, exhibit very little repression in the absence of ammonium (at most 2-fold; Data Set S2). In order to make the comparison more meaningful, we compiled a list of the most highly expressed protein-encoding genes in ammonium-replete conditions (i.e., with normalized transcript levels >300 RPKM). When the log$_2$ of the expression change (-NH$_4^+$ compared to +NH$_4^+$) is plotted against the log$_2$ of the expression level in +NH$_4^+$ medium of these 42 genes, something striking emerges: the structural genes for the HbRC (*pshA*, *pshX*), its electron donor (*cyhA*) and acceptor (*pshB1*, *pshB2*) form a discrete group of genes that are highly repressed by the lack of ammonium (red dots in Figure 4B).

To test the effects of ammonium depletion on expression of the HbRC, we grew cells in PYE or PYE-NH$_4^+$ media, took samples at a few points along the growth curve, and directly assayed functional HbRC levels by a quantitative P$_{800}$ photobleaching experiment. Briefly, cells were excited with a saturating laser flash and the absorbance at 800 nm was monitored (Figure 4C). The primary electron donor of the HbRC (P$_{800}$) absorbs maximally at 800 nm and is bleached upon light-driven electron transfer to the ferredoxin acceptors. The bleaching recovers within a few milliseconds due to reduction of oxidized P$_{800}$ by the membrane-attached cyt c_{553} donor [43,44]. We observed only a modest drop (~30%) in cellular P$_{800}$ levels in PYE-NH$_4^+$ cultures relative to the PYE cultures (Table S4), and in both cases there was a small drop (~20%) as the cultures aged. Moreover, the kinetics of P$_{800}$ recovery were about the same in both (t$_{1/2}$ ≈ 2.5 ms; Figure 4C), which is inconsistent with a large drop in cyt c_{553}; the reduction of P$_{800}^+$ by cyt c_{553} is a second-order reaction and is thus sensitive to the concentration of the electron donor [44]. Thus, despite the large drop in *pshA* and *cyhA* mRNA levels, we find no evidence for correspondingly large drops in the levels of the encoded proteins. Thus, the transcript levels of the

structural genes for the RC and its donors and acceptors (*pshA*, *pshX*, *cyhA*, and *pshB1/2*) may be much higher in the presence of ammonium than actually necessary, although it is unclear why this would be the case.

Hydrogenase enzymes can catalyze the two-electron reduction of protons to molecular hydrogenase (H₂) and/or the two-electron oxidation of H₂ [45]. The genome of *H. modesticaldum* contains two copies of the genes *(hupSLC)* for a membrane-bound [NiFe] uptake hydrogenase; one of these operons also contains the genes encoding the maturation factors (*hupD hypABCDEF*) [11]. This type of hydrogenase is often used by N₂-fixing organisms to recover the H₂ produced by nitrogenase [46]. However, we found that expression of the *hup* and *hyp* genes was not induced by a shift to ammonium-deplete conditions, but was instead slightly repressed by a factor of 7 ± 5 (Data Set S4, pathway hmo00190c). The genome also appears to encode two copies of a [FeFe] hydrogenase, which are often used by Clostridia to get rid of excess electrons during fermentation. The predicted products of HM1_1617 and HM1_2613 resemble each other (59% identical residues) and the catalytic subunits of clostridial [FeFe] hydrogenases, but genes for maturation factors were not found. It is thus unclear if this organism assembles an active [FeFe] hydrogenase. In any case, these genes are also not induced by the shift to N₂-fixing conditions.

Regulation of fundamental carbon metabolism by ammonium

As mentioned, *H. modesticaldum* cannot grow as a photoautotroph, as it lacks at least one key enzyme for each carbon fixation pathway. For example, heliobacteria seem to possess every enzyme for the reverse TCA (rTCA) cycle except for ATP-citrate lyase, which prevents them from using this cycle to fix CO_2 in an autotrophic manner [11,17]. Thus, *H. modesticaldum* is an obligate heterotroph; it uses pyruvate, lactate, or acetate as carbon sources [16]. Figure 2 summarizes central carbon metabolism and connections to nitrogen metabolism in heliobacteria.

Pyruvate was the sole carbon source during the growth of heliobacteria in these experiments. Expression of genes encoding enzymes in pyruvate metabolism was reduced by approximately 6-fold in the absence of ammonia (Data Set S5, pathway hmo00620). Pyruvate can be oxidatively decarboxylated to produce acetyl-CoA with concomitant reduction of 2 ferredoxins by the pyruvate:ferredoxin oxidoreductase (PFOR), as this organism lacks an NAD(P)-dependent pyruvate dehydrogenase. During fermentative growth, acetyl-CoA synthetase converts acetyl-CoA to acetyl-phosphate, which can be used as a phosphate donor to ADP to make ATP via acetate kinase. These latter 2 enzymes are moderately expressed (7-16 RPKM) under both conditions, while PFOR is highly expressed (30-168 RPKM). This may be because the cultures used were grown in the light and could use photosynthetic cET to produce ATP, allowing pyruvate to be used primarily as a carbon source.

This organism appears to lack pyruvate carboxylase to convert pyruvate to oxaloacetate (OAA). However, it can convert pyruvate to phosphoenolpyruvate (PEP) by pyruvate phosphate dikinase (PPDK). Incorporation of carbon into biomolecules can then be achieved through PEP carboxykinase (PEPCK) [17], which carboxylates PEP to OAA. OAA can then be converted to malate, fumarate, and succinate by the "left arm" of the TCA cycle; these enzymes are highly expressed
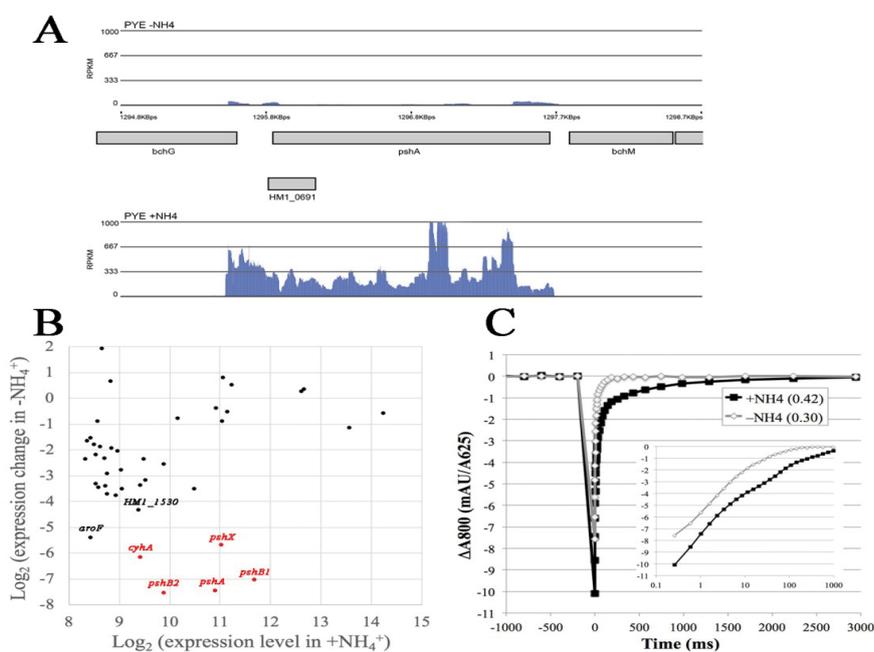

Figure 4: Ammonium regulation of the photosynthetic reaction center. (A) The absence of ammonium repressed the *pshA* gene encoding the HbRC core polypeptide. (B) Log-log plot of expression change vs. expression for the 42 most highly expressed protein-coding genes. This list includes all protein-encoding genes with normalized transcript levels of >300 RPKM (in the presence of NH₄⁺) that do not overlap with rRNA or tRNA genes (potentially raising their transcript counts). The x-axis is the log₂ of the normalized transcript level (in RPKM). The y-axis is the log₂ of the ratio of the normalized transcript levels in ammonium-deplete to ammonium-replete conditions. The names of the six genes with repression levels >16-fold are marked. The five red dots correspond to structural genes for the HbRC (*pshA*, *pshX*), ferredoxins (*pshB1*, *pshB2*), and cyt c_{553} (*cyhA*). (C) Photobleaching of P₈₀₀ induced by a 6-ns laser flash (at time 0) was monitored in cells using 10-μs pulses of light from an 803-nm LED. The ammonium-replete and ammonium-deplete cultures had A625s of 0.42 and 0.30, respectively, at the time of the measurement. Absorption changes at 803 nm are reported as mAU normalized to the A625 values of the culture to provide an estimate of cellular HbRC content. The inset is the post-flash data plotted on a log₁₀ time scale.

in *H. modesticaldum* (Data Set S5, pathway hmo00020). The transcript level of *pckA* encoding PEPCK was reduced by ~16-fold during N$_2$-fixing conditions, which is the largest fold change for any gene in the TCA cycle, which exhibited an average ~3.5-fold down-regulation in transcript level. Malate dehydrogenase, which would reduce OAA to malate in the reverse direction, exhibited an ~11-fold reduction in transcript level, the second highest change. It may be that the organism slows the flow of carbon into the left arm of the TCA cycle when ammonium is low.

We note that there are two genes annotated to be similar to the alpha subunit of OAA decarboxylase (oadA), which are moderately expressed under both conditions (30-120 RPKM). The OAA decarboxylase complex usually consists of 3 subunits – the beta and gamma subunits are integral membrane proteins required for coupling Na+ efflux to decarboxylation of OAA by the alpha subunit [39]. These latter two subunits are not annotated in the genome, so the existence of this enzyme is unclear at this time. We note, however, that the first oadA gene exhibits similarity to pyruvate carboxylase, and is preceded by HM1_0379, which is annotated as a biotin carboxyl carrier protein; both genes are moderately expressed (80-120 RPKM in +NH4+) and exhibit little repression in the absence of ammonium. There is a similar story with the second oadA gene, which is followed by a gene similar to the biotin carboxylase subunit of acetyl-CoA carboxylase (accC). Thus, it is possible that at least one of these putative oadA genes is part of a functional pyruvate carboxylase, which could allow synthesis of OAA in low ammonium.

The key enzyme of the TCA cycle – citrate synthase – appears to be missing in the *H. modesticaldum* genome. There is no gene encoding the typical (*Si*)-citrate synthase, but it has been suggested that this organism uses an (*Re*)-citrate synthase, an enzyme evolutionarily related to homocitrate synthases [7]. There are two genes in the *H. modesticaldum* genome that show high similarity (>40% identity) to the putative (*Re*)-citrate synthase of *Clostridium kluyveri* [46]. The most similar gene (50% identity) is *aksA* (HM1_2993), which is well expressed under both conditions (21-28 RPKM). The other is *nifV* (42% identity) and is part of the *nif* operon. Expression of this gene was increased by a factor of ~33-fold in the absence of ammonium. We thus propose that this gene is a true homocitrate synthase, and this increase in expression is expected, as homocitrate is an essential part of the FeMoCo cofactor in dinitrogenase, serving as one of the ligands to the Mo metal [47,48]. Heliobacteria do not have the lysine biosynthetic pathway via homocitrate – they use the pathway via aspartate 4-semialdehyde (Data Set S5, pathway hmo00300) – and there is no other known use for homocitrate. Thus, it is likely that the *aksA* gene product is actually a (*Re*)-citrate synthase, as proposed by Tang et al. [7]. If so, then OAA could be converted to citrate by this enzyme, and thence to 2-oxoglutarate via isocitrate by the "right arm" of the TCA cycle (Figure 2). This would allow the two arms of the TCA cycle to work in the opposite directions and OAA would be a major branch point compound.

PEP would also be a major branch point, as it can be converted to OAA by PEPCK to feed into the TCA cycle, or converted to carbohydrate via the enzymes of gluconeogenesis. As *H. modesticaldum* does not utilize carbohydrates as a carbon or energy source, it is most likely that the enzymes of glycolysis are all used in the anabolic direction (i.e. gluconeogenesis) for making fructose-6-phosphate, which then can be converted into whichever hexoses are needed. The genes involved in the gluconeogenesis pathway were all down-regulated in the absence of ammonium by an average of ~8-fold (Data Set S5, pathway

hmo00010). The fructose-6-phosphate generated by this pathway can also be converted to pentose-phosphates via the pentose-phosphate pathway. It should be noted that the oxidative enzymes of this pathway (glucose-6-phosphate dehydrogenase and 6-phosphogluconate dehydrogenase) are absent in this organism, so it would have to use primarily transketolase and transaldolase to convert the hexose-phosphates and triose-phosphates from gluconeogenesis into pentose-phosphates. The genes involved in the pentose-phosphate metabolism were all down-regulated in the absence of ammonia by a factor of ~7.5 (Data Set S5, pathway hmo00030). Gluconeogenesis products also feed into the fructose/mannose and galactose metabolic pathways, allowing production of any needed hexoses. The genes encoding enzymes in these pathways were not significantly affected by the absence of NH$_4^+$ (Data Set S5, pathways hmo00051 and hmo00052).

There has so far been no report of carbon storage polymers in *H. modesticaldum*, such as polyhydroxybutyrate or glycogen/starch. While the genome lacks enzymes for synthesis of the former, there are genes encoding a putative glycogen synthase (*glgA*) and glycogen branching enzyme (HM1_2990) in the genome. There are two genes for glycogen phosphorylase (*glgP*); one is highly expressed (108 -135 RPKM) and the other is barely expressed (0.3 – 0.7 RPKM). There is also a debranching enzyme (*malQ*). Thus, the minimum complement of enzymes is present for glycogen/starch synthesis and breakdown; they are down-regulated only ~3-fold in the absence of ammonium (Data Set S5, pathway hmo00500). If the rate of carbon incorporation exceeds that of ammonia synthesis during N$_2$-fixing conditions, it is possible that the cells store some of the excess carbon as glycogen.

Expression of enzymes in fatty acid biosynthesis was unaffected by the absence of ammonia (Data Set S5, pathway hmo00061). *H. modesticaldum* does not appear to use fatty acids as a carbon or energy source [5,9]. The only genes that showed significant expression in this pathway were HM1_0073 and *aas*, which are involved in other pathways, such as fatty acid biosynthesis. The metabolic pathways of lipopolysaccharide and glycerolipid synthesis were not significantly affected by the absence of ammonium, while glycerophospholipid synthesis genes collectively showed a ~6-fold decrease in gene expression (Data Set S5, pathways hmo00540, hmo00561 and hmo00564). A deviation from this trend was seen with *glpK*, which was up-regulated by a factor of ~3.5-fold. This gene encodes glycerol kinase, which can be involved in other pathways.

Regulation of intermediary metabolism and cofactor synthesis

The transcripts of genes involved in the synthesis of alanine, aspartate, glutamate, lysine, arginine, proline, glycine, serine, threonine, cysteine, methionine, and D-alanine, as well as the sulfur relay system, were all reduced by 2- to 3-fold on average in the absence of ammonium (Data Set S6). The genes involved in the nitrogen metabolism pathway—which utilizes the products formed during nitrogen fixation in order to synthesize glutamine, arginine and proline—displayed a 5-fold decrease in expression. The same was seen for genes involved in selenocompound metabolism. However, in both the cysteine/methionine and selenocompound metabolism pathways, there was an exception to both trends. The gene *metC* (cystathionine β-lyase) deviated significantly with an increase in expression of 4.5-fold (Data Set S6, pathway hmo00270). Expression of genes in basal sulfur metabolism was unaffected by the absence of ammonium, except for one of the *cysD* genes, which was repressed ~9-fold (Data Set S6, pathway hmo00920). Metabolism of the branched aliphatic amino acids (valine, leucine and isoleucine) demonstrated a 7-fold reduction in gene expression, while biosynthesis of the aromatic amino

acids (histidine, phenylalanine, tyrosine and tryptophan) displayed a higher level of repression in the absence of ammonium (~12-fold). An interesting exception is the two *aroF* genes, which encode the enzyme catalyzing the first committed step of aromatic amino acid biosynthesis (3-deoxy-7-phosphoheptulonate synthase); the more highly expressed one is strongly repressed (~60-fold; Data Set S6, pathway hmo00400; see Figure 4B), while the less highly expressed one is unaffected by the absence of ammonium. Thus, the repression of genes involved in biosynthesis of most amino acids was not very different from the average, with the exception of the aromatic and (to a lesser extent) the branched aliphatic amino acids.

The purine and pyrimidine pathways of *H. modesticaldum* synthesize the ribo- and deoxyribonucleotides that are polymerized to make DNA and RNA, while the amino/sugar nucleotide pathway involves the incorporation of sugars into nucleotides. Peptidoglycan, which forms the cell wall that encases the cell membrane and provides structural support and protection, is synthesized as amino-sugar (rather than nucleotide-sugar) polymers. In the absence of ammonium, the overall expression of genes in the nucleotide and peptidoglycan synthetic pathways decreased by an average of 4-fold. All individual genes were down-regulated by a factor of less than 17-fold or up-regulated by a factor less than 3-fold. Overall, nucleotide metabolism and peptidoglycan synthesis do not seem to be greatly affected by the absence of ammonia, as one might expect (Data Set S6, pathways hmo00230 and hmo00240).

There are several pathways for the biosynthesis of molecules used as cofactors or prosthetic groups. The flux through these pathways is thus smaller than the pathways for intermediary metabolism. For most of these pathways, changes in transcript levels of the enzymes was either not significant or not very large (~6-fold, on average; Data Set S6, pathways hmo00670 to hmo00790). Synthesis of porphyrin is initiated by the conversion of glutamyl-tRNA to 5-aminolevulinate; it then diverges into two pathways: one leads to the synthesis of Vitamin B12, while the other path leads to synthesis of heme and chlorophyll. The porphyrin and chlorophyll biosynthesis pathway was one of the few pathways for which the change in transcript abundance (~9-fold) was statistically significant after the Bonferroni correction (Table 2, Data Set S6, pathway hmo00860). The *thiD* gene, which encodes phosphomethylpyrimidine kinase involved in thiamine and Vitamin B6 synthesis, was an outlier, being repressed ~23-fold in the absence of ammonium (Data Set S6, pathways hmo00730 and hmo00750). Other outliers were *bchG* (encoding chlorophyll synthase, 28-fold), as well as HM1_2396 and *cbiB* (encoding cobalamin biosynthesis protein, CbiB/CobD; 20-fold and 24-fold, respectively).

DNA, RNA, and protein metabolism

The genes involved in replication and repair of genomic DNA, including base excision repair, nucleotide excision repair, mismatch repair, and homologous recombination, were reduced by an average of ~5-fold in the absence of ammonium (Data Set S7). The expression of genes encoding subunits of RNA polymerase and those involved in RNA degradation were unaffected by the absence of ammonium (Data Set S7, pathways hmo03020 and hmo03018).

The ribosome is a major cellular component – ribosomal RNA (rRNA) is by far the most abundant form of RNA and ribosomal proteins are also abundant proteins. On average, expression of rRNA and ribosomal proteins was not affected by the absence of ammonium (Data Set S7, pathway hmo03010). Expression of tRNAs and aminoacyl-tRNA synthetases exhibited a ~3.4-fold reduction in message level (Data Set S7, pathway hmo00970).

Like all Gram-positive bacteria, heliobacteria lack an outer membrane and periplasm. Only one membrane must be crossed to direct a protein outside the cell. Heliobacteria possess both the Sec-SRP pathway and the Twin-arginine targeting (Tat) pathways for protein export. The expression of genes involved in the bacterial secretion systems was not significantly affected by ammonium depletion (Data Set S7, pathway hmo03060).

Regulation of active transport, flagella, and chemotaxis

ABC (ATP binding cassette) transporters couple ATP hydrolysis to active transport of a wide variety of molecules. *H. modesticaldum* has 11 different ABC transporter systems, which import phosphate, sulfate/thiosulfate, mineral ions (zinc, nickel/cobalt, tungstate, molybdate), monosaccharides (ribose/xylose), amino acids (glutamine, spermidine/putrescine, branch chain amino acids), and vitamins (biotin). There are also two ABC-2 type systems for export of teichoic acid (for lipopolysaccharide synthesis) and bacitracin (presumably for resistance to this antibiotic), as well as the FtsEX complex involved in cell division. Expression of genes encoding ABC transporters decreased by an average of 2.6-fold in the absence of ammonia (Data Set S8, pathway hmo02010).

The phosphotransferase system (PTS) is responsible for importing carbohydrates, such as hexoses and disaccharides. In order to prevent the acquired substrate from prematurely exiting the cell, phosphotransferase chemically modifies the substrate via phosphorylation using PEP as the phosphoryl donor. Only one gene in this pathway was identified in *H. modesticaldum*, *ptsI* (encoding phosphoenolpyruvate-protein phosphotransferase), and it was down-regulated by 7-fold in the absence of ammonia (Data Set S8, pathway hmo02060). It is unclear if the PTS system exists in this organism.

Two-component signal transduction systems consisting of a sensor and regulator allow cells to sense changes in environment, such as carbon sources, pH, quorum signals, and osmolarity. Expression of the genes involved in these systems decreased 3.5-fold on average in the absence of ammonium (Data Set S8, pathway hmo02020). A notable exception to this trend is *csrA*, which underwent a 7-fold increase in expression. This gene encodes the carbon storage regulator protein, an RNA-binding protein that is a pleiotropic regulator of many genes. The CsrA protein facilitates the destruction of many mRNAs, presumably by binding to their Shine-Dalgarno sequences and inhibiting translation, thus removing the protection from nucleases provided by the ribosome [49-53]. It is possible that increased expression of this protein is at least partially responsible for the decrease in the steady-state level of many other mRNAs.

Bacterial chemotaxis allows *H. modesticaldum* to regulate the speed of rotation and direction of its flagella in response to perceived chemical gradients. Transcripts of genes in this pathway were reduced an average of ~6-fold in the absence of ammonium (Data Set S8, pathway hmo02030). Similarly, the expression of genes involved in flagellar assembly (*i.e.* flagellar components or assembly proteins) was reduced by ~7-fold (Data Set S8, pathway hmo02040). There were 6 genes in this group that exhibited a moderate increase in expression; exclusion of these 6 genes was required to calculate a change in expression that was statistically significant. The gene that demonstrated the largest increase in expression, *flgM*, encodes for a negative regulator for flagellin synthesis. It is unclear why the transcripts of some genes involved in flagellar synthesis are increasing, while those of most of the genes are decreasing in abundance after the shift to ammonium-deplete medium.

Importance

The Heliobacteria appear to be the result of a lateral gene transfer of photosynthetic genes into an ancestral bacterium likely resembling a Clostridial cell, and indeed they can grow in the dark fermentatively like their non-photosynthetic cousins. The transcriptome reported here provides insights into how such a hybrid organism operates under normal conditions, and gives us the first view at a large scale of how the genetic information in its genome can be used under different conditions, such as when it is forced to use atmospheric N$_2$ as a source of nitrogen. These results should prove to be foundational for future studies on this particular organism (*H. modesticaldum*), which is rapidly becoming a model organism for the Heliobacteria.

This study reports the first transcriptome of *H. modesticaldum*. We also report the genome-wide effect that the absence of NH$_4^+$ has on expression. These findings show that the shift from ammonium-replete to N$_2$-fixing conditions is accompanied by a general decrease of ~4-fold in 74% of the genome. There was also increased expression of nitrogenase and the high-affinity ammonium transporter, along with a few other proteins, including some regulators (*e.g.* PII proteins and *csrA*). We also observed a significant decrease in the level of the mRNAs encoding the photosynthesis RC core polypeptide and electron donors/acceptors (cyt c_{553} and ferredoxins), which are among the most highly expressed protein-coding genes in ammonium-replete conditions. However, functional assays of the RC revealed only a modest decrease of the protein *in vivo*, indicating a significant role for regulation of translation or post-translational processes. Proteomic studies will be required in the future to unravel these contradictory responses. Future applications of the two *H. modesticaldum* transcriptomes reported will likely benefit from additional RNA-seq studies using a variety of growth conditions and comparative RNA-seq studies of other heliobacterial species.

References

1. Gest H, Favinger JL (1983) Heliobacterium chlorum, an anoxygenic brownish-green bacterium containing a `new' form of bacteriochlorophyll. Arch Microbiol 136: 11-16.

2. Asao M, Madigan MT (2010) Taxonomy, phylogeny, and ecology of the heliobacteria. Photosynth Res 104: 103-111.

3. Amesz J (1995) The antenna-reaction center complex of heliobacteria. Advances in Photosynthesis 2: 687-697.

4. Kimble LK, Stevenson AK, Madigan MT (1994) Chemotrophic growth of heliobacteria in darkness. FEMS Microbiol Lett 115: 51-55.

5. Asao M, Madigan MT, Family IV. Heliobacteriaceae, in The Firmicutes, de Vos P, et al., Editors. 2009, Springer: Dordrecht. Pp: 913-920.

6. Kimble L, Madigan MT (1992) Nitrogen fixation and nitrogen metabolism in heliobacteria. Arch Microbiol 158: 155-61.

7. Tang KH, Feng X, Zhuang WQ, Alvarez-Cohen L, Blankenship RE, et al. (2010) Carbon flow of heliobacteria is related more to clostridia than to the green sulfur bacteria. J Biol Chem 285: 35104-35112.

8. Sattley WM, Madigan MT, Swingley WD, Cheung PC, Clocksin KM, et al. (2008) The genome of Heliobacterium modesticaldum, a phototrophic representative of the Firmicutes containing the simplest photosynthetic apparatus. J Bacteriol 190: 4687-4696.

9. Kimble LK, Mandelco L, Woese CR, Madigan MT (1995) Heliobacterium modesticaldum, sp. nov., a thermophilic heliobacterium of hot springs and volcanic soils. Arch Microbiol 163: 259-267.

10. Xiong J, Inoue K, Bauer CE (1998) Tracking molecular evolution of photosynthesis by characterization of a major photosynthesis gene cluster from Heliobacillus mobilis. Proc Natl Acad Sci USA 95: 14851-14856.

11. Sattley WM, Blankenship RE (2010) Insights into heliobacterial photosynthesis and physiology from the genome of Heliobacterium modesticaldum. Photosynth Res 104: 113-122.

12. Burris RH, Roberts GP (1993) Biological nitrogen fixation. Annu Rev Nutr 13: 317-335.

13. Howard JB, Rees DC (1996) Structural Basis of Biological Nitrogen Fixation. Chem Rev 96: 2965-2982.

14. Raymond J, Siefert JL, Staples CR, Blankenship RE (2004) The natural history of nitrogen fixation. Mol Biol Evol 21: 541-554.

15. Kimble L, Madigan M (1992) Nitrogen fixation and nitrogen metabolism in heliobacteria. Arch Microbiol 158: 155-161.

16. Kimble LK, Stevenson AK, Woese CR, Madigan MT (1995) Heliobacterium modesticaldum, sp. nov., a thermophilic heliobacterium of hot springs and volcanic soils. Arch Microbiol 163: 259-267.

17. Tang KH, Yue H, Blankenship RE (2010) Energy metabolism of Heliobacterium modesticaldum during phototrophic and chemotrophic growth. BMC Microbiol 10: 150.

18. Lin CC, Casida LE (1984) GELRITE as a Gelling Agent in Media for the Growth of Thermophilic Microorganisms. Appl Environ Microbiol 47: 427-429.

19. Schmieder R, Edwards R (2011) Quality control and preprocessing of metagenomic datasets. Bioinformatics 27: 863-864.

20. Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nature Methods 9: 357-359.

21. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B, et al. (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nature Methods 5: 621-628.

22. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, et al. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13: 2498-2504.

23. Kanehisa M, Goto S, Kawashima S, Okuno Y, Hattori M, et al. (2004) The KEGG resource for deciphering the genome. Nucleic Acids Res 32: D277-D280.

24. Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25: 402-408.

25. Sarrou I, Khan Z, Cowgill J, Lin S, Brune D, et al. (2012) Purification of the photosynthetic reaction center from Heliobacterium modesticaldum. Photosynth Res 111: 291-302.

26. Giannoukos G, Ciulla DM, Huang K, Haas BJ, Izard J, et al. (2012) Efficient and robust RNA-seq process for cultured bacteria and complex community transcriptomes. Genome Biol 13: R23.

27. Waters LS, Storz G (2009) Regulatory RNAs in bacteria. Cell 136: 615-628.

28. Georg J, Voss B, Scholz I, Mitschke J, Wilde A, et al. (2009) Evidence for a major role of antisense RNAs in cyanobacterial gene regulation. Mol Syst Biol 5: 305.

29. Kimble LK, Madigan MT (1992) Evidence for an alternative nitrogenase in Heliobacterium gestii. FEMS Microbiol Lett 100: 255-260.

30. Drepper T, Gross S, Yakunin AF, Hallenbeck PC, Masepohl B, et al. (2003) Role of GlnB and GlnK in ammonium control of both nitrogenase systems in the phototrophic bacterium Rhodobacter capsulatus. Microbiology 149: 2203-2212.

31. Dodsworth JA, Leigh JA (2006) Regulation of nitrogenase by 2-oxoglutarate-reversible, direct binding of a PII-like nitrogen sensor protein to dinitrogenase. Proc Natl Acad Sci USA 103: 9779-9784.

32. Ninfa AJ, Atkinson MR (2000) PII signal transduction proteins. Trends Microbiol 8: 172-179.

33. Enkh-Amgalan J, Kawasaki H, Seki T (2006) Molecular evolution of the nif gene cluster carrying nifl1 and nifl2 genes in the Gram-positive phototrophic bacterium Heliobacterium chlorum. Int J Syst Evol Microbiol 56: 65-74.

34. Enkh-Amgalan J, Kawasaki H, Oh-oka H, Seki T (2006) Cloning and characterization of a novel gene involved in nitrogen fixation in Heliobacterium chlorum: a possible regulatory gene. Arch Microbiol 186: 327-337.

35. Nygaard TP, Rovira C, Peters GH, Jensen MO (2006) Ammonium recruitment and ammonia transport by E. coli ammonia channel AmtB. Biophys J 91: 4401-4412.

36. Huergo LF, Pedrosa FO, Muller-Santos M, Chubatsu LS, Monteiro RA, et al. (2012) PII signal transduction proteins: pivotal players in post-translational control of nitrogenase activity. Microbiology 158: 176-190.

37. Zhang T, Yan Y, He S, Ping S, Alam KM, et al. (2012) Involvement of the ammonium transporter AmtB in nitrogenase regulation and ammonium excretion in Pseudomonas stutzeri A1501. Res Microbiol 163: 332-339.

38. Amesz J (1989) Energy transfer and electron transport in Heliobacterium chlorum. Photosynthetica 23: 403-410.

39. Heinnickel M, Golbeck JH (2007) Heliobacterial photosynthesis. Photosynth Res 92: 35-53.

40. Sattley WM, Asao M, Tang JK-H, Collins AM (2014) Energy Conservation in Heliobacteria: Photosynthesis and Central Carbon Metabolism, in The Structural Basis of Biological Energy Generation, Hohmann-Marriott MF, Editor. 2014, Springer Science+Business Media: Dordrecht. Pp: 231-247.

41. Gisriel C, Sarrou I, Ferlez B, Golbeck JH, Redding KE, et al. (2017) Structure of a symmetric photosynthetic reaction center-photosystem. Science 357: 1021-1025.

42. Romberger SP, Golbeck JH (2012) The FX iron-sulfur cluster serves as the terminal bound electron acceptor in heliobacterial reaction centers. Photosynth Res 111: 285-290.

43. Redding KE, Sarrou I, Rappaport F, Santabarbara S, Lin S, et al. (2014) Modulation of the fluorescence yield in heliobacterial cells by induction of charge recombination in the photosynthetic reaction center. Photosynth Res 120: 221-235.

44. Kashey TS, Cowgill JB, McConnell MD, Flores M, Redding KE, et al. (2014) Expression and characterization of cytochrome c553 from Heliobacterium modesticaldum. Photosynth Res 120: 291-299.

45. Vignais PM (2008) Hydrogenases and H(+)-reduction in primary energy conservation. Results Probl Cell Differ 45: 223-252.

46. Ludwig M, Schulz-Friedrich R, Appel J (2006) Occurrence of hydrogenases in cyanobacteria and anoxygenic photosynthetic bacteria: implications for the phylogenetic origin of cyanobacterial and algal hydrogenases. J Mol Evol 63: 758-768.

47. Buckel W (2001) Sodium ion-translocating decarboxylases. Biochim Biophys Acta 1505: 15-27.

48. Li F, Hagemeier CH, Seedorf H, Gottschalk G, Thauer RK, et al. (2007) Re-citrate synthase from Clostridium kluyveri is phylogenetically related to homocitrate synthase and isopropylmalate synthase rather than to Si-citrate synthase. J Bacteriol 189: 4299-4304.

49. Seefeldt LC, Hoffman BM, Dean DR (2009) Mechanism of Mo-dependent nitrogenase. Annu Rev Biochem 78: 701-722.

50. Liu MY, Gui G, Wei B, Preston JF, Oakford L, et al. (1997) The RNA molecule CsrB binds to the global regulatory protein CsrA and antagonizes its activity in Escherichia coli. J Biol Chem 272: 17502-17510.

51. Gudapaty S, Suzuki K, Wang X, Babitzke P, Romeo T, et al. (2001) Regulatory interactions of Csr components: the RNA binding protein CsrA activates csrB transcription in Escherichia coli. J Bacteriol 183: 6017-6027.

52. Wei BL, Brun-Zinkernagel AM, Simecka JW, Pruss BM, Babitzke P, et al. (2001) Positive regulation of motility and flhDC expression by the RNA-binding protein CsrA of Escherichia coli. Mol Microbiol 40: 245-256.

53. Yang TY, Sung YM, Lei GS, Romeo T, Chak KF, et al. (2010) Posttranscriptional repression of the cel gene of the ColE7 operon by the RNA-binding protein CsrA of Escherichia coli. Nucleic Acids Res 38: 3936-3951.