Identification of Missing Genes and Enzymes for Autotrophic Carbon Fixation in Crenarchaeota

W. Hugo Ramos-Vera, Michael Weiss, Eric Strittmatter, Daniel Kockelkorn, and Georg Fuchs*  
Mikrobiologie, Fakultät Biologie, Albert-Ludwigs-Universität, Freiburg, Germany  
Received 28 September 2010/Accepted 12 December 2010

Two autotrophic carbon fixation cycles have been identified in Crenarchaeota. The dicarboxylate/4-hydroxybutyrate cycle functions in anaerobic or microaerobic autotrophic members of the Thermoproteales and Desulfurococcales. The 3-hydroxypropionate/4-hydroxybutyrate cycle occurs in aerobic autotrophic Sulfolobales; a similar cycle may operate in autotrophic aerobic marine Crenarchaeota. Both cycles form succinyl-coenzyme A (CoA) from acetyl-CoA and two molecules of inorganic carbon, but they use different means. Both cycles have in common the (re)generation of acetyl-CoA from succinyl-CoA via identical intermediates. Here, we identified several missing enzymes/genes involved in the seven-step conversion of succinyl-CoA to two molecules of acetyl-CoA in Thermoproteus neutrophilus (Thermoproteales), Ignicoccus hospitalis (Desulfurococcales), and Metallosphaera sedula (Sulfolobales). The identified enzymes/genes include succinyl-CoA reductase, succinic semialdehyde reductase, 4-hydroxybutyrate-CoA ligase, bifunctional crotonyl-CoA hydratase/(S)-3-hydroxybutyryl-CoA dehydrogenase, and beta-ketothiolase. 4-Hydroxybutyryl-CoA dehydratase, which catalyzes a mechanistically intriguing elimination of water, is well conserved and rightly can be considered the key enzyme of these two cycles. In contrast, several of the other enzymes evolved from quite different sources, making functional predictions based solely on genome interpretation difficult, if not questionable.

For a long time, Crenarchaeota (Archaea) generally were thought to live in volcanic habitats requiring a thermophilic lifestyle. Many of them can grow on purely inorganic media, and they often use various inorganic sulfur compounds, such as hydrogen sulfide, elemental sulfur, or pyrite, either as electron donor or electron acceptor, depending on an aerobic or anaerobic lifestyle. They are, therefore, often referred to as sulfur-dependent Archaea (for a review, see reference 32). Many chemolithotrophic members of this group are facultative autotrophs. However, in the past years, mesophilic, chemolithoautotrophic, ammonia-oxidizing, aerobic Crenarchaeota were found that are abundant mainly in the oceans, although some occur in soils; they are only distantly related to the other phyla of the Crenarchaeota (20, 23, 34, 35).

Two autotrophic carbon fixation cycles have been identified in Crenarchaeota (reviewed in reference 9). The dicarboxylate/4-hydroxybutyrate cycle (in short, dicarboxylate/hydroxybutyrate cycle) functions in the anaerobic or microaerobic autotrophic members of the Thermoproteales (such as Thermoproteus neutrophilus) (7, 41, 43–45, 51) and Desulfurococcales (such as Ignicoccus hospitalis) (24, 31) (Fig. 1B). Many of them grow as strict anaerobes by reducing elemental sulfur with H2S, but some grow under microaerobic or denitrifying conditions (e.g., Pyrolobus fumarii [Desulfurococcales]) (10, 28). The cycle can be divided into two parts. In the first part, acetyl-coenzyme A (CoA), one CO2 unit, and one bicarbonate unit are transformed via C4 dicarboxylic acids to succinyl-CoA. This part uses an oxygen-sensitive enzyme (pyruvate synthase) and electron donor (ferredoxin). This may explain why these reactions are restricted to anaerobes or microaerobes, at best.

In the second part, succinyl-CoA is reduced in two steps to 4-hydroxybutyrate. 4-Hydroxybutyrate transformation proceeds via 4-hydroxybutyryl-CoA and requires the key enzyme 4-hydroxybutyryl-CoA dehydratase. The product of the dehydratase, crotonyl-CoA, is converted into two molecules of acetyl-CoA via β-oxidation reactions. One acetyl-CoA unit can be used for biosynthesis, and the second one serves as a CO2 acceptor for the next round of the cycle.

The 3-hydroxypropionate/4-hydroxybutyrate cycle (in short, hydroxypropionate/hydroxybutyrate cycle) functions in autotrophic Sulfolobales such as Metallosphaera sedula (8, 29) (Fig. 1A). This group comprises extreme thermoacidophiles from volcanic areas that grow best at pH ~2 and a temperature of 60 to 90°C. Most Sulfolobales species can grow chemoautotrophically on sulfur, pyrite, or H2 under (micro)aerobic conditions (25, 27), but some became secondarily facultative anaerobic or even strictly anaerobic (e.g., Stygiolobus azoricus) (47). Their carbon fixation cycle differs basically in the way that succinyl-CoA is formed from acetyl-CoA and CO2. The key carboxylating enzyme is the bifunctional biotin-dependent acetyl-CoA/propionyl-CoA carboxylase (30). The enzymes of the hydroxypropionate/hydroxybutyrate cycle are oxygen tolerant. Therefore, this cycle fits in with the lifestyle of the aerobic Crenarchaeota.

Whereas the formation of succinyl-CoA from acetyl-CoA and CO2 in the two cycles is well understood and most genes could be identified, various enzymes and genes of the common 4-hydroxybutyrate part of the two cycles were not defined. It was the aim of this study to close this gap. Fortunately, the genomes of M. sedula (5), T. neutrophilus (data available from the DOE Joint Genome Institute at http://www.jgi.doe.gov/),
and *I. hospitalis* (DOE Joint Genome Institute) (39) have been sequenced. Many common traits became apparent, but a substantial disparity that causes difficulties to predict enzyme functions *in silico* also became apparent. This caveat is noteworthy in silico and with regard to our proposal (8) that the abundant autotrophic, nitrifying “marine group I” *Crenarchaeota* use a (modified) hydroxypropionate/hydroxybutyrate cycle. The proposal is based on the presence of the conserved genes coding for the characteristic enzymes acetyl-CoA/propionyl-CoA carboxylase, methylmalonyl-CoA mutase, and 4-hydroxybutyryl-CoA dehydratase in the genomes of *Cenarchaeum symbiosum* and *Ignicoccus hospitalis* (see reference 9 and literature cited therein). (B) Dicarboxylate/4-hydroxybutyrate cycle, as proposed for *Desulfuroccales* and *Thermoproteales* (24, 41). The enzymes studied here in *Metallosphaera sedula*, *Thermoproteus neutrophilus*, and *Ignicoccus hospitalis* are numbered 10 to 16 below; the identified genes are given in parentheses. Enzymes: 1, acetyl-CoA carboxylase; 2, malonyl-CoA reductase (NADPH); 3, malonic semialdehyde dehydrogenase (NAD); 4, 3-hydroxypropionate-CoA ligase (AMP forming); 5, 3-hydroxypropionyl-CoA dehydratase; 6, acryloyl-CoA reductase (NADPH); 7, propionyl-CoA carboxylase; 8, methylmalonyl-CoA epimerase; 9, methylmalonyl-CoA mutase; 10, succinyl-CoA reductase (NADPH reductase); 11, succinic semialdehyde dehydrogenase (NAD); 12, 4-hydroxybutyrate-CoA ligase (AMP forming) (Msed_0399, Tneu_0420, Ighi_?); 13, 4-hydroxybutyryl-CoA dehydratase (Msed_0399, Tneu_0422, Ighi_?); 14, crotonyl-CoA ligase (AMP forming) (Msed_0399, Tneu_0420, Ighi_?); 15, (S)-3-hydroxybutyryl-CoA dehydrogenase (NAD<sup>+</sup>) (Msed_0399, Tneu_0541, Ighi_1058); 16, acetoacetyl-CoA beta-ketothiolase (Msed_0656, Tneu_0249, Ighi_1401); 17, pyruvate synthase (reduced methyl viologen); 18, pyruvate:water dikinase; 19, PEP carboxylase; 20, malate dehydrogenase (NADH); 21, fumarate hydratase; 22, fumarate reductase (reduced methyl viologen); 23, succinyl-CoA synthetase (ADP forming). Fdred, reduced ferredoxin; MV, methyl viologen.

**MATERIALS AND METHODS**

**Materials.** Chemicals were obtained from Fluka (Neu-Ulm, Germany), Sigma-Aldrich (Deisenhofen, Germany), Qiagen (Hilden, Germany), GE Healthcare (Munich, Germany), VWR (Darmstadt, Germany), and Roth (Karlsruhe, Germany). Gases were obtained from Sauerstoffwerke Friedrichshafen (Friedrichshafen, Germany). Enzymes and primers were obtained from MBI Fermentas (St. Leon-Rot, Germany) and Sigma-Aldrich. Materials and equipment for protein purification were obtained from GE Healthcare, Millipore (Bedford, MA), and Sigma-Aldrich. Plasmids were obtained from Novagen (Darmstadt, Germany).

**Strains and culture conditions.** *T. neutrophilus* (DSM 2338) was a kind gift of K. O. Stetter and H. Huber from the culture collection of the Lehrstuhl für Mikrobiologie, University of Regensburg. It was grown anaerobically and autotrophically on a defined mineral medium with elemental sulfur under gassing with a mixture of 80% H<sub>2</sub> and 20% (vol/vol) CO<sub>2</sub> at 85°C and pH 6.8 (41, 43). *M. sedula* TH2 (DSM 5348) was grown autotrophically at 75°C on a chemically defined medium (pH 2.0) under gassing with a mixture of 19% CO<sub>2</sub>, 3% O<sub>2</sub>, and 78% H<sub>2</sub> with a generation time of 20 h (1, 26). Autotrophically grown *I. hospitalis* cells were a kind gift of H. Huber, Regensburg, Germany. *Escherichia coli* strains DH5<sup>a</sup> and Rosetta 2 (DE3) (Merck, Germany) were grown at 37°C in lysogeny broth (LB) medium. Antibiotics were added to E. coli cultures to a final concentration of 100 μg of ampicillin ml<sup>−1</sup> and 34 μg of chloramphenicol ml<sup>−1</sup>. Cells were stored frozen in liquid nitrogen until use.

**Preparation of cell extracts.** Cell extracts were prepared under anoxic conditions. Cells were suspended in an equal amount of 10 mM Tris-HCl (pH 7.8), and the cell suspension was passed through a chilled French pressure cell at 137 MPa. The lysate was ultra centrifuged for 1 h (100,000 × g; 4°C), and aliquots of the supernatant (cell extract) were stored anoxically at −70°C until use.

**Syntheses.** Acetoacetyl-CoA was synthesized from diketene (48). Succinyl-CoA and crotonyl-CoA were synthesized from their anhydrides by a slightly modified method (48); the deviations were anoxic conditions and room temperature. (R)- and (S)-3-hydroxybutyryl-CoA were synthesized by the mixed anhydride method (50). The dry powders of the CoA esters were stored at −20°C.
Enzyme assays. Spectrophotometric enzyme assays (0.5 ml) were performed in 0.5 ml glass cuvettes at 65°C, unless otherwise indicated. None of the enzymes was oxygen sensitive; we controlled this by using a strict anaerobic incubation, including 5 mM DTT. The pH of the buffers was adjusted to the indicated value at 20°C. Reactions involving NAD(P)H were measured at 365 nm ($\epsilon_{max} = 3.4 \times 10^4 M^{-1} cm^{-1}$; $\epsilon_{max} = 3.5 \times 10^4 M^{-1} cm^{-1}$). [11]

The reaction mixture for succinyl-CoA reductase (EC 1.2.1.) contained 100 mM Tris-HCl (pH 7.8), 5 mM MgCl$_2$, 5 mM DTT, 0.5 mM NADPH, and purified enzyme. The reaction was started by the addition of 0.2 mM succinyl-CoA. For $K_m$ determination, the concentration of one substrate was varied (succinyl-CoA, 0.05 to 3 mM; NADPH, 0.05 to 2 mM) while keeping the concentration of the other substrate constant (succinyl-CoA, 1 mM; NADPH, 0.5 mM).

The enzyme succinic semialdehyde reductase (EC 1.1.1.) was determined in a reaction mixture containing 100 mM 3-(N-morpholino)propanesulfonic acid (MOPS)-KOH (pH 7.8), 5 mM MgCl$_2$, 5 mM DTT, 0.5 mM NADPH, and purified enzyme. The reaction was started by the addition of 1 mM succinic semialdehyde. For $K_m$ determination, the concentration of one substrate was varied (succinic semialdehyde, 0.01 to 0.5 mM; NADPH, 0.01 to 1 mM) while keeping the concentration of the other substrate constant (succinic semialdehyde, 1 mM; NADPH, 0.5 mM).

4-Hydroxybutyryl-CoA ligase (EC 6.2.1.1.) activity was measured at 75 or 85°C by monitoring CoA ester formation. The reaction mixture (0.35 ml) contained 100 mM MOPS-KOH (pH 7.8), 10 mM MgCl$_2$, 6 mM ATP, 0.5 mM CoA, 10 mM 4-hydroxybutyrate, and purified enzyme. Acetate, propionate, 3-hydroxypropionate, (S)-3-hydroxybutyrate, or crotonate also was tested. The reaction (0.1 ml) was stopped after 0, 1, and 2 min by the addition of 10 µl of 1 M HCl, and the products were analyzed by reversed-phase high-pressure liquid chromatography (RP-HPLC) using an RP-C$_{18}$ column (18).

A discontinuous assay based on the determination of the amount of AMP formed was used to measure $K_m$ values. The reaction mixture (0.35 ml) contained 100 mM MOPS-KOH (pH 7.8), 5 mM MgCl$_2$, 3 mM ATP, 0.5 mM CoA, and 10 mM 4-hydroxybutyrate. The reaction mixture was then supplemented with the purifying enzyme. After 0, 2, and 4 min of incubation at 85°C, 0.1 ml of the reaction mixture was removed and diluted in 0.5 ml 100 mM MOPS-KOH (pH 7.8), 5 mM MgCl$_2$, 5 mM ATP, 20 mM KCl, 0.75 mM NADH, 1 mM phosphoenolpyruvate, 4 U myokinase, 3 U pyruvate kinase, and 2 U lactate dehydrogenase (at 30°C). NADH consumption was monitored at 365 nm.

Crotonyl-CoA hydratase/3-hydroxybutyryl-CoA dehydrogenase (EC 4.2.1.17 and EC 1.1.1.157, respectively) was measured spectrophotometrically at 65°C by monitoring CoA ester formation. The reaction mixture (0.35 ml) contained 100 mM MOPS-KOH (pH 7.8), 10 mM MgCl$_2$, 6 mM ATP, 0.5 mM CoA, 10 mM 4-hydroxybutyrate, and purified enzyme. The reaction was started by the addition of purified enzyme. After 0, 2, and 4 min of incubation at 85°C, 0.1 ml of the reaction mixture was removed and diluted in 0.5 ml 100 mM MOPS-KOH (pH 7.8), 5 mM MgCl$_2$, 5 mM ATP, 20 mM KCl, 0.75 mM NADH, 1 mM phosphoenolpyruvate, 4 U myokinase, 3 U pyruvate kinase, and 2 U lactate dehydrogenase (at 30°C). NADH consumption was monitored at 365 nm.

4-Hydroxybutyrate-CoA ligase (EC 2.3.2.1) was measured at 65°C by monitoring CoA ester formation from 4-hydroxybutyric-CoA. The reaction mixture (0.1 ml) contained 100 mM MOPS-KOH (pH 7.2), 5 mM MgCl$_2$, 5 mM DTT, 1 mM CoA, 1 mM acetoacetyl-CoA, and purified enzyme. The reaction was stopped after 0, 0.5, and 1 min by the addition of 10 µl of 1 M HCl, and the products were analyzed by RP-HPLC using an RP-C$_{18}$ column (18). For the determination of the apparent $K_m$ the concentration of acetoacetyl-CoA was varied (0.1 to 1 mM).

Heterologous expression of genes from T. neutrophilus, M. sedula, and I. hospitalis in E. coli. Standard protocols were used for the purification, preparation, cloning, transformation, and amplification of DNA (6). Chromosomal DNA from T. neutrophilus, M. sedula, and I. hospitalis was isolated using the Illustra bacteria genomic Prep mini spin kit (GE Healthcare). Plasmid DNA was isolated using the QIAprep spin miniprep kit (Qiagen). The genes were amplified from chromosomal DNA by PCR using Taq polymerase. Primers are listed in Table S1 in the supplemental material. After purification with a QIAquick PCR purification kit, the PCR product was treated with the corresponding restriction enzymes and normally ligated into the expression vectors pET16b (N-terminal His$_{10}$ tag) and pET7-7 (52) in the case of T_neu_0421 and Mseq_0399 and pEXP5-CT TOPO (Invitrogen, Karlsruhe, Germany) in the case of Mseq_1422 (C-terminal His$_{10}$ tag omitted). The constructs were transformed into E. coli DH5α. The inserted sequences were sequenced by GATC-Biotech (Konstanz, Germany). The proteins were overexpressed in E. coli Rosetta 2 (DE3). The cells were transformed with the construct, grown in 1-liter cultures at 37°C in LB media containing 100 µg of ampicillin ml$^{-1}$ and 34 µg of chloramphenicol ml$^{-1}$. E. coli Rosetta 2 (DE3). The cells were transformed with the construct, grown in 1-liter cultures at 37°C in LB media containing 100 µg of ampicillin ml$^{-1}$ and 34 µg of chloramphenicol ml$^{-1}$. E. coli Rosetta 2 (DE3). The cells were transformed with the construct, grown in 1-liter cultures at 37°C in LB media containing 100 µg of ampicillin ml$^{-1}$ and 34 µg of chloramphenicol ml$^{-1}$. E. coli Rosetta 2 (DE3). The cells were transformed with the construct, grown in 1-liter cultures at 37°C in LB media containing 100 µg of ampicillin ml$^{-1}$ and 34 µg of chloramphenicol ml$^{-1}$.
FIG. 2. Transcriptional organization of a gene cluster in T. neutrophilus that is involved in the dicarboxylate/4-hydroxybutyrate cycle. cbs, three genes coding for proteins containing cystathionine β-synthase domains; pepc, PEP carboxylase gene; ssr, succinic semialdehyde reductase gene; 4hbl, 4-hydroxybutyrate-CoA ligase gene; scr, succinyl-CoA reductase gene; 4hbd, 4-hydroxybutyryl-CoA dehydratase gene; frdA, fumarate reductase subunits A and B genes. Bar, 500-bp scale.

Msed_1422. The enzyme Msed_1422 was produced untagged and purified from 3 g (wet mass) E. coli cells. It was partially purified via heat precipitation at 75°C for 15 min, cooled on ice for 15 min, and centrifuged at 100,000 × g at 4°C for 1 h. Heat-precipitated cell extract was stored at −20°C with 25% glycerol.

Msed_1353 and Msed_1291. The His6-tagged enzymes Msed_1353 and Msed_1291 were purified from E. coli cells as described above for Tneu_0421, only the heat precipitation step was performed at 75°C.

Acetate-CoA ligase Ignt_0256-7. The His6-tagged enzyme acetate-CoA ligase Ignt_0256-7 was purified from 5 g (wet mass) E. coli cells as described above for Tneu_0421, only the heat precipitation step was performed at 90°C.

LC-MS/MS and data analysis of the purified proteins. For in-gel digestion, the excised gel bands were destained with 30% (vol/vol) acetonitrile, shrunk with 100% acetonitrile, and dried in an Eppendorf vacuum concentrator. Digests with trypsin and elastase were performed overnight at 37°C in 0.05 M NH4HCO3 (pH 8). Digests with thermolysin were performed for 2 h at 60°C in 0.05 M NH4HCO3 (pH 8). Approximately 0.1 μg of protease was used for one gel band. Peptides were extracted from the gel slices with 5% (vol/vol) formic acid. All liquid chromatography-mass spectrometry/mass spectrometry (LC-MS/MS) analyses were performed on a Q-TOF mass spectrometer (Agilent 6520; Agilent Technologies, Böblingen, Germany) and on an electron transfer dissociation (ETD) ion trap (Agilent 6340), both coupled to a 1200 Agilent nanoflow system via an HPLC-chip cube ESI (electrospray ionization) interface. ETD analyses on the ion trap were performed using the data-dependent acquisition mode. After an MS scan (standard enhanced mode), a maximum of three peptides were selected for ETD-MS/MS (standard enhanced mode). The automated gain control (IC) for MS scans was set to 350,000. The maximum accumulation time was set to 300 ms. The following ETD parameters were used: ICC target, 400,000; reaction time, 100 ms; cutoff, 140. Resonance excitation (Smart Decomp) was used for doubly charged peptides. Peptides were separated on an HPLC-chip with an analytical column of 75-μm inner diameter and 150-mm length and with a 40-nl trap column, both packed with Zorbax 300SB C18 bonded phase (5-μm particle size). Peptides were eluted with a linear acetonitrile gradient with 1%/min at a flow rate of 300 nL/min (starting with 5% acetonitrile). The Q-TOF was operated in the 2-GHz extended dynamic range mode. MS/MS analyses were performed using the data-dependent acquisition mode. After an MS scan (two spectra), a maximum of three peptides were selected for MS/MS (two spectra/s). Singly charged precursor ions were excluded from selection. Internal calibration was applied using two reference masses. Mascot Distiller 2.1 was used for raw data processing and for generating peak lists, essentially with standard settings for the Agilent Q-ToF. Mascot server 2.2 was used for database searching with the following parameters: peptide mass tolerance of 15 ppm, MS/MS mass tolerance of 0.02 Da, 13C of L. enzyme type “trypsin” with two uncharged sites allowed for trypsin and “none” for elastase and thermolysin, and variable modifications of carbamidomethyl (cysteine), Gln→pyroGlu (N-terminal Gln), oxidation (methionine), and acetylation (lysine). For protein identification the NCBI protein database was used. For protein modification (PTM) analysis, a small custom database containing the protein sequences of Tneu_0421 and Msed_1353 were used. All relevant MS/MS spectra were validated manually.

Protein-analyzing methods. Protein was determined by the method of Bradford (12) using bovine serum albumin as the standard. Protein fractions were analyzed by SDS−12.5% polyacrylamide gel electrophoresis (36). Proteins were visualized by Coomassie brilliant blue R-250 staining. The native molecular mass was determined on a 24-mL Superdex 200 10/300 GL gel filtration column (GE Healthcare) calibrated with vitamin B12 (1.35 kDa), RNase A (13.7 kDa), chymotrypsinogen A (25 kDa), ovalbumin (43 kDa), bovine serum albumin (67 and 134 kDa), aldolase (158 kDa), catalase (240 kDa), ferritin (440 kDa), and blue dextran 2000 (2,000 kDa).

Computational analysis. The BLAST searches were performed via the NCBI BLAST server (http://www.ncbi.nlm.nih.gov/BLAST) (3). The amino acid sequences were aligned using CLUSTALW (55) implemented within BioEdit software (http://www.mbio.ncsu.edu/BioEdit/bioedit.html).

RESULTS AND DISCUSSION

Succinyl-CoA reductase. Succinyl-CoA reductase was identified in M. sedula, where this enzyme catalyzes not only the NADPH-dependent reduction of succinyl-CoA to succinic semialdehyde but also that of malonyl-CoA to malonic semialdehyde (1, 33). In contrast, in T. neutrophilus only an NADPH-dependent succinyl-CoA reductase activity was detectable in extract of autotrophically grown cells, which was 10-fold downregulated in acetate-grown cells (41, 42). Searching the genome did not give an unmistakable hint at a succinyl-CoA reductase gene. A putative succinyl-CoA reductase gene (scr; Tneu_0421) was localized in a cluster that included the putative genes for fumarate reductase subunits A and B, succinic semialdehyde reductase, 4-hydroxybutyrate-CoA ligase, 4-hydroxybutyryl-CoA dehydratase, and PEP carboxylase (Fig. 2). We refer to this cluster as the autotrophy-related gene cluster (41, 42), which is also present in other autotrophic members of Thermoproteales, such as Pyrobaculum islandicum and Pyrobaculum calidifontis. The scr gene, which coded for a 51-kDa protein, was overexpressed as soluble N-terminal His-tagged protein in E. coli. The purified recombinant enzyme indeed catalyzed the NADPH-dependent reduction of succinyl-CoA to succinic semialdehyde (10% activity with NADH). Its kinetic constants measured at 65°C and an optimal pH of 8 were a Vmax of 35 μmol min−1 mg−1 (U) and apparent Km values of 80 μM for succinyl-CoA and 160 μM for NADPH. The enzyme, which acted neither on malonyl-CoA nor on methylmalonyl-CoA, showed high similarity (>72 and >85% amino acid sequence [aa] identity/similarity) to putative aldehyde dehydrogenases of other autotrophic members of the Thermoproteales, i.e., Pyrobaculum islandicum, Pyrobaculum calidifontis, and Thermoproteus tenax. These enzymes likely also act as succinyl-CoA reductases, since no other obvious candidate gene was found in the genome; still, this annotation needs to be proven owing to the fact that a related aldehyde dehydrogenase in M. sedula is non-CoA acylating and oxidizes succinic semialdehyde to succinate (19, 33).

Hence, at least three different kinds of succinyl-CoA reductase enzymes operate in Crenarchaeota: (i) the NADPH-dependent enzyme in Thermoproteales, as disclosed here, a member of the aldehyde dehydrogenase family; (ii) the enzyme in M. sedula (Msed_0709) and in other autotrophic Sulfolobales, an NADPH-dependent paralog of aspartate semialdehyde dehydrogenase (1, 33); (iii) the I. hospitalis enzyme (24), which
differs from both the *Thermoproteus* and the *Metallosphaera* enzyme, as it does not use NAD(P)H. Rather, it may use reduced ferredoxin as natural electron donor, and its gene has not been identified yet. Marine *Crenarchaeota (N. maritimus* and *C. symbiosum*) do not contain an apparent candidate gene for succinyl-CoA reductase similar to reductases of types i to iii. The *N. maritimus* gene *Nmar* 1608 shows some similarity to *Tnue* 0421 (33 and 53% aa identity and similarity); however, its function needs to be tested. Succinyl-CoA reductase from *Clostridium kluyveri* (sucD; accession number AAA92347) (49) has only low similarity (26 and 41% aa identity and similarity) to succinyl-CoA reductase from *T. neutrophilus* (*Tnue* 0421).

There may be reasons for this unexpected diversity of succinyl-CoA reductases. The promiscuous enzyme of *Sulfolobales*, which apparently is derived from a duplication of the aspartate semialdehyde dehydrogenase gene and further mutations, also serves for malonyl-CoA reduction, thus saving extra costs for making another enzyme. The *Ignicoccus* enzyme uses reduced ferredoxin, thus pulling away the heat-labile succinyl-CoA (approximate standard free energy change [\(\Delta G^\circ\)] of \(-6 \text{ kJ}\) instead of \(+12 \text{ kJ}\) with NADPH); this may be important for extreme thermophiles.

It should be stressed that predictions of gene function in *Crenarchaeota* are not easy, even if gene products show high similarity to known enzymes. This issue is illustrated by *M. sedula*, which has two genes coding for enzymes similar to the *Thermoproteus* succinyl-CoA reductase (*Msed* 1774, 50 and 72% aa identity and similarity; *Msed* 1119, 43 and 65% aa identity and similarity); however, none reduced succinyl-CoA, and the true function of both enzymes remained obscure (33). The first one oxidized succinic semialdehyde to succinate irreversibly, which may link the \(\text{CO}_2\) fixation cycle to the central carbon metabolism (19); the other had only low succinic semialdehyde dehydrogenase (non-CoA acylating) activity, and its true substrate remained unknown.

**Succinic semialdehyde reductase.** Succinic semialdehyde reductase was identified in *M. sedula* (33) but not in *T. neutrophilus*. *T. neutrophilus* contains a gene in the autotrophy-related gene cluster that is similar (53 and 75% aa identity and similarity) to *Msed* 1774 and purified the recombinant enzyme. It acted on 4-hydroxybutyrate (100%) but also slowly on crotonate (22%), acetate (13%), 3-hydroxypropionate (9%), and 3-hydroxybutyrate (8%). Its specific activity (85°C, pH 7.8) was only 1.6 U mg\(^{-1}\) and the apparent \(K_m\) for 4-hydroxybutyrate was 0.7 mM. The low activity indicates that the recombinant enzyme is modified and partly inactive. To test this possibility, the recombinant protein was digested in gel with trypsin, elastase, and thermolysin, and the generated peptide mixtures were analyzed by nano-LC-MS/MS on a Q-TOF instrument and on an ion trap instrument by applying electron transfer dissociation (ETD). Data analysis resulted in the identification of four different peptides acetylated on Lys594. VAILDLKPTRT GaK and LDKLPKTRT GaK were identified from the digest with thermolysin, and DKLKPRTGKaK and aKVMMRVL were identified from the digest with elastase. As an example, Fig. S1 in the supplemental material shows the ETD MS/MS spectrum of the peptide VAILDLKPRTGaK, proving acetylation on the C-terminal lysine residue. All detected peptides containing Lys594 are acetylated on Lys594. No peptide with unmodified Lys594 was detectable. Thus, we conclude that the degree of acetylation on this site is very high. This posttranslational inactivation by acetylation and reactivation by deacetylation has been reported for various bacterial organic acid-CoA ligases (17). The observed Lys acetylation may be an artifact of the *E. coli* expression system, which may hamper the overproduction of active CoA ligases. The genome of *T. neutrophilus* does not contain obvious candidate genes for acetylating or deacetylating enzymes related to the reported enzymes (17), but their existence cannot be excluded due to the lack of knowledge of the archaeal system. There are two related CoA ligase genes in *T. neutrophilus*. *Tnue* 1843 is closely related to acetate-CoA ligase (AMP-forming) PAE2867 from *Pyrobaculum aerophilum* (13), suggesting that it...
codes for this enzyme, whereas the function of Tneu_0385 is unknown.

The _M. sedula_ gene Msed_1456 codes for 3-hydroxypropionate-CoA ligase (2); the enzyme acted slowly on 4-hydroxybutyrate (22%). The purified 4-hydroxybutyrate-CoA ligase was tentatively identified before as Msed_1422 (8). Since a His tag may interfere with the formation of the correct active site, resulting in insoluble or soluble but inactive enzyme, Msed_1422 was overproduced in the nontagged form. The overexpression of the gene in _E. coli_ resulted in a soluble enzyme that acted on neither 4-hydroxybutyrate nor on acetate, propionate, 3-hydroxypropionate, 3-hydroxybutyrate, or crotonate. This enzyme has some similarity to LuxE acyl-protein synthetase; possibly it acts by the acylation of proteins rather than by the CoA activation of low-molecular-mass organic acids, and it may be involved in regulation by covalent modification. Alternatively, it was produced in an inactive form despite the fact that it was soluble. Candidates for the true 4-hydroxybutyrate-CoA ligase, Msed_1353 and Msed_1291, showed similarity to 4-hydroxybutyrate-CoA ligase from _T. neutralus_ as well as to 3-hydroxypropionate-CoA ligase Msed_1456 from _M. sedula_. The gene Msed_1353 (expected protein molecular mass, 75 kDa) was overexpressed as soluble, N-terminal His10-tagged protein in _E. coli_. However, the purified recombinant enzyme did not act on 4-hydroxybutyrate but on propionate (11 U mg\(^{-1}\) at 75°C and pH 7.8, 100%), acetate (75%), and 3-hydroxypropionate (32%) and therefore is probably a promiscuous acetate/propionate-CoA ligase. Recombinant, soluble N-terminal His10-tagged Msed_1291 protein did not act on any of these organic acids; possibly it was produced as soluble but inactive enzyme. Thus, there is no proven 4-hydroxybutyrate-CoA ligase gene in _M. sedula_ so far.

_I. hospitalis_ lacked a candidate gene for 4-hydroxybutyrate-CoA ligase. The gene Igni_0379 codes for a putative COA ligase that shows little similarity to 4-hydroxybutyrate-CoA ligase of _T. neutralus_ (23 and 40% aa identity and similarity). However, the open reading frames (ORF) Igni_0256 and Igni_0257 together showed high similarity to the _Thermoproteus_ enzyme. We therefore checked the integrity of both ORFs. It turned out that the 3’ region of Igni_0256 incorrectly contained a stretch of an additional 38 bp, indicating that the corrected Igni_0256 and Igni_0257 genes together constituted the gene for 4-hydroxybutyrate-CoA ligase (expected protein molecular mass, 73 kDa; 52 and 69% aa identity and similarity with the _T. neutralus_ enzyme). However, when the corrected gene was overexpressed, the purified N-terminal His10-tagged protein did not act on 4-hydroxybutyrate but on acetate (3 U mg\(^{-1}\) at 85°C and pH 7.8, 100%), propionate (50%), and 3-hydroxypropionate (33%). Marine _Crenarchaeota_ (_N. maritimus, C. symbiosum_) also do not contain an obvious 4-hydroxybutyrate-CoA ligase candidate gene when searched with the _T. neutralus_ enzyme (e.g., Nmar_0701, 34% identity). However, a gene encoding a putative ADP-forming acyl-CoA ligase similar to archaeal acetyl-CoA synthetase (ADP forming) (CENSYa_0021, Nmar_0206) is located in front of the 4-hydroxybutyryl-CoA dehydratase gene (CENSYa_0022, Nmar_0207), which may be a promising candidate; however, it has only marginal similarity to known AMP-forming CoA ligases.

**4-Hydroxybutyryl-CoA dehydratase.** A common key enzyme of the two carbon fixation cycles, 4-hydroxybutyryl-CoA dehydratase, also occurs in fermenting clostridia (14, 37), where it plays a role in gamma-aminobutyrate fermentation. Although the dehydratase is inactivated by oxygen in clostridia (46), it obviously is sufficiently oxygen insensitive in _Sulfobolales_ (10) to operate under microoxic or even oxic conditions. Cell extracts of autotrophic _Sulfobolales, Thermoproteales, and Desulfurococcales_ contained high 4-hydroxybutyryl-CoA dehydratase activity (8, 10, 24, 41), and their genomes contained the corresponding gene (Msed_1321, Tneu_0422, and Igni_0595, respectively). The crenarchaeal dehydratases are highly related to each other and to the enzyme from clostridia or _Acidaminococcus_ (∼40 and ∼60% aa identity and similarity). The characteristic gene also is present in marine _Crenarchaeota_ (_N. maritimus, C. symbiosum_) (Nmar_0207, CENSYa_0022). The overproduction of this enzyme (even in _Eubacteria_) has not been achieved yet (21), possibly owing to the fact that the assembly of the oxygen-sensitive 4Fe-4S cluster in _E. coli_ does not function. _Sulfobolales_ contain two genes for putative 4-hydroxybutyryl-CoA dehydratase (e.g., _Msed_1220 and _Msed_1321), one of which (_Msed_1220) lacks catalytically essential amino acids. Transcriptome studies of _M. sedula_ showed that of the two genes, only the true dehydratase _Msed_1321 was triggered by autotrophic versus heterotrophic conditions (4). Hence, the identity of this gene and its role in autotrophic carbon fixation seems to be established. The function of the second, related gene, however, remains unknown.

**Crotonyl-CoA hydratase and 3-hydroxybutyryl-CoA dehydrogenase.** The nature of the following two enzymes and their genes was less clear, namely, crotonyl-CoA hydratase yielding (S)-3-hydroxybutyryl-CoA and (S)-3-hydroxybutyryl-CoA dehydrogenase yielding acetoacetyl-CoA. We used 3-hydroxypropionyl-CoA dehydratase from _M. sedula_ (_Msed_2001; 28-kDa protein) as the search sequence for crotonyl-CoA (enoyl-CoA) hydratase, since _Msed_2001 also catalyzes the hydration of crotonyl-CoA (53). A similar enoyl-CoA hydratase gene was not found in the genome of _T. neutralus_ or _I. hospitalis_. However, the best hit was a domain of a gene coding for an apparent fusion protein (∼70 kDa) of an enoyl-CoA hydratase (C-terminal domain, ∼20 kDa) and a dehydrogenase (N-terminal domain, ∼40 kDa) (_Msed_0399, Tneu_0541, Igni_1058). Autotrophic _Sulfobolales_ contained additional, less similar candidate genes coding for members of the enoyl-CoA hydratase family (23 to 46 and 48 to 70% aa identity and similarity). A search with 3-hydroxyacyl-CoA dehydrogenase from _E. coli_ yielded similar results. It should be noted that fusion proteins are quite common in fatty acid metabolism; e.g., in _E. coli_, the tetrafunctional FadB, which is involved in fatty acid degradation, contains these two domains in reverse order and catalyzes also (R)/(S) epimerase and cis/trans isomerase reactions (15).

The expected bifunctional hydratase/dehydrogenase activity was present in similar amounts in extracts of autotrophically and heterotrophically grown cells of _M. sedula_ (0.6 U mg\(^{-1}\), 65°C, pH 8). The enzyme was purified more than 63-fold to near homogeneity with 1.4% yield from extracts of autotrophically grown cells using crotonyl-CoA and NAD\(^+\) as substrates in the assay (see Table S2 in the supplemental material). The native enzyme preparation catalyzed both hydratase and dehydrogenase reactions and had a specific activity of 38 U mg\(^{-1}\).
SDS-PAGE showed two major protein bands of 70 and 50 kDa. Mass-spectrometric analysis identified the 70-kDa protein as Msed_0399. To prove that the 70-kDa protein catalyzed both reactions, we cloned and overexpressed Msed_0399 in E. coli and purified the recombinant protein. Substrates were indeed crotonyl-CoA ($K_m$, 0.3 mM), 3-(S)-3-hydroxybutyryl-CoA ($K_m$, 0.2 mM), and NAD$^+$ ($K_m$, 0.25 mM). NADP$^+$, acryloyl-CoA, 3-hydroxypropionyl-CoA, and the (R) stereoisomer of 3-hydroxybutyl-CoA were not transformed. The $V_{max}$ was 13.8 U mg$^{-1}$ (at 65°C) measured at the optimal pH 8, no whether crotonyl-CoA or (S)-3-hydroxybutyl-CoA was tested as the substrate.

This bi-functional enzyme likely also operates in autotrophic carbon fixation in anaerobic Crenarchaeota, owing to the fact that it is the only candidate gene for the two enzyme functions in autotrophic Thermoproteales and Desulfurococcales. Some Sulfolobales contain an additional gene copy (<50 and <70% aa identity and similarity); the genomic environment of this gene does not provide a clue as to its function. The hydratase/dehydrogenase fusion gene is not present in autotrophic marine group 1 Crenarchaeota. Rather, separate genes for enoyl-CoA hydratase and 3-hydroxyacyl-CoA dehydrogenase exist that likely catalyze these two reactions; their function needs to be defined.

Beta-ketothiolase. M. sedula, T. neutrophilus, and I. hospitalis contained a beta-ketothiolase gene (Msed_0656, Tneu_0429, and Igi_1401, respectively) that was highly similar to the eubacterial enzyme (~36 and ~53% aa identity and similarity). We overexpressed the Thermoproteus gene Tneu_0429 (expected protein molecular mass, 42 kDa) as soluble, N-terminal His$^{10}$-tagged protein in E. coli. The purified recombinant enzyme catalyzed the expected thiolic cleavage of acetoacetyl-CoA into two molecules of acetyl-CoA. Its $V_{max}$ (85°C, pH 7.8) was 55 U mg$^{-1}$, and the apparent $K_m$ for acetoacetyl-CoA was 0.15 mM. In M. sedula, Msed_0656 exhibited upregulation under autotrophic conditions (4). Therefore, these similar genes most likely code for beta-ketothiolase, catalyzing the last step in the two carbon fixation cycles. In marine Crenarchaeota (N. maritimus, C. symbiosum), no highly similar gene was found.

Methylmalonyl-CoA mutase. An intermediate step in the hydroxypropionate/hydroxybutyrate cycle, but not in the dicarboxylic/hydroxybutyrate cycle, is the conversion of methylmalonyl-CoA to succinyl-CoA, which is catalyzed by methylmalonyl-CoA epimerase and methylmalonyl-CoA mutase (MCM). The majority of MCMs in bacteria are heterodimers of a large and a small subunit, in which the small subunit binds the substrate and the cofactor. M. sedula as well as the other autotrophic members of the Sulfolobales clearly encode MCMs in their genome (e.g., small subunit Msed_2055 and large subunit Msed_0638). Interestingly, Msed_2056 (ArgK-like) overlaps with the ORF encoding the small MCM subunit (cobalamin B12-binding domain protein Msed_2055), indicating the potential for cotranscription (4); the function of this gene is unknown. Banerjee and associates (16) claimed that our annotation of MCMs in Crenarchaeota was erroneous, and that instead MCM-like proteins may actually represent isobutyryl-CoA mutases. This is clearly not true for Sulfolobales, and in marine Crenarchaeota (N. maritimus, C. symbiosum) the apparent MCM candidate genes also were found in addition to the epimerase gene. In contrast, the anaerobic I. hospitalis and T. neutrophilus lack any MCM or MCM-like genes: methylmalonyl-CoA is not an intermediate in the dicarboxylic/hydroxybutyrate cycle, and therefore MCM is not required.

Fumarate reductase. Autotrophic carbon fixation in M. sedula requires succinate dehydrogenase rather than fumarate reductase; it was membrane bound and in vitro was coupled to the reduction of dichlorophenindophenol (8). This enzyme is required to convert succinate to malate and oxaloacetate, from which intermediates most metabolites of the central carbon metabolism are derived (19). In contrast, T. neutrophilus and I. hospitalis both contained soluble fumarate reductases that acted with reduced viologen dyes as the electron donor but not with NAD(P)H. The autotrophy-related gene cluster of T. neutrophilus contained the putative genes for two subunits of fumarate reductase (Tneu_0423, Tneu_0424), which were induced in autotrophically grown cells (42). The search of the I. hospitalis genome resulted in two genes, Igi_0276 and Igi_0445 (24). Recently, a soluble, five-component, NADH-dependent fumarate reductase was identified in Hydrogenobacter thermophilus (38) that operates in the reductive citric acid cycle. The fumarate reductase subunit A (HTH_1421) shows high similarity to Tneu_0423 and Igi_0276 (43 and 60% and 50 and 67% aa identity/similarity), whereas the subunit B (HTH_9983) has lower similarity to Tneu_0424 and Igi_0445 (37 and 56% and 29 and 50% aa identity and similarity, respectively).

Transcriptomic evidence in M. sedula. A comprehensive transcriptome analysis of autotrophically, mixotrophically, and heterotrophically grown M. sedula cells by Auernik and Kelly provided valuable supporting information (4). Some of their results were discussed already in the previous sections. Several of the known inorganic carbon fixation cycle enzyme transcripts were upregulated under autotrophic conditions compared to expression under heterotrophic conditions. Two additional ORFs merit further autotrophy-related investigation include Msed_2087, annotated as encoding an acyl-CoA dehydrogenase domain-containing protein (upregulated 8-fold during both autotrophy versus heterotrophy and autotrophy versus mixotrophy), and Msed_1994 (COG 3435; upregulated 4-fold during autotrophy versus heterotrophy), located on the strand opposite the malonic semialdehyde dehydrogenase gene.

Proteomic and genomic evidence in T. neutrophilus. In T. neutrophilus, the characteristic enzyme activities of the carbon fixation cycle, fumarate hydratase, fumarate reductase, succinyl-CoA synthetase, and enzymes catalyzing the conversion of succinyl-CoA to acetyl-CoA, were differentially downregulated in the presence of acetate and, to a lesser extent, in the presence of other organic substrates (42). This regulation pattern correlated well with the differential expression profile of the proteome and with the transcription of the encoding genes. Autotrophy-induced enzymes included succinyl-CoA synthetase (two subunits, Tneu_1463 and Tneu_1464), 4-hydroxybutyrate-CoA ligase (Tneu_0420), 4-hydroxybutyryl-CoA dehydratase (Tneu_0422), and the fusion protein crotonyl-CoA hydratase (S)-3-hydroxybutyl-CoA dehydrogenase (Tneu_0541). The genes encoding PEP carboxylase, fumarate reductase, and four enzymes catalyzing the conversion of succinyl-CoA to crotonyl-CoA are clustered. These findings strongly corroborate some of our present functional annotations of enzymesgenes
<table>
<thead>
<tr>
<th>Classification and no.</th>
<th>Reaction catalyzed by the individual enzyme and equilibrium constant [kinetic data for purified enzymes]</th>
<th>$\Delta G^{\circ}$</th>
<th>Gene(s) in <em>M. sedula</em></th>
<th>Gene(s) in <em>T. neutrophilus</em></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Succinyl-SCoA formation in <em>M. sedula</em></strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Acetyl-SCoA $+$ HCO$_3^{-}$ $+$ ATP $\rightarrow$ malonyl-SCoA$^{-}$ $+$ H$<em>2$O $+$ ADP $+$ P$<em>i$ ($K</em>{eq} = 280$ M; $v</em>{max} = 3.2$ U mg$^{-1}$ (65°C); apparent $K_m = 60$ μM acetyl-SCoA, 40 μM ATP, 300 μM HCO$<em>3^{-}$; $pH</em>{opt} = 7.5$)</td>
<td>$-14$</td>
<td>Msed_0148, Msed_0147</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Malonyl-SCoA$^{-}$ $+$ NADPH $+$ H$^+$ $\rightarrow$ malonic semialdehyde$^{-}$ $+$ NAD$^+$ $+$ CoASH ($K_{eq} = 0.02$; $v_{max} = 44$ U mg$^{-1}$ (65°C); apparent $K_m = 40$ μM malonyl-SCoA, 25 μM NADPH; $pH_{opt} = 7.2$)</td>
<td>$+10$</td>
<td>Msed_0709$^a$</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Malonic semialdehyde$^{-}$ $+$ NADPH $+$ H$^+$ $\rightarrow$ 3-hydroxypropionate$^{-}$ $+$ NAD$^+$ ($K_{eq} = 1.6 \times 10^5$ M$^{-1}$; $v_{max} = 200$ U mg$^{-1}$ (65°C); apparent $K_m = 70$ μM malonic semialdehyde, 70 μM NADPH; $pH_{opt} = 6.7$)</td>
<td>$-24$</td>
<td>Msed_1993$^b$</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>3-Hydroxypropionate$^{-}$ $+$ H$^+$ $+$ CoASH $+$ ATP $\rightarrow$ 3-hydroxypropionyl-SCoA $+$ H$<em>2$O $+$ AMP $+$ PP ($K</em>{eq} = 11$; $v_{max} = 18$ U mg$^{-1}$ (65°C); apparent $K_m = 180$ μM 3-hydroxypropionyl-SCoA, 45 μM ATP (CoASH not analyzed); pH 8.4)</td>
<td>$-6$</td>
<td>Msed_1456$^{a,b}$</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>3-Hydroxypropionyl-SCoA $+$ acryloyl-SCoA $+$ H$<em>2$O ($K</em>{eq} = 7$ M; $v_{max} = 151$ U mg$^{-1}$ (65°C); apparent $K_m = 60$ μM 3-hydroxypropionyl-SCoA; $pH_{opt} = 8.1$)</td>
<td>$-5$</td>
<td>Msed_2001$^{a,b}$</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Acryloyl-SCoA $+$ NADPH $+$ H$^+$ $\rightarrow$ propionyl-SCoA $+$ NAD$^+$ ($K_{eq} = 6 \times 10^6$ M$^{-1}$; $v_{max} = 18.7$ U mg$^{-1}$ (65°C); apparent $K_m = 3$ μM acryloyl-SCoA, 36 μM NADPH; $pH_{opt} = 6.0$)</td>
<td>$-50$</td>
<td>Msed_1426$^{a,b}$</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Propionyl-SCoA $+$ HCO$_3^{-}$ $+$ ATP $\rightarrow$ (S)-methylmalonyl-SCoA$^{-}$ $+$ H$<em>2$O $+$ ADP $+$ P$<em>i$ ($K</em>{eq} = 120$ M; $v</em>{max} = 3.3$ U mg$^{-1}$ (65°C); apparent $K_m = 70$ μM propionyl-SCoA, 40 μM ATP, 300 μM HCO$<em>3^{-}$; $pH</em>{opt} = 7.5$)</td>
<td>$-11$</td>
<td>Msed_0147, Msed_0148, Msed_1375$^{a,b}$</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>(S)-Methylmalonyl-SCoA$^{-}$ $\rightarrow$ (R)-methylmalonyl-SCoA$^{-}$ ($K_{eq} = 1$)</td>
<td>$0$</td>
<td>Msed_0638$^{a,b}$</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>(R)-Methylmalonyl-SCoA$^{-}$ $\rightarrow$ Succinyl-SCoA$^{-}$ ($K_{eq} = 20$)</td>
<td>$-7$</td>
<td>Msed_0638/205$^{a,b}$</td>
<td></td>
</tr>
<tr>
<td><strong>Succinyl-SCoA formation in <em>T. neutrophilus</em></strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Acetyl-SCoA $+$ CO$<em>2$ $+$ ferredoxin$</em>{red}^{-}$ $+$ H$^+$ $\rightarrow$ pyruvate$^{-}$ $+$ ferredoxin$<em>{ox}$$^{-}$ $+$ CoASH ($K</em>{eq} = 0.0005$)</td>
<td>$+19$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Pyruvate$^{-}$ $+$ ATP $+$ H$_2$O $\rightarrow$ phosphoenolpyruvate$^{-}$ $+$ AMP $+$ P$<em>i$ ($K</em>{eq} = 120$)</td>
<td>$-12$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Phosphoenolpyruvate$^{-}$ $+$ HCO$_3^{-}$ $\rightarrow$ oxaloacetate$^{2-}$ $+$ P$<em>i$ ($K</em>{eq} = 1.6 \times 10^4$)</td>
<td>$-24$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Oxaloacetate$^{2-}$ $+$ NADH $+$ H$^+$ $\rightarrow$ (5)-malate$^{2-}$ $+$ NAD$^+$ ($K_{eq} = 1.8 \times 10^4$ M$^{-1}$)</td>
<td>$-30$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>(5)-Malate$^{2-}$ $\rightarrow$ fumarate$^{2-}$ $+$ H$<em>2$O ($K</em>{eq} = 0.3$ M$^{-1}$)</td>
<td>$+3$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>Fumarate$^{2-}$ $+$ 2 &quot;H&quot; $\rightarrow$ succinate$^{2-}$ ($K_{eq} = 8 \times 10^{11}$)</td>
<td>$-68$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>Succinate$^{2-}$ $+$ CoASH $+$ H$^+$ $+$ ATP $\rightarrow$ succinyl-SCoA$^{-}$ $+$ H$_2$O $+$ ADP $+$ P$<em>i$ ($K</em>{eq} = 0.3$)</td>
<td>$+3$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Acetyl-SCoA regeneration
from succinyl-SCoA in
M. sedula and T. neutrophilus

10 Succinyl-SCoA + NADPH + H⁺ (65°C); apparent Km 150 μM succinyl-SCoA, 190 μM NADPH; pHopt 7.2; T. neutrophilus vmax 35 μM mg⁻¹ (65°C); apparent Km 80 μM succinyl-SCoA, 160 μM NADPH; pHopt 8]

11 Succinic semialdehyde + NADPH + H⁺ (65°C); apparent Km 1.5 x 10⁻⁵ M⁻¹; M. sedula vmax 700 μM (65°C); apparent Km 52 μM succinic semialdehyde, 6 μM NADPH, 180 μM NADH; pHopt 7.5; T. neutrophilus vmax 580 μM mg⁻¹ (65°C); apparent Km 30 μM succinic semialdehyde, 190 μM NADPH; pH 7.5]

12 4-Hydroxybutyrate + H⁺ + ATP + CoASH →4-hydroxybutyl-SCoA + H₂O + AMP + PP₃ [Keq = 11; T. neutrophilus vmax 1.6 μM mg⁻¹ (85°C); apparent Km 700 μM 4-hydroxybutyrate; pH 7.8]

13 4-Hydroxybutyl-SCoA →crotonyl-SCoA + H₂O [Keq = 3 M]

14 Crotonyl-SCoA + H₂O (S)-3-hydroxybutyl-SCoA [Keq = 3 M⁻¹; M. sedula vmax 38 μM mg⁻¹ (native enzyme)/18 μM mg⁻¹ (recombinant enzyme) (65°C); apparent Km 300 μM crotonyl-SCoA; pHopt 8]

15 (S)-3-Hydroxybutyl-SCoA + NAD⁺ →acetoacetyl-SCoA + NADH + H⁺ [Keq = 0.002 M; M. sedula vmax 13.8 μM mg⁻¹ (recombinant enzyme) (65°C); apparent Km 200 μM (S)-3-hydroxybutyl-SCoA, 250 μM NAD⁺; pHopt 8]

16 Acetoacetyl-SCoA + CoASH →2 acetyl-SCoA [Keq = 5 x 10⁻⁶ M⁻¹; T. neutrophilus vmax 27 μM mg⁻¹ (85°C); apparent Km 150 μM acetoacetyl-SCoA]

The corresponding enzyme was purified and characterized from the same species.

The substrate specificity of the enzymes, encoded by the corresponding putative genes, cannot be determined on the basis of sequence comparison.

The corresponding genes are colocalized with some other genes involved in autotrophic CO₂ fixation.

The corresponding protein was heterologously expressed in E. coli and characterized.

The homologous protein from Sulfolobus tokodaii T. neutrophilus 2 "H" is assumed to be Fd red, where Fd is ferredoxin.

The numbers in column 1 refer to the reaction numbers in Fig. 1. The kinetic data determined in this and previous works were included. The approximate standard free energy change (ΔG°) in kJ reaction⁻¹) was calculated from the standard free energies of formation at 25°C (from reference 54) or from reported equilibrium constants (note that these standard values do not refer to the high temperature conditions of the organisms). Standard redox potentials also were taken from reference 54, and in some cases they were used for calculations of ΔG°. Gases were taken as being in the gaseous state, and all other compounds were in the aqueous dissolved form. The redox potential (E°) of ferredoxin was set at −414 mV, corresponding to that of the hydrogen/H⁺ couple at pH 7. For the sake of simplicity, equations of ATP consuming/forming reactions considered neither the participation of water nor the changes and H⁺ associated with these reactions. Note that [H⁺] in all equilibrium constants (Keq, calculated from ΔG° values referring to pH 7) needs to be at set at 1, since pH 7 ([H⁺] = 10⁻⁷ M) already was taken into consideration in the ΔG° values. Also, in Keq [H₂O] needs to be set at 1 (rather than 55 M), since according to convention for aqueous systems, the free energies of formation on which the ΔG° values are based refer to compounds in the aqueous state (solutions in water). The free energy changes at cellular concentrations of the reactants (ΔG') need to be corrected according to ΔG' = ΔG° + RT lnK, with ΔG' = −RT lnK. The equilibrium of a reaction may be far on the product side in vivo even if the free energy change under standard conditions is positive. For instance, reaction 2 (ΔG° + 10 kJ) has a Keq = [malonic semialdehyde]/[malonyl-SCoA]/[NADH]/[NAD⁺] [H⁺] = approximately 0.05. However, assuming cellular concentrations, for instance, 0.12 mM NADPH, 0.002 mM NAD⁺, and 1.4 mM CoASH ([H⁺] set at 1, since Keq derived from ΔG° already was corrected for pH 7), the resulting ratio of [malonic semialdehyde]/[malonyl-SCoA]/[NADH]/[NAD⁺]/[H⁺] at equilibrium (ΔG' = 0) under in vivo conditions would be around 700:1. pHopt, optimal pH; SCoA, coenzyme A with sulfur; CoASH, coenzyme A (indicating thiol group).
involved in the autotrophic carbon fixation cycle in \textit{T. neothrippus}.

**Status report on the autotrophic pathways in Crenarchaeota.** Most genes involved in the autotrophic carbon fixation cycles in \textit{M. sedula} and \textit{T. neothrippus} have been identified, and the less common enzymes have been studied (Table 1). Also, the thermodynamics of the individual steps can be determined fairly well. This provides a basis for future pathway modeling. Most autotrophy-related enzymes in autotrophic marine and soil \textit{Crenarchaeota} remain to be shown. In cases of anaerobes like \textit{T. neothrippus}, the generation of low-potential, reduced ferredoxin is unsolved, as is the case for most anaerobes that use pyruvate synthase for assimilation purposes. Finally, a dehydrogenase is unsolved, as is the case for most anaerobes that like soil 2,3-dicarboxylic acids from \textit{C}_{4}-units involving crotonyl-CoA carboxylase/reductase: the ethylmalonyl-CoA pathway. Proc. Natl. Acad. Sci. U. S. A. 104:10631–10636.


