

“Candidatus Galacturonibacter soehngeni” Shows Acetogenic Catabolism of Galacturonic Acid but Lacks a Canonical Carbon Monoxide Dehydrogenase/Acetyl-CoA Synthase Complex

Valk, Laura C.; Diender, Martijn; Stouten, Gerben R.; Petersen, Jette F.; Nielsen, Per H.; Dueholm, Morten S.; Pronk, Jack T.; van Loosdrecht, Mark C.M.

DOI

[10.3389/fmicb.2020.00063](https://doi.org/10.3389/fmicb.2020.00063)

Publication date

2020

Document Version

Final published version

Published in

Frontiers in Microbiology

Citation (APA)

Valk, L. C., Diender, M., Stouten, G. R., Petersen, J. F., Nielsen, P. H., Dueholm, M. S., Pronk, J. T., & van Loosdrecht, M. C. M. (2020). “Candidatus Galacturonibacter soehngeni” Shows Acetogenic Catabolism of Galacturonic Acid but Lacks a Canonical Carbon Monoxide Dehydrogenase/Acetyl-CoA Synthase Complex. *Frontiers in Microbiology*, 11, Article 63. <https://doi.org/10.3389/fmicb.2020.00063>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.



“*Candidatus Galacturonibacter soehngeni*” Shows Acetogenic Catabolism of Galacturonic Acid but Lacks a Canonical Carbon Monoxide Dehydrogenase/Acetyl-CoA Synthase Complex

Laura C. Valk^{1*}, Martijn Diender², Gerben R. Stouten¹, Jette F. Petersen³, Per H. Nielsen³, Morten S. Dueholm³, Jack T. Pronk¹ and Mark C. M. van Loosdrecht¹

¹ Department of Biotechnology, Delft University of Technology, Delft, Netherlands, ² Laboratory of Microbiology, Wageningen University and Research, Wageningen, Netherlands, ³ Department of Chemistry and Bioscience, Centre for Microbial Communities, Aalborg University, Aalborg, Denmark

OPEN ACCESS

Edited by:

Mirko Basen,
University of Rostock, Germany

Reviewed by:

Daan R. Speth,
California Institute of Technology,
United States
Emma Jayne Gagen,
University of Queensland, Australia

*Correspondence:

Laura C. Valk
l.c.valk@tudelft.nl;
lcv@bio.aau.dk

Specialty section:

This article was submitted to
Microbial Physiology and Metabolism,
a section of the journal
Frontiers in Microbiology

Received: 26 September 2019

Accepted: 13 January 2020

Published: 29 January 2020

Citation:

Valk LC, Diender M, Stouten GR, Petersen JF, Nielsen PH, Dueholm MS, Pronk JT and van Loosdrecht MCM (2020) “*Candidatus Galacturonibacter soehngeni*” Shows Acetogenic Catabolism of Galacturonic Acid but Lacks a Canonical Carbon Monoxide Dehydrogenase/Acetyl-CoA Synthase Complex. *Front. Microbiol.* 11:63. doi: 10.3389/fmicb.2020.00063

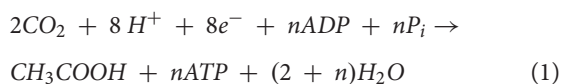
Acetogens have the ability to fixate carbon during fermentation by employing the Wood-Ljungdahl pathway (WLP), which is highly conserved across Bacteria and Archaea. In a previous study, product stoichiometries in galacturonate-limited, anaerobic enrichment cultures of “*Candidatus Galacturonibacter soehngeni*,” from a novel genus within the *Lachnospiraceae*, suggested the simultaneous operation of a modified Entner-Doudoroff pathway for galacturonate fermentation and a WLP for acetogenesis. However, a draft metagenome-assembled genome (MAG) based on short reads did not reveal homologs of genes encoding a canonical WLP carbon-monoxide-dehydrogenase/acetyl-Coenzyme A synthase (CODH/ACS) complex. In this study, NaH¹³CO₃ fed to chemostat-grown, galacturonate-limited enrichment cultures of “*Ca. G. soehngeni*” was shown to be incorporated into acetate. Preferential labeling of the carboxyl group of acetate was consistent with acetogenesis via a WLP in which the methyl group of acetate was predominately derived from formate. This interpretation was further supported by high transcript levels of a putative pyruvate-formate lyase gene and very low transcript levels of a candidate gene for formate dehydrogenase. Reassembly of the “*Ca. G. soehngeni*” MAG with support from long-read nanopore sequencing data produced a single-scaffold MAG, which confirmed the absence of canonical CODH/ACS-complex genes homologs. However, high CO-dehydrogenase activities were measured in cell extracts of “*Ca. G. soehngeni*” enrichment cultures, contradicting the absence of corresponding homologs in the MAG. Based on the highly conserved amino-acid motif associated with anaerobic Ni-CO dehydrogenase proteins, a novel candidate was identified which could be responsible for the observed activities. These results demonstrate operation of an acetogenic pathway, most probably as a yet unresolved variant of the Wood-Ljungdahl pathway, in anaerobic, galacturonate-limited cultures of “*Ca. G. soehngeni*.”

Keywords: acetogenesis, ¹³C-labeling, meta-transcriptomics, chemostat enrichment culture, Wood-Ljungdahl pathway

INTRODUCTION

Over the course of multiple decades, seven carbon-fixing pathways capable of supporting autotrophic growth have been identified and intensively studied; the Calvin-Benson-Bassham (CCB) reductive pentose-phosphate cycle, the reductive citric-acid cycle (Arnon-Buchanan (AB) cycle), the hydroxypropionate (Fuchs-Holo) bi-cycle, the 3-hydroxypropionate/4-hydroxybutyrate cycle, dicarboxylate/hydroxybutyrate cycle, the reductive acetyl-CoA (Wood-Ljungdahl) pathway and the reductive glycine pathway (Berg, 2011; Fuchs, 2011; Figueroa et al., 2018). The first five pathways are primarily used for carbon fixation and the reductive glycine pathway for recycling of electron carriers. Only the Wood-Ljungdahl pathway (WLP) also acts as a primary pathway for energy conservation in anaerobes (Fuchs, 2011; Bar-Even et al., 2012b; Schuchmann and Müller, 2014).

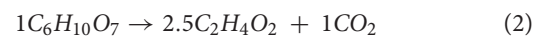
The WLP is highly conserved across Archaea and Bacteria, with only two known variations, one found predominantly in methanogenic archaea and one in acetogenic bacteria. The first has formyl-methanofuran rather than formate as first intermediate, and uses ATP-independent formyl-MFR:tetrahydromethanopterin formyltransferase instead of ATP-consuming formyl-tetrahydrofolate ligase (consuming an ATP). Moreover, methanogens use methanofuran (MFR), tetrahydromethanopterin and coenzyme-F₄₂₀ as cofactors while acetogens rely on NAD(P)H, tetrahydrofolate (THF) and ferredoxin (Fd) (Fuchs, 2011; Adam et al., 2018). Reduction of CO₂ to acetate via the WLP requires 8 electrons (Equation 1, Ragsdale and Pierce, 2008; Schuchmann and Müller, 2014).



The WLP consists of two branches. In acetogens, the WLP methyl branch reduces CO₂ to a methyl group by first reducing CO₂ to formate via formate dehydrogenase (*fdhA*; EC 1.17.1.9), after which formate is bound to tetrahydrofolate (THF) by formate-tetrahydrofolate ligase (*fhs*, EC 6.3.4.3). Formyl-THF is then further reduced to methenyl-THF, methylene-THF and lastly to methyl-THF by formyl-THF cyclohydrolase and methylene-THF dehydrogenase (*folD*; EC 3.5.4.9 and EC 1.5.1.5) and methylene-THF reductase (*metF*, EC 1.5.1.20), respectively (Ragsdale, 2008; Ragsdale and Pierce, 2008). A methyl transferase then transfers the methyl group from THF to a corrinoid iron-sulfur protein (*acsE*, EC 2.3.1.258), which is a subunit of the carbon monoxide (CO) dehydrogenase/acetyl-CoA synthase complex. The carbonyl branch of the WLP reduces CO₂ to CO in a reaction catalyzed by another subunit of the canonical WLP, the CO dehydrogenase/acetyl-CoA synthase complex (CODH/ACS, EC 2.3.1.169). Alternatively, CO can be formed by a separate CO dehydrogenase (CODH, EC 1.2.7.4) (Ragsdale and Kumar, 1996; Doukov et al., 2002; Jeoung and Dobbek, 2011). The CODH/ACS complex then links the two WLP branches by coupling the CO- and CH₃-groups with CoA, yielding acetyl-CoA (Menon and Ragsdale, 1996a; Ragsdale and Kumar, 1996; Ragsdale, 2008). The high degree of conservation of WLP genes and their genomic

co-localization suggests that their evolution involved interspecies gene transfer events (Techtmann et al., 2012; Adam et al., 2018). However, two recent studies suggested carbon fixation occurred in the absence of a full complement of structural genes for canonical WLP enzymes (Figueroa et al., 2018; Valk et al., 2018). These observations suggest that variants of the canonical WLP may still await discovery.

In a recent study on D-galacturonate-limited, anaerobic enrichment cultures, we identified the dominant bacterium as a species from a novel genus within the *Lachnospiraceae*, for which we proposed the name “*Candidatus Galacturonibacter soehngeni*.” The *Lachnospiraceae* family is part of the phylum Firmicutes, which includes several genera that harbor acetogens (Drake et al., 2008; Ragsdale and Pierce, 2008; Schuchmann and Muller, 2013; Valk et al., 2018). Fermentation product stoichiometries of the enrichment cultures were consistent with an acetogenic dissimilation of galacturonate. The overall stoichiometry is shown in Equation (2) (Valk et al., 2018).



Metagenome analysis of the enrichment culture revealed homologs of most structural genes for WLP enzymes, but no homologs were found for genes encoding subunits of the canonical CODH/ACS complex (EC 2.3.1.169) (Valk et al., 2018).

The goal of the present study was to further investigate the presence of a possible alternative configuration of the WLP in “*Ca. G. soehngeni*.” To analyze *in vivo* activity of the WLP, D-galacturonate-limited enrichment cultures were co-fed with ¹³C-labeled bicarbonate, followed by analysis of ¹³C in the methyl and carboxyl groups of acetate. To investigate whether canonical WLP genes might have been overlooked in the initial metagenomics analysis, a fully closed metagenome-assembled genome (MAG) sequence of “*Ca. G. soehngeni*” was constructed using long-read nanopore sequencing, and meta-transcriptome analysis was performed to analyze the expression levels of genes of interest. Additionally, CO dehydrogenase activity was analyzed in cell extracts.

MATERIALS AND METHODS

Reactor Setup and Operation

Chemostat cultures were grown in 1.2 L laboratory bioreactors (Applikon, Delft, The Netherlands), which were stirred at 300 rpm and kept at 30°C. Anaerobic conditions were maintained by flushing the headspace with nitrogen gas, at a flow rate of 120 mL min⁻¹. Culture pH was controlled at 8 ± 0.1 by automatic titration (ADI 1030 Biocontroller, Applikon, Delft, The Netherlands) of 1 M NaOH. The dilution rate was 0.09 ± 0.01 h⁻¹ and the working volume of 0.5 L was kept constant by peristaltic effluent pumps (Masterflex, Cole-Parmer, Vernon Hills, IL, United States) coupled to electrical level sensors. Bioreactors were inoculated (10% v/v) with 50 mL samples of D-galacturonate-limited, anaerobic chemostat enrichment cultures (Valk et al., 2018), stored in 30% v/v glycerol at -20°C. Cultures were run in continuous mode and after at least

6 days (18 generations) stable product composition and biomass concentration were established. System stability was assessed by online monitoring of CO₂ production and offline monitoring of fermentation products and optical density. When measurements varied by less than 10% over multiple volume changes, without a clear upward or downward trend, samples were taken during subsequent cycles.

Medium

The cultivation medium contained (g L⁻¹): D-galacturonate 4.3; NH₄Cl 1.34; KH₂PO₄ 0.78; Na₂SO₄·10H₂O 0.130; MgCl₂·6H₂O 0.120; FeSO₄·7H₂O 0.0031; CaCl₂ 0.0006; H₃BO₄ 0.0001; Na₂MoO₄·2H₂O 0.0001; ZnSO₄·7H₂O 0.0032; CoCl₂·H₂O 0.0006; CuCl₂·2H₂O 0.0022; MnCl₂·4H₂O 0.0025; NiCl₂·6H₂O 0.0005; EDTA 0.10. Nineteen liter of mineral solution (mineral concentration adjusted to the final volume, 20 L) was autoclaved for 20 min at 121°C after which 1 L (86 g L⁻¹) D-galacturonate solution was filter sterilized (0.2 μm Mediakap Plus, Spectrum Laboratories, Rancho Dominguez, CA, United States) into the media. 1.5 mL Pluronic PE 6100 antifoam (BASF, Ludwigshafen, Germany) was added per 20 L of mineral solution to avoid excessive foaming.

Analysis of Substrate and Extracellular Metabolite Concentrations

To determine substrate and extracellular metabolite concentration, reactor sample supernatant was obtained by centrifugation of culture samples (Heraeus Pico Microfuge, Thermo Fisher Scientific, Waltham, MA, United States). Concentrations of D-galacturonate and extracellular metabolites were analyzed with an Agilent 1100 Affinity HPLC (Agilent Technologies, Amstelveen, The Netherlands) equipped with an Aminex HPX-87H ion-exchange column (BioRad, Hercules, CA, United States) operated at 60°C with a mobile phase of 5 mM H₂SO₄ and a flow rate of 0.6 mL min⁻¹. CO₂ and H₂ concentrations in the bioreactor exhaust gas were measured using a Prima BT Bench Top mass spectrometer (Thermo Fisher Scientific, Waltham, MA, United States) after the gas was cooled by a condenser (4°C).

Biomass Dry Weight

Twenty milliliter of culture broth samples were filtered over pre-dried and pre-weighed membrane filters (0.2 μm Supor-200, Pall Corporation, New York, NY, United States), which were then washed with demineralized water, dried in a microwave oven (Robert Bosch GmbH, Gerlingen, Germany) for 20 min at 360 W and reweighed. Carbon and electron balances were constructed

based on the number of carbon atoms and electrons per mole, while biomass composition was assumed to be CH_{1.8}O_{0.5}N_{0.2} (Roels, 1983).

Quantitative Fluorescent *in situ* Hybridization (qFISH) Analysis

Fluorescent *in situ* hybridization was performed as described previously (Daims et al., 2005), using a hybridization buffer containing 35% (v/v) formamide. Probes were synthesized and 5' labeled with either 5(6)-carboxyfluorescein-N-hydroxysuccinimide ester (FLUOS) or with one of the sulfoindocyanine dyes (Cy3 and Cy5; Thermo Hybaid Interactiva, Ulm, Germany) (Table 1). The general probe EUB338mix, labeled at both 3' and 5' ends with Cy5, was used to identify all eubacteria in the sample. Microscopic analysis was performed with a LSM510 Meta laser scanning confocal microscope (Carl Zeiss, Oberkochen, Germany). The qFISH analysis was based on at least 29 fields of view at 6730 × magnification, using DAIME (version 2.1) software (DOME, Vienna, Austria; Daims et al., 2006). The bio-volume fractions of “*Ca. G. soehngenii*” and *Enterobacteriaceae* populations were calculated as the ratio of the area hybridizing with specific probes relative to the total area hybridizing with the universal EUBmix probe set (Amann et al., 1990; Daims et al., 1999).

Labeling Experiment ¹³C-Labeled Sodium Bicarbonate Addition

A 1 M NaH¹³CO₃ solution was used to replace the regular 1 M NaOH solution as a pH titrant in steady-state D-galacturonate-limited enrichment cultures (pH 7.8 ± 0.1, D = 0.1 h⁻¹, T = 30°C). Broth was collected on ice every 2 h for 8 consecutive hours and centrifuged (12,000 × g, Heraeus Pico Microfuge, Thermo Fisher Scientific, Waltham, MA, United States) before the supernatant was collected and stored at -20°C until analysis by NMR. CO₂, H₂ and ¹³CO₂ concentrations in the exhaust gas were measured by MS (Prima BT Bench Top MS, Thermo Fisher Scientific, Waltham, MA, United States) after the gas had been cooled by a condenser (4°C).

Illumina and Nanopore Sequencing, Metagenome Assembly, and Genome Binning DNA

The metagenomic-assembled genome of “*Candidatus Galacturonibacter soehngenii*” described by Valk et al. (2018) was used as template for preparing the metagenome libraries. The DNA extraction, Illumina sequencing, metagenomic assembly and binning process is described

TABLE 1 | Oligonucleotide probes used for the quantitative fluorescence *in situ* hybridization analysis.

Probe	Sequence (5'–3')	Specificity	References
EUB338mix	GCWGCCWCCGTAGGWT	All bacteria	Daims et al., 1999
ENT	CTCTTTGGTCTTGCACG	<i>Enterobacteriaceae</i>	Kempf et al., 2000
Lac87	GTGGCGATGCAAGTCTGA	“ <i>Ca. G. soehngenii</i> ”	This study

With W indicating A or T.

in Valk et al. (2018). Long-read genomic DNA sequencing was conducted using 1D nanopore sequencing (Oxford Nanopore Technologies, Oxford, United Kingdom), following the manufacturer's protocol (LSK-108), omitting the optional DNA shearing and DNA repair steps. The library was loaded on a flow cell (FLO-MIN106) and the MinION Mk1B DNA sequencer (Oxford Nanopore Technologies, Oxford, United Kingdom) was used for sequencing combined with the MinKNOW v. 1.7.3 (Oxford Nanopore Technologies, Oxford, United Kingdom) software with the 48 h sequencing workflow (NC_48h_Sequencing_Run_FLO_MIN106_SQK-LSK108.py). Albacore v. 1.2.1 (Oxford Nanopore Technologies, Oxford, United Kingdom) was used to base-call the sequencing reads.

Genome Assembly

The assembling of the contigs from the “*Candidatus Galacturonibacter soehngenii*” genome bin into a single scaffold based on the long Nanopore reads was done using SSPACE-LongRead scaffolder v. 1.1 (Boetzer and Pirovano, 2014). GapFiller v. 1.11 (Boetzer and Pirovano, 2012) or by manual read mapping and extension in CLC Genomics Workbench v. 9.5.2 (Qiagen, Hilden, Germany) were used to close gaps in the draft genome with the previously assembled Illumina data. Finally, manual polishing of the complete genome was done to remove SNPs and ensure a high-quality assembly. The meta-genome has been submitted to the sequence read archive (SRA)¹ with accession number SRR10674409, under the BioProject ID PRJNA566068.

Genome Annotation and Analysis

The metagenome-assembled genome was uploaded to the automated Microscope platform (Vallenet et al., 2006, 2017). Manual assessment of pathway annotations was assisted by the MicroCyc (Caspi et al., 2008), KEGG (Kyoto Encyclopedia of Genes and Genomes; Kanehisa et al., 2014) and SwissProt alignment (BLASTP version 2.2.28+; Altschul et al., 1997) databases. The predicted proteome of “*Ca. G. soehngenii*” was submitted to InterProScan (version 5.25-64.0), to identify predictive Pfam domains (El-Gebali et al., 2018). The annotated genome sequence of “*Candidatus Galacturonibacter soehngenii*” has been submitted to the European Nucleotide Archive (ENA) under the BioProject ID PRJNA566068.

Genome-Centric Meta-Transcriptomic Analyses; RNA Extraction and Purification

During pseudo-steady state, broth samples were taken from the enrichment culture, directly frozen in liquid nitrogen and subsequently stored at -80°C . Five hundred microliter samples were thawed on ice, pelleted by centrifugation ($21,000 \times g$, 2 min, 4°C) and used for total RNA extraction with the RNeasy PowerMicrobiome Kit (Qiagen, Hilden, Germany), following the manufacturer's instruction with the addition of phenol:chloroform:isoamyl alcohol (25:25:1) and

β -mercaptoethanol ($10 \mu\text{L mL}^{-1}$ final concentration). Cell lysis was with a FastPrep-24 bead beater (MP Biomedicals, Fisher Scientific, Hampton, VA, United States, four successive cycles of 40 s at 6.0 m s^{-1} , 2 min incubation on ice between cycles). Total RNA extracts were subjected to DNase treatment to remove DNA contaminants by using the DNase Max Kit (Qiagen, Hilden, Germany) and further cleaned up and concentrated with the Agencourt AMPure XP magnetic beads (Beckman Coulter, Brea, CA, United States) before rRNA depletion. Integrity and quality of purified total RNA were assessed on a TapeStation 2200 (Agilent, Santa Clara, CA, United States) with the Agilent RNA screen-tapes (Agilent, Santa Clara, CA, United States) and the concentration was measured using Qubit RNA HS Assay Kit (Thermo Scientific Fisher, Waltham, MA, United States).

rRNA Depletion, Library Preparation, and Sequencing

Five hundred nanogram of total RNA from each sample was obtained after rRNA was depleted using the Ribo-Zero rRNA Removal (Bacteria) Kit (Illumina, San Diego, CA, United States), with 2 μg total RNA as input. Quality of extracted mRNA was checked with Agilent RNA HS screen-tapes (Agilent, Santa Clara, CA, United States) and RNA concentration was determined with a Qubit RNA HS Assay Kit (Thermo Scientific Fisher, Waltham, MA, United States). The TruSeq Stranded mRNA Sample Preparation Kit (Illumina, San Diego, CA, United States) was used to prepare cDNA sequencing libraries according to the manufacturer's instruction. Libraries were sequenced on an Illumina HiSeq2500 using the TruSeq PE Cluster Kit v3-cBot-HS and TruSeq SBS kit v3-HS sequencing kit ($1 \times 50 \text{ bp}$; Illumina, San Diego, CA, United States). The raw meta-transcriptome reads have been submitted to the sequence read archive (SRA)¹ with accession number SRR10674118-23, under the BioProject ID PRJNA566068.

Trimming and Mapping of rRNA Reads

Raw RNA reads in FASTQ format were imported into CLC Genomics Workbench v. 9.5.5 and trimmed for quality, requiring a minimum phred score of 20 and a read length of 45. Reads from each sample were hereafter mapped to CDSs obtained from the MAG of “*Ca. G. soehngenii*” with a minimum similarity of 98% over 80% of the read length. Reads per kilobase of transcript per million mapped reads (RPKM) were calculated based on raw read-counts and the length of each CDS. The meta-transcriptome mapped to the genome of “*Ca. G. soehngenii*” are shown in **Supplementary Data Sheet S2**.

Plasmid and Strain Construction

Gene F7O84_RS11645 was codon optimized for expression in *Escherichia coli* with the GeneArt online tool and integrated behind the TEV recognition site of the pET151/D-TOPO expression vector by GeneArt (GeneArt GmbH, Regensburg, Germany). The resulting plasmid was transformed into a chemically competent *E. coli* strain BL21 according to manufacturer's instructions (NEBuilder HiFi DNA Assembly Master Mix chemical transformation protocol (E2621),

¹<https://www.ncbi.nlm.nih.gov/sra/>

New England Biolabs, Ipswich, MA, United States) and named pUD1074. The plasmid sequence of pUD1074 has been deposited at the NCBI GenBank² with the corresponding accession number MN498128.

Heterologous Expression of the Putative CO Dehydrogenase Candidate

All *E. coli* cultures were performed in 120 mL capped bottles with 50 mL of mineral medium (Diender et al., 2016). Prior to inoculation, the bottles were autoclaved at 120°C after which the mineral media was supplemented with autoclaved (120°C, 20 min); glucose 5 g L⁻¹, peptone (BD Bacto Difco, Thermo Fisher Scientific, Waltham, MA, United States) 1 g L⁻¹, yeast extract (BD Bacto Difco, Thermo Fisher Scientific, Waltham, MA, United States) 2 g L⁻¹ and cysteine 1 g L⁻¹. Additionally, 0.05 g L⁻¹ ampicillin was added and the gas phase was exchanged with air, with a final pressure of 170 kPa. All *E. coli* cultures used for measurements were inoculated with overnight grown pre-cultures (1:50 v/v) and incubated at 37°C and shaken (300 rpm) until oxygen was depleted (2–3 h). Subsequently 1 mL (250 g L⁻¹) glucose, 1 mL reducing agent (0.4 M cysteine) and 1 mL IPTG (40 mM) were added.

After 3 h (at 30°C, unshaken) of incubation, the cells were harvested and processed anaerobically according to Diender et al. (2016). Enzymatic activity analysis was conducted using a modified method initially described by Diender et al. (2016). The essays were performed in an anaerobic environment using 100–300 µL of cell extract with both CO and hydroxylamine as substrate. To increase metal cofactor availability, 1:200 (v/v) metals solution was added to the assay buffer which contained in (g L⁻¹); HCl 1.8, H₃BO₃ 0.0618, MnCl₂ 0.06125, FeCl₂ 0.9435, CoCl₂ 0.0645, NiCl₂ 0.01286, ZnCl₂ 0.0677, CuCl₂ 0.01335.

Homology Protein BLAST Analysis

The sequence of the putative CODH (F7O84_RS11645) was blasted with the BLASTp (version 2.2.28+; Altschul et al., 1997) tool of the JGI-IMG/M database (Markowitz et al., 2012), with default parameter settings. Finished genomes from members of the *Lachnospiraceae* family in the public JGI-IMG/M database (Markowitz et al., 2012) were selected for analysis, **Supplementary Table S4**. The stains identified in the BLAST search, or closely related strains (**Supplementary Table S5**) were subsequently analyzed in KEGG (Kanehisa et al., 2014) for presences of the CODH/ACS complex with pathway map 1200.

RESULTS

Physiological Characterization of D-Galacturonate-Limited Enrichment Cultures Dominated by “*Ca. G. soehngenii*”

Anaerobic, galacturonate-limited chemostat enrichment cultures were used to study the physiology of “*Ca. G. soehngenii*”

cultures. In a previous study (Valk et al., 2018), the relative abundance of “*Ca. G. soehngenii*” in such cultures did not exceed 65%, based on metagenomic analysis, and formate and H₂ were detected in the liquid and gas phases, respectively. It was hypothesized that, in these experiments, a low *in situ* hydrogen partial pressure limited *in vivo* WLP activity, as it was expected that hydrogen was used as reductant for the production of acetate from formate or CO₂. To investigate this possibility, head space flushing instead of sparging was applied, using N₂ gas. This caused an increase in the hydrogen partial pressure in the media broth (De Kok et al., 2013). Additionally, the dilution rate was decreased from 0.125 to 0.1 h⁻¹. Analysis of the abundance of “*Ca. G. soehngenii*” in the resulting enrichment cultures by quantitative fluorescence *in situ* hybridization (qFISH) indicated that 86.5 ± 2.6% of the bio-volume of qFISH-detectable cells consisted of “*Ca. G. soehngenii*.” The major side population *Enterobacteriaceae* represented 13.8 ± 2.4% of the bio-volume. As these two subpopulations together accounted for 100.2 ± 5.0% of the bio-volume, it was assumed that any other, minor, subpopulations did not significantly influence the stoichiometry of catabolic fluxes.

Product yields and biomass-specific conversion rates of the D-galacturonate-limited anaerobic enrichment cultures dominated by “*Ca. G. soehngenii*” (**Table 2**) showed acetate as dominant catabolic product (0.57 ± 0.03 Cmol (Cmol galacturonate⁻¹)). Carbon and electron recoveries were 94 and 92%, respectively, indicating that all major fermentation products were identified. As observed previously (Valk et al., 2018), this acetate yield on galacturonic acid was significantly higher than the combined yields of formate and hydrogen. This difference was interpreted as indicative for acetogenesis by one of the dominant organisms, of which only the “*Ca. G. soehngenii*” MAG was shown to harbor homologs for most WLP structural genes (Ragsdale and Pierce, 2008; Valk et al., 2018). Yields of hydrogen

TABLE 2 | Yields (in Cmol (Cmol galacturonate⁻¹), unless stated otherwise) and biomass-specific conversion rates (q; mmol g_x⁻¹ h⁻¹) of anaerobic, galacturonate-limited chemostat enrichment cultures dominated by “*Ca. Galacturonibacter soehngenii*.”

	Yield (Cmol; Cmol _s ⁻¹)	Biomass specific conversion rates (mmol (g _x) ⁻¹ h ⁻¹)
D-galacturonate	–	–4.0 ± 0.2
Biomass	0.17 ± 0.02	–
Acetate	0.57 ± 0.03	6.9 ± 0.4
Formate	0.02 ± 0.01	0.4 ± 0.2
CO ₂	0.18 ± 0.02	4.3 ± 0.3
H ₂ (mol Cmol ⁻¹)	0.02 ± 0.01	0.2 ± 0.1
H ₂ + Formate (mol Cmol _s ⁻¹)	0.04 ± 0.02	
Acetyl-CoA derivatives (mol Cmol _s ⁻¹)	0.29 ± 0.02	

Chemostat cultures were operated at dilution rate of 0.1 h⁻¹, pH, 8 and at 30°C, with galacturonate the sole carbon-source. Data are presented as average ± mean deviations, derived from nine measurements each on duplicate steady-state enrichment cultures.

²www.ncbi.nlm.nih.gov/genbank

and formate on galacturonate (0.02 ± 0.01 mol Cmol galacturonate⁻¹) and 0.02 ± 0.01 (Cmol galacturonate⁻¹), respectively were significantly lower than found in a previous study on “*Ca. G. soehngeni*” (Valk et al., 2018). This observation is consistent with a higher *in vivo* contribution of the WLP as a result of a higher hydrogen partial pressure and/or lower specific growth rate in the present study.

Incorporation of ¹³C-Labeled Bicarbonate Into Acetate Corroborates Acetogenic Fermentation

A simple model was constructed to predict formation of labeled acetate, using biomass-specific conversion rates measured in pseudo-steady state enrichment cultures as inputs (Supplementary Calculations S1, S2 and Supplementary Figure S1). Model simulations predicted that, after 8 h, approximately 15% of the acetate produced by the enrichment culture should be labeled. To investigate if CO₂ was indeed incorporated into acetate via acetogenic fermentation, ¹³C-labeled bicarbonate was fed to a “*Ca. G. soehngeni*” enrichment chemostat culture. However, after 8 h, the fraction of ¹³C in the methyl group of acetate increased to 2.0%. This increase represented only a small increase relative to the 1% natural abundance of ¹³C (Table 3; Rumble et al., 2017). In contrast, after 8 h of ¹³C-bicarbonate feeding, the enrichment culture showed a 21.5% abundance of ¹³C in the carbonyl-group of acetate (Table 3).

Significant Activity of CO Dehydrogenase in Cell Extracts of “*Ca. G. soehngeni*” Enrichment Cultures

In the WLP, ¹³C-labeled CO₂ incorporation into the carbonyl-group of acetate involves activity of CO dehydrogenase (COOS, EC 1.2.7.4). To investigate the presence of this key enzyme in “*Ca. G. soehngeni*,” an anaerobic enzyme activity assay was performed on cell extracts of enrichment cultures, using CO as electron donor and methyl viologen (MV) as electron acceptor (Diender et al., 2016). These assays revealed a CO dehydrogenase activity of $2.1 \pm 0.6 \mu\text{mol min}^{-1}$ (mg protein)⁻¹. Reduction of MV in the absence of either CO or cell extract was below detection limit [$<0.05 \mu\text{mol min}^{-1}$ (mg protein)⁻¹].

TABLE 3 | Percentages of ¹³C-labeled methyl and carbonyl groups in total-culture acetate, calculated from proton and carbon NMR spectra.

	Time (h)	% ¹³ C
Methyl (CH ₃)	0	1.0
	4	1.6
	8	2.0
Carbonyl (CO)	8	21.8

Samples were taken from the “*Ca. G. soehngeni*” chemostat enrichment cultures in bioreactor 2 after switching the alkali supply line from 1 M NaOH to 1 M NaH¹³CO₃ (Time = 0 h).

Identification of Two Putative Novel CO Dehydrogenase Genes in a Newly Obtained Single-Scaffold MAG of “*Ca. G. soehngeni*”

Previous analysis of the “*Ca. G. soehngeni*” MAG (Valk et al., 2018) was based on an assembly made with short-read DNA sequencing data. To identify if putative CODH/ACS complex genes had been missed in this analysis due to incomplete assembly, long-read Oxford Nanopore sequencing (Deamer et al., 2016; Jain et al., 2016) was used to improve the previously assembled “*Ca. G. soehngeni*” MAG. The resulting genome assembly consisted of 8 contigs and was estimated to have a 98% completeness and contained no genetic contamination with sequences from other organisms according to checkM (Table 4). As in the previous study, homologs were detected for most structural genes associated with the WLP (Table 5), but none of the annotated genes in the predicted proteome showed homology with known CODH/ACS genes (Vallenet et al., 2006; Ragsdale, 2008; Valk et al., 2018). A search in the newly assembled “*Ca. G. soehngeni*” MAG sequence for homologs of signature genes of the six other known pathways for inorganic carbon fixation did not point toward their involvement in carbon metabolism (Supplementary Table S2).

CO dehydrogenases contain highly conserved amino-acid motifs (Pfam or protein-family domains) associated with their nickel-iron-sulfur clusters (Eggen et al., 1991, 1996; Maupin-Furlow and Ferry, 1996; Jeoung and Dobbek, 2011; Techtmann et al., 2012; El-Gebali et al., 2018). The newly assembled “*Ca. G. soehngeni*” MAG sequence did not reveal hits for the Pfam domain of the CO dehydrogenase α-subunit of the CODH/ACS complex (PF18537) (Darnault et al., 2003). However, two open reading frames F7O84_RS02405 and F7O84_RS11645, harbored the PF03063 Pfam domain, which is associated with the hybrid cluster protein (HCP) and the catalytic center of the Ni-CODH family (van den Berg et al., 2000; Wolfe et al., 2002). Although HCP has been associated with hydroxylamine reductase activity, its catalytic activity has

TABLE 4 | Statistics of the metagenome-assembled genome (MAG) of “*Ca. Galacturonibacter soehngeni*.”

“ <i>Candidatus Galacturonibacter soehngeni</i> ”	
Genome size (Mbp)	4.1
Scaffolds	1
Contigs	8
Contigs N50	1033779
Max contig size	1514059
Completeness (%)	98
Contamination (%)	0
GC content (%)	34.4
Protein coding density (%)	89
CDS	3924
rRNA copies	5

Completeness and contamination were estimated with CheckM (Parks et al., 2015).

TABLE 5 | Genes of the Wood-Ljungdahl pathway from the predictive proteome of the MAG “*Ca. G. soehngeni*” with gene names, EC number, gene or homolog and *E*-value based on SwissProt alignment (BLASTP version 2.2.28+, MicroScope platform v3.13.2).

Encoded protein	EC	Gene name	<i>E</i> -value	Gene ID
Formate dehydrogenase	1.17.1.9	<i>fdhA</i>	1 e ⁻⁶⁰	F7O84_RS07405
Formate-tetrahydrofolate ligase	6.3.4.3	<i>fhs</i>	0.0	F7O84_RS05385
Methenyl-tetrahydrofolate cyclohydrolase/methylene-tetrahydrofolate dehydrogenase	3.5.4.9 and 1.5.1.5	<i>folD</i>	5 e ⁻¹⁵²	F7O84_RS05380
Methyl-tetrahydrofolate reductase	1.5.1.20	<i>metF</i>	1 e ⁻⁸⁷	F7O84_RS08335
5-Methyl-tetrahydrofolate:corrinoid/iron-sulfur protein methyltransferase	2.1.1.258	<i>acsE</i>	5 e ⁻³⁷	F7O84_RS02745
CO-Methylating acetyl-CoA synthase	2.3.1.169	<i>acsBCD</i>	>10	
Carbon-monoxide dehydrogenase	1.2.7.4	<i>cooS</i>	>10	

not been experimentally confirmed and, moreover, sequence motifs in HCP showed high similarity with the functional domain of Ni-CODHs making it an interesting candidate genes for the CODH function of the WLP in “*Ca. G. soehngeni*” (Heo et al., 2002; Wolfe et al., 2002; Aragão et al., 2003; Almeida et al., 2006). A closer inspection of the genetic context of both genes showed many flanking genes encoding hypothetical proteins in their close vicinity, but no genes previously associated with acetogenesis.

Homologs of Acetogenesis Genes Are Transcribed in D-Galacturonate-Limited “*Ca. G. soehngeni*” Enrichment Cultures

A meta-transcriptome analysis of the enrichment cultures showed significant transcript levels of most homologs of known WLP genes, which were approximately 10-fold lower than those of homologs of structural genes encoding Entner-Doudoroff-pathway enzymes involved in galacturonate catabolism (Table 6). A notable exception was the extremely low transcript level of a putative formate dehydrogenase gene (F7O84_RS07405; EC 1.17.1.9). A candidate gene for pyruvate-formate lyase (PFL, EC 6.2.1.3) was highly transcribed (F7O84_03160, Table 6). These observations suggested that formate generated by PFL, rather than CO₂, was the major substrate for the methyl branch of the WLP in “*Ca. G. soehngeni*.”

Homologs of Rnf cluster (F7O84_03275-3295; EC 7.2.1.2) and hydrogenase (F7O84_0945-50, F7O84_04820; EC 1.12.7.2) genes, which were previously implicated in acetogenesis (Biegel and Müller, 2010; Schuchmann and Müller, 2014, 2016), showed high transcript levels (Table 6). Of the two candidate genes for CO dehydrogenase, F7O84_RS11645 showed the highest transcript level (Table 6). As, under the experimental conditions, no hydroxylamine reductase activity was expected, this result reinforced the candidature of F7O84_RS11645 as possible CO dehydrogenase gene. In an attempt to directly investigate if F7O84_RS11645 encoded a functional CO dehydrogenase, its open reading frame was cloned into high-copy-number *E. coli* expression vector. However, enzyme assays with cell extracts of the resulting *E. coli* strain did not yield consistent evidence for either CO dehydrogenase or hydroxylamine dehydrogenase activity (Supplementary Table S3).

Identification of Proteins With a High Homology of the Putative CODH Within Other Members of the *Lachnospiraceae* Species

A protein BLAST search (Altschul et al., 1997) of the putative CODH (F7O84_RS11645) was done to investigate if presence of the putative CODH gene also coincided with an apparently incomplete WLP in other members of the *Lachnospiraceae* family. Indeed, 13 sequenced members of the *Lachnospiraceae* family showed predicted proteins with a high homology with the putative CODH (Supplementary Table S4). 9 of the 13 *Lachnospiraceae* members were present in the KEGG database (Kanehisa et al., 2014; Supplementary Table S5), and subsequently analyzed on the presence or absence of the CODH/ACS complex. All organisms contained only a partial WLP, with the ACS genes not identified. In seven of the members, respectively *Lachnoclostridium saccharolyticum*, *Lachnoclostridium phytofermentans*, *Pseudobutyribrio xylanivorans*, *Butyribrio fibrisolvens*, *Pseudobutyribrio xylanivorans*, and both *Roseburia* species the full CODH/ACS complex was not identified. Further study is required to elucidate the relevance of the putative CODH for acetogenic metabolism.

DISCUSSION

Incorporation of carbon from ¹³C labeled bicarbonate into the carbonyl group of acetate supported our previous conclusion, based on product profiles, that acetogenesis occurs in anaerobic, galacturonate-limited enrichment culture of “*Ca. G. soehngeni*” (Valk et al., 2018). A much lower labeling of the methyl group of acetate indicated that, instead of carbon dioxide, the methyl branch of the WLP in the “*Ca. G. soehngeni*” enrichment cultures predominantly used formate as a substrate, generated in the anaerobic fermentation of galacturonate (Figure 1). This conclusion is consistent with the low transcript levels of the only putative formate dehydrogenase gene (F7O84_RS07405; EC 1.17.1.9; Table 6) identified in the “*Ca. G. soehngeni*” MAG, the high transcript level of a putative pyruvate-formate lyase gene (F7O84_RS03160, EC 6.2.1.3; Table 5) and the low net production rates of formate in the anaerobic enrichment cultures (Table 2). In contrast, previous labeling studies on acetogens

TABLE 6 | Transcript levels of putative key genes of the adapted Entner-Doudoroff pathway for galacturonate metabolism and the Wood-Ljungdahl pathway for acetogenesis in meta-transcriptome samples of the “*Ca. G. soehngeni*” chemostat enrichment cultures expressed as reads per kilobase million (RPKM, average \pm average deviation) based on technical triplicates of duplicate enrichment cultures.

Protein function	EC number	Gene ID	RPKM
Adapted entner-doudoroff pathway			
Uronate isomerase	5.3.1.12	F7084_RS17360	5852 \pm 2398
Tagaturonate reductase	1.1.1.58	F7084_RS17370	3067 \pm 1236
Altronate dehydratase	4.2.1.7	F7084_RS17375	8426 \pm 3296
2-Dehydro-3-deoxygluconokinase	2.7.1.45	F7084_RS17390	3863 \pm 1343
2-Dehydro-3-deoxyphosphogluconate aldolase	4.1.2.14	F7084_RS17395	1752 \pm 245
Acetate production			
Pyruvate:ferredoxin oxidoreductase	1.2.7.1	F7084_RS03200	4145 \pm 278
Pyruvate formate lyase	6.2.1.3	F7084_RS03160	1893 \pm 651
Phosphate acetyltransferase	2.3.1.8	F7084_RS05985	1500 \pm 176
Acetate kinase	2.7.2.1	F7084_RS05980	1625 \pm 200
Wood-Ljungdahl pathway			
Formate dehydrogenase	1.17.1.9	F7084_RS07405	14 \pm 3
Formate-tetrahydrofolate ligase	6.3.4.3	F7084_RS05385	256 \pm 58
Methenyl-tetrahydrofolate cyclohydrolase/methylene-tetrahydrofolate dehydrogenase	3.5.4.9 and 1.5.1.5	F7084_RS05385	236 \pm 9
Methyl-tetrahydrofolate reductase	1.5.1.20	F7084_RS08335	126 \pm 13
5-methyl-tetrahydrofolate:corrinoid/iron-sulfur protein methyltransferase	2.1.1.258	F7084_RS02745	144 \pm 19
CO-methylating acetyl-CoA synthase	2.3.1.169		n.d.
CO dehydrogenase	1.2.7.4		n.d.
Prismane/CO dehydrogenase family	1.7.99.1	F7084_RS02405	40 \pm 8
Prismane/CO dehydrogenase family	1.7.99.1	F7084_RS11645	315 \pm 51
Energy-metabolism associated genes			
Electron transport complex protein A	7.2.1.2	F7084_RS03295	58 \pm 5
Electron transport complex protein B	7.2.1.2	F7084_RS03300	261 \pm 40
Electron transport complex protein C	7.2.1.2	F7084_RS03275	329 \pm 22
Electron transport complex protein DG	7.2.1.2	F7084_RS03290	101 \pm 13
Electron transport complex protein E	7.2.1.2	F7084_RS03285	143 \pm 9
Ferredoxin hydrogenase subunit A	1.12.7.2	F7084_RS09545	196 \pm 100
Ferredoxin hydrogenase subunit B	1.12.7.2	F7084_RS09550	356 \pm 32
Ferredoxin hydrogenase subunit C	1.12.7.2	F7084_RS04820	124 \pm 86

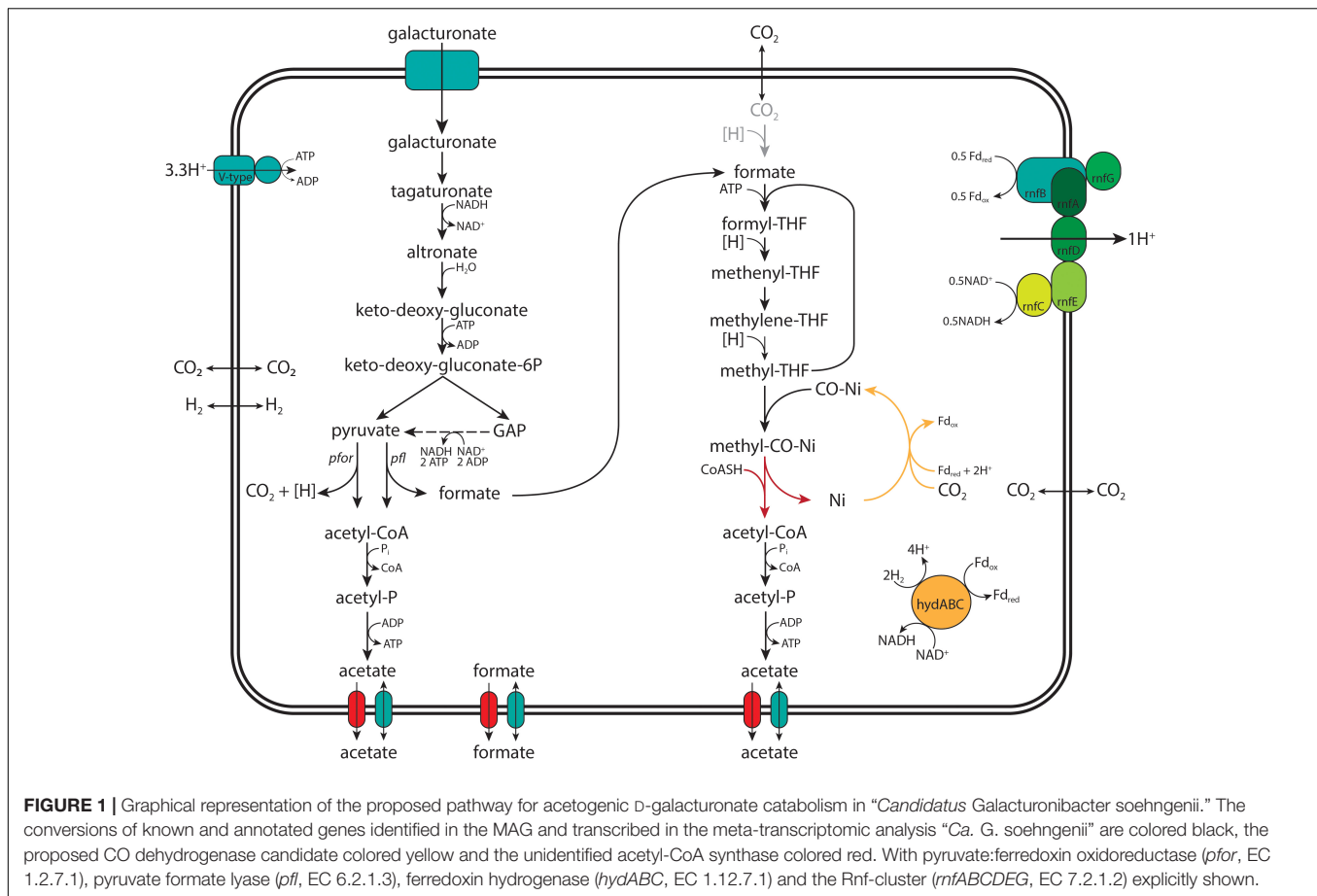
N.d., not detected.

harboring the WLP showed marginal preferential labeling of the carboxyl moiety of acetate (Wood and Harris, 1952; O'Brien and Ljungdahl, 1972; Schulman et al., 1972), indicating the use of extracellular CO₂ as substrate for both the methyl- and carbonyl-groups of acetate.

While the observed labeling pattern was consistent with acetogenic metabolism of galacturonate via a WLP, this did not rule out involvement of another pathway for carbon fixation in acetate. Involvement of the hydroxypropionate bi-cycle, 3-hydroxypropionate/4-hydroxybutyrate cycle and dicarboxylate/hydroxybutyrate cycle were excluded since no homologs were found in the “*Ca. G. soehngeni*” MAG for the majority of genes associated with these three pathways (Supplementary Table S2). Key genes were also missing for the reductive pentose phosphate cycle (rPPP) and reductive citric acid cycle (rTCA) (Supplementary Table S2) and, moreover, neither of these pathways could explain preferential labeling of the carboxyl group of acetate (Alberts et al., 2002; Shimizu et al., 2015). No gene candidates were identified for the glycine cleavage (GCV) system (Supplementary Table S2 and Supplementary

Figure S2) and ¹³C-labeled bicarbonate fed into this pathway should result in equal labeling of the methyl and carbonyl groups of acetate (Figueroa et al., 2018; Supplementary Figure S2). Additionally, none of the routes would require the high CO dehydrogenase enzyme activity measured in cell extracts of the “*Ca. G. soehngeni*” enrichment culture. This analysis leaves the WLP as the only known carbon fixation pathway consistent with the observed stoichiometry of fermentation products, the labeling pattern of acetate and, with the notable exception of the CODH complex, genome and transcriptome analysis of “*Ca. G. soehngeni*.”

Homologs of structural genes encoding enzymes of an adapted Entner-Doudoroff pathway for galacturonate metabolism were highly expressed in the galacturonate-limited, anaerobic “*Ca. G. soehngeni*” enrichment cultures (Table 6). Since conversion of one mole of galacturonate into two moles of pyruvate via this pathway is redox-cofactor neutral, redox equivalents for acetogenesis needed to be derived from pyruvate dissimilation (van Maris et al., 2006; Kuivanen et al., 2019). Pyruvate:ferredoxin oxidoreductase (F7084_RS03200,



EC 1.2.7.1) has been reported to couple fermentation and WLP in other anaerobes (Drake et al., 1981; Menon and Ragsdale, 1996b; Schuchmann and Müller, 2014). Strong, highly transcribed homologs of structural genes for PFOR and for a ferredoxin hydrogenase (EC 1.12.7.2) (Table 6; F7O84_RS03200 and F7O84_0945-50, F7O84_04820 respectively) indicated that it also fulfils this role in “*Ca. G. soehngeni*.”

The significant CO dehydrogenase (CODH) (Weghoff and Müller, 2016) activities in cell extracts enrichment cultures, combined with the incorporation of ^{13}C from bicarbonate in acetate strongly suggested the presence of a functional CODH enzyme in “*Ca. G. soehngeni*.” Two highly conserved classes of CODH enzymes have been described (King and Weber, 2007; Techtmann et al., 2012). Aerobic CODH enzymes (*coxSML* complex; EC 1.2.5.3) have a Mo-Cu-Se associated active site and only use CO as substrate (Schübel et al., 1995; Dobbek et al., 1999). Strictly anaerobic Ni-Fe-S associated CODH (*cooS*, EC 1.2.7.4) can use also CO_2 as substrate (Doukov et al., 2002; Ragsdale, 2008; Techtmann et al., 2012). A close functional relationship between Ni-CO dehydrogenases and hydroxylamine reductases was shown when a single amino-acid substitution was shown to change a Ni-CO dehydrogenase into a hydroxylamine reductase (Heo et al., 2002). Since no strong homologs of canonical aerobic or anaerobic CODH genes were identified, the HCP homolog F7O84_RS11645 is

therefore the best candidate for the observed CODH activity. Our inability to demonstrate stable CODH activity in cell extracts upon expression of F7O84_RS11645 in *E. coli* could have many causes, including improper folding, metal or co-factor requirements (Ensign et al., 1990; Kerby et al., 1997) or requirement of additional subunits or other proteins (Bonams and Luddent, 1987; Bonam et al., 1989; Ensign and Luddent, 1991; Aragão et al., 2008; Bar-Even et al., 2012a). The immediate genetic context of F7O84_RS11645 showed many ORFs encoding predicted conserved proteins with unknown function. Co-expression of fosmid libraries (Shizuya et al., 1992; Ho et al., 2018) of the “*Ca. G. soehngeni*” MAG together with the plasmid used in this study in an *E. coli* strain, may be helpful in resolving the genetic requirements for CODH activity in this organism.

It remains unclear how the CODH-dependent carbonyl branch and formate-dependent methyl branch of a WLP pathway in “*Ca. G. soehngeni*” organism are linked. The present study is not the first in which carbon fixation linked to the WLP was observed in the absence of a full complement of canonical WLP structural genes (Zhuang et al., 2014; Figueroa et al., 2018). However, no clear physiological nor phylogenetic connections were detected between “*Ca. G. soehngeni*” and the organisms studied previously, a strict dehalogenide-respiring *Dehalococcoides mccartyi* strain from the Chloroflexi phylum

and the phosphite-oxidizing Deltaproteobacterium “*Candidatus Phosphitovorax anaerolimi*” Phox-21, respectively.

This study illustrates how quantitative analysis of metabolite formation by chemostat enrichment cultures, combined with ¹³C-labeling, (meta-)genome assembly and annotation, meta-transcriptome analysis and biochemical assays can raise new and surprising questions about intensively studied metabolic pathways. Based on our results, involvement of a novel inorganic carbon assimilation pathway, which produces a similar labeling and product profile as the WLP, cannot be fully excluded. However, despite the wide distribution of the CODH/ACS complex in Bacteria and Archaea (Schuchmann and Müller, 2016), the available evidence appears to point in the direction of an as yet unidentified link between the methyl and carbonyl branches of the WLP. Further research to resolve this issue may benefit from additional labeling studies with ¹³C-bicarbonate, ¹³C-formate or partially labeled D-galacturonate combined with metabolome analysis and *in vitro* enzyme activity studies of formate dehydrogenase. Such studies are complicated by our current inability to grow “*Ca. G. soehngeni*” in pure cultures (Valk et al., 2018). The organisms shown in the **Supplementary Table S4** might be interesting alternative organisms to study in more detail, as they are available in pure culture. It would therefore be relevant to identify if any of these organisms exhibit a similar acetogenic metabolism, with an incomplete complement of WLP enzymes, to further explore this intriguing metabolic conundrum.

DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in the European Nucleotide Archive (ENA) under the BioProject ID PRJNA566068, NCBI GenBank accession number MN498128.

REFERENCES

- Adam, P. S., Borrel, G., and Gribaldo, S. (2018). Evolutionary history of carbon monoxide dehydrogenase/acetyl-CoA synthase, one of the oldest enzymatic complexes. *Proc. Natl. Acad. Sci. U.S.A.* 115, E1166–E1173. doi: 10.1073/pnas.1716667115
- Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (2002). *Molecular Biology of the Cell*, 4th Edn. New York, NY: Garland Science.
- Almeida, C. C., Romão, C. V., Lindley, P. F., Teixeira, M., and Saraiva, L. M. (2006). The role of the hybrid cluster protein in oxidative stress defense. *J. Biol. Chem.* 281, 32445–32450. doi: 10.1074/jbc.M605888200
- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402. doi: 10.1093/nar/25.17.3389
- Amann, R. L., Binder, B. J., Olson, R. J., Chisholm, S. W., Devereux, R., and Stahl, D. A. (1990). Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. *Appl. Environ. Microbiol.* 56, 1919–1925. doi: 10.1128/aem.56.6.1919-1925.1990
- Aragão, D., Macedo, S., Mitchell, E. P., Romão, C. V., Liu, M. Y., Frazão, C., et al. (2003). Reduced hybrid cluster proteins (HCP) from *Desulfovibrio desulfuricans* ATCC 27774 and *Desulfovibrio vulgaris* (Hildenborough): X-ray structures at high resolution using synchrotron radiation. *J. Biol. Inorg. Chem.* 8, 540–548. doi: 10.1007/s00775-003-0443-x
- Aragão, D., Mitchell, E. P., Frazão, C. F., Carrondo, M. A., and Lindley, P. F. (2008). Structural and functional relationships in the hybrid cluster protein family:

AUTHOR CONTRIBUTIONS

ML, JTP, and LV designed the experiments, interpreted the results, and wrote the manuscript. LV did all cultivations and labeling study. LV and MD performed the enzyme activity assays and heterologous experiment. LV and JFP performed the qFISH analysis. GS made the model. MSD performed the experimental work for the meta-transcriptomic and meta-genomic analysis. MSD, LV, and PN analyzed the data. All authors read and approved of the final manuscript.

FUNDING

This research was supported by the SIAM Gravitation Grant 024.002.002, the Netherlands Organisation for Scientific Research.

ACKNOWLEDGMENTS

We thank J. L. Rombouts for the design of the FISH probe Lac87 used in this study and P. de Waard for the assistance with NMR analysis. We additionally thank Dr. A. M. Vos and Dr. M. D. Verhoeven for critical reading of the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2020.00063/full#supplementary-material>

- structure of the anaerobically purified hybrid cluster protein from *Desulfovibrio vulgaris* at 1.35 Å resolution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 64, 665–674. doi: 10.1107/S0907444908009165
- Bar-Even, A., Flamholz, A., Noor, E., and Milo, R. (2012a). Thermodynamic constraints shape the structure of carbon fixation pathways. *Biochim. Biophys. Acta - Bioenerg.* 1817, 1646–1659. doi: 10.1016/j.bbapbio.2012.05.002
- Bar-Even, A., Noor, E., and Milo, R. (2012b). A survey of carbon fixation pathways through a quantitative lens. *J. Exp. Bot.* 63, 2325–2342. doi: 10.1093/jxb/err417
- Berg, I. A. (2011). Ecological aspects of the distribution of different autotrophic CO₂ fixation pathways. *Appl. Environ. Microbiol.* 77, 1925–1936. doi: 10.1128/AEM.02473-10
- Biegel, E., and Müller, V. (2010). Bacterial Na⁺-translocating ferredoxin: NAD⁺ oxidoreductase. *Proc. Natl. Acad. Sci. U.S.A.* 107, 18138–18142. doi: 10.1073/pnas.1010318107
- Boetzer, M., and Pirovano, W. (2012). Toward almost closed genomes with GapFiller. *Genome Biol.* 13:R56. doi: 10.1186/gb-2012-13-6-r56
- Boetzer, M., and Pirovano, W. (2014). SSPACE-LongRead: scaffolding bacterial draft genomes using long read sequence information. *BMC Bioinformatics* 15:211. doi: 10.1186/1471-2105-15-211
- Bonam, D., Lehman, L., Roberts, G. P., and Ludden, P. W. (1989). Regulation of carbon monoxide dehydrogenase and hydrogenase in *Rhodospirillum rubrum*: effects of CO and oxygen on synthesis and activity. *J. Bacteriol.* 171, 3102–3107. doi: 10.1128/jb.171.6.3102-3107.1989
- Bonams, D., and Ludden, P. W. (1987). Purification and characterization of carbon monoxide. *J. Biol. Chem.* 262, 2980–2987.

- Caspi, R., Foerster, H., Fulcher, C. A., Kaipa, P., Krummenacker, M., Latendresse, M., et al. (2008). The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Res.* 36, 623–631. doi: 10.1093/nar/gkm900
- Daims, H., Brühl, A., Amann, R., Schleifer, K.-H., and Wagner, M. (1999). The domain-specific probe EUB338 is insufficient for the detection of all bacteria: development and evaluation of a more comprehensive probe set. *Syst. Appl. Microbiol.* 22, 434–444. doi: 10.1016/s0723-2020(99)80053-8
- Daims, H., Lückner, S., and Wagner, M. (2006). Daime, a novel image analysis program for microbial ecology and biofilm research. *Environ. Microbiol.* 8, 200–213. doi: 10.1111/j.1462-2920.2005.00880.x
- Daims, H., Stoecker, K., and Wagner, M. (2005). “Fluorescence in situ hybridization for the detection of prokaryotes,” in *Molecular Microbial Ecology*, eds A. M. Osborn, and C. J. Smith, (New York, NY: Taylor & Francis), 213–239.
- Darnault, C., Volbeda, A., Kim, E. J., Legrand, P., Vernède, X., Lindahl, P. A., et al. (2003). Ni-Zn-[Fe4-S4] and Ni-Ni-[Fe4-S4] clusters in closed and open α subunits of acetyl-CoA synthase/carbon monoxide dehydrogenase. *Nat. Struct. Biol.* 10, 271–279. doi: 10.1038/nsb912
- De Kok, S., Meijer, J., van Loosdrecht, M. C. M., and Kleerebezem, R. (2013). Impact of dissolved hydrogen partial pressure on mixed culture fermentations. *Appl. Microbiol. Biotechnol.* 97, 2617–2625. doi: 10.1007/s00253-012-4400-x
- Deamer, D., Akeson, M., and Branton, D. (2016). Three decades of nanopore sequencing. *Nat. Biotechnol.* 34, 518–524. doi: 10.1038/nbt.3423
- Diender, M., Pereira, R., Wessels, H. J. C. T., Stams, A. J. M., and Sousa, D. Z. (2016). Proteomic analysis of the hydrogen and carbon monoxide metabolism of *Methanothermobacter marburgensis*. *Front. Microbiol.* 7:1049. doi: 10.3389/fmicb.2016.01049
- Dobbek, H., Gremer, L., Meyer, O., and Huber, R. (1999). Crystal structure and mechanism of CO dehydrogenase, a molybdo iron-sulfur flavoprotein containing S-selenylcysteine. *Proc. Natl. Acad. Sci. U.S.A.* 96, 8884–8889. doi: 10.1073/pnas.96.16.8884
- Doukov, T. I., Iverson, T. M., Seravalli, J., Ragsdale, S. W., and Drennan, C. L. (2002). A Ni-Fe-Cu center in a bifunctional carbon monoxide dehydrogenase/acetyl-CoA synthase. *Science* 298, 567–573. doi: 10.1126/science.1075843
- Drake, H. L., Gößner, A. S., and Daniel, S. L. (2008). Old acetogens, new light. *Ann. N. Y. Acad. Sci.* 1125, 100–128. doi: 10.1196/annals.1419.016
- Drake, H. L., Hu, S. I., and Wood, H. G. (1981). Purification of five components from *Clostridium thermoaceticum* which catalyze synthesis of acetate from pyruvate and methyltetrahydrofolate. *J. Biol. Chem.* 256, 11137–11144.
- Eggen, R. I. L., Geerling, A. C. M., Jetten, M. S. M., and De Vos, W. M. (1991). Cloning, expression, and sequence analysis of the genes for carbon monoxide dehydrogenase of *Methanotherx soehngenii*. *J. Biol. Chem.* 266, 6883–6887.
- Eggen, R. I. L., van Kranenburg, R., Vriesema, A. J. M., Geerling, A. C. M., Verhagen, M. F. J. M., Hagen, W. R., et al. (1996). Carbon monoxide dehydrogenase from *Methanosarcina frisia* Gö1. *J. Biol. Chem.* 271, 14256–14263. doi: 10.1074/jbc.271.24.14256
- El-Gebali, S., Mistry, J., Bateman, A., Eddy, S. R., Luciani, A., Potter, S. C., et al. (2018). The Pfam protein families database in 2019. *Nucleic Acids Res.* 47, 427–432. doi: 10.1093/nar/gky995
- Ensign, S. A., Campbell, M. J., and Ludden, P. W. (1990). Activation of the nickel-deficient carbon monoxide dehydrogenase from *Rhodospirillum rubrum*: kinetic characterization and reductant requirement. *Biochemistry* 29, 2162–2168. doi: 10.1021/bi00460a029
- Ensign, S. A., and Ludden, P. W. (1991). Characterization of the CO oxidation/H₂ evolution system of *Rhodospirillum rubrum*: role of a 22-kDa iron-sulfur protein in mediating electron transfer between carbon monoxide dehydrogenase and hydrogenase. *J. Biol. Chem.* 266, 18395–18403.
- Figuerola, I. A., Barnum, T. P., Somasekhar, P. Y., Carlström, C. I., Engelbrektson, A. L., and Coates, J. D. (2018). Metagenomics-guided analysis of microbial chemolithoautotrophic phosphite oxidation yields evidence of a seventh natural CO₂ fixation pathway. *Proc. Natl. Acad. Sci. U.S.A.* 115, E92–E101. doi: 10.1073/pnas.1715549114
- Fuchs, G. (2011). Alternative pathways of carbon dioxide fixation: insights into the early evolution of life? *Annu. Rev. Microbiol.* 65, 631–658. doi: 10.1146/annurev-micro-090110-102801
- Heo, J., Wolfe, M. T., Staples, C. R., and Ludden, P. W. (2002). Converting the NiFeS carbon monoxide dehydrogenase to a hydrogenase and a hydroxylamine reductase. *J. Bacteriol.* 184, 5894–5897. doi: 10.1128/JB.184.21.5894-5897.2002
- Ho, J. C. H., Pawar, S. V., Hallam, S. J., and Yadav, V. G. (2018). An improved whole-cell biosensor for the discovery of lignin-transforming enzymes in functional metagenomic screens. *ACS Synth. Biol.* 7, 392–398. doi: 10.1021/acssynbio.7b00412
- Jain, M., Olsen, H. E., Paten, B., and Akeson, M. (2016). The oxford nanopore MinION: delivery of nanopore sequencing to the genomics community. *Genome Biol.* 17:239.
- Jeoung, J.-H., and Dobbek, H. (2011). Ni, Fe-Containing Carbon Monoxide Dehydrogenases. *Encycl. Inorg. Bioinorg. Chem.* 179, 1–11. doi: 10.1002/9781119951438.eibc0640
- Kanehisa, M., Goto, S., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M. (2014). Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Res.* 42, D199–D205. doi: 10.1093/nar/gkt1076
- Kempf, V. A. J., Trebesius, K., and Autenrieth, I. B. (2000). Fluorescent in situ hybridization allows rapid identification of microorganisms in blood cultures. *J. Clin. Microbiol.* 38, 830–838. doi: 10.1128/jcm.38.2.830-838.2000
- Kerby, R. L., Ludden, P. W., and Roberts, G. P. (1997). In vivo nickel insertion into the carbon monoxide dehydrogenase of *Rhodospirillum rubrum*: molecular and physiological characterization of cooCTJ. *J. Bacteriol.* 179, 2259–2266. doi: 10.1128/jb.179.7.2259-2266.1997
- King, G. M., and Weber, C. F. (2007). Distribution, diversity and ecology of aerobic CO-oxidizing bacteria. *Nat. Rev. Microbiol.* 5, 107–118. doi: 10.1038/nrmicro1595
- Kuivanen, J., Biz, A., and Richard, P. (2019). Microbial hexuronate catabolism in biotechnology. *AMB Express* 9, 1–11. doi: 10.1186/s13568-019-0737-1
- Markowitz, V. M., Chen, I. M. A., Palaniappan, K., Chu, K., Szeto, E., Grechkin, Y., et al. (2012). IMG: the integrated microbial genomes database and comparative analysis system. *Nucleic Acids Res.* 40, 115–122. doi: 10.1093/nar/gkr1044
- Maupin-Furlow, J. A., and Ferry, J. G. (1996). Analysis of the CO dehydrogenase/acetyl-coenzyme A synthase operon of *Methanosarcina thermophila*. *J. Bacteriol.* 178, 6849–6856. doi: 10.1128/jb.178.23.6849-6856.1996
- Menon, S., and Ragsdale, S. W. (1996a). Evidence that carbon monoxide is an obligatory intermediate in anaerobic acetyl-CoA synthesis. *Biochemistry* 35, 12119–12125. doi: 10.1021/bi961014d
- Menon, S., and Ragsdale, S. W. (1996b). Unleashing hydrogenase activity in carbon monoxide dehydrogenase/acetyl-CoA synthase and pyruvate: ferredoxin oxidoreductase. *Biochemistry* 35, 15814–15821. doi: 10.1021/bi9615598
- O'Brien, W. E., and Ljungdahl, L. G. (1972). Fermentation of fructose and synthesis of acetate from carbon dioxide by *Clostridium formicoaceticum*. *J. Bacteriol.* 109, 626–632. doi: 10.1128/jb.109.2.626-632.1972
- Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., and Tyson, G. W. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* 25, 1043–1055. doi: 10.1101/gr.186072.114
- Ragsdale, S. W. (2008). Enzymology of the Woods-Ljungdahl pathway of acetogenesis. *Ann. N. Y. Acad. Sci.* 1125, 129–136. doi: 10.1196/annals.1419.015
- Ragsdale, S. W., and Kumar, M. (1996). Nickel-containing carbon monoxide dehydrogenase/Acetyl-CoA synthase. *Chem. Rev.* 96, 2515–2539.
- Ragsdale, S. W., and Pierce, E. (2008). Acetogenesis and the Wood-Ljungdahl pathway of CO₂ fixation. *Biochim. Biophys. Acta* 1784, 1873–1898. doi: 10.1016/j.bbapap.2008.08.012
- Roels, J. (1983). *Energetics and Kinetics in Biotechnology*. Amsterdam: Biomedical Press.
- Rumble, J. R., Linde, D. R., and Bruno, T. J. (2017). *CRC Handbook of Chemistry and Physics*, 98th Edn. Boca Raton, FL: CRC Press LLC.
- Schübel, U., Kraut, M., Mörsdorf, G., and Meyer, O. (1995). Molecular characterization of the gene cluster coxMSL encoding the molybdenum-containing carbon monoxide dehydrogenase of *Oligotropha carboxidovorans*. *J. Bacteriol.* 177, 2197–2203. doi: 10.1128/jb.177.8.2197-2203.1995
- Schuchmann, K., and Müller, V. (2013). Direct and reversible hydrogenation of CO₂ to formate by a bacterial carbon dioxide reductase. *Science* 342, 1382–1386. doi: 10.1126/science.1244758

- Schuchmann, K., and Müller, V. (2014). Autotrophy at the thermodynamic limit of life: a model for energy conservation in acetogenic bacteria. *Nat. Rev. Microbiol.* 12, 809–821. doi: 10.1038/nrmicro3365
- Schuchmann, K., and Müller, V. (2016). Energetics and application of heterotrophy in acetogenic bacteria. *Appl. Environ. Microbiol.* 82, 4056–4069. doi: 10.1128/AEM.00882-16
- Schulman, M., Donald Parker, Ljungdahl, L. G., and Wood, H. G. (1972). Total synthesis of acetate from CO₂. V. Determination by mass analysis of the different types of acetate formed from ¹³CO₂ by heterotrophic bacteria. *J. Bacteriol.* 109, 633–644. doi: 10.1146/annurev.mi.23.100169.002503
- Shimizu, R., Dempo, Y., Nazkayama, Y., Nakamura, S., Bamba, T., Fukusaki, E., et al. (2015). New insight into the role of the Calvin cycle: reutilization of CO₂ emitted through sugar degradation. *Sci. Rep.* 5:11617. doi: 10.1038/srep11617
- Shizuya, H., Birren, B., Kim, U.-J., Mancino, V., Slepak, T., Tachiiri, Y., et al. (1992). Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. *Proc. Natl. Acad. Sci. U.S.A.* 89, 8794–8797. doi: 10.1073/pnas.89.18.8794
- Teichtmann, S. M., Lebedinsky, A. V., Colman, A. S., Sokolova, T. G., Woyke, T., Goodwin, L., et al. (2012). Evidence for horizontal gene transfer of anaerobic carbon monoxide dehydrogenases. *Front. Microbiol.* 3:132. doi: 10.3389/fmicb.2012.00132
- Valk, L. C., Frank, J., de la Torre-Cortes, P., van 't Hof, M., van Maris, A. J. A., and Pronk, J. T. (2018). Galacturonate metabolism in anaerobic chemostat enrichment cultures: combined fermentation and acetogenesis by the dominant sp. nov. “*Candidatus Galacturonibacter soehngeni*.” *Appl. Environ. Microbiol.* 84:e01370-18. doi: 10.1128/AEM.01370-18
- Vallenet, D., Calteau, A., Cruveiller, S., Gachet, M., Lajus, A., Josso, A., et al. (2017). MicroScope in 2017: an expanding and evolving integrated resource for community expertise of microbial genomes. *Nucleic Acids Res.* 45, D517–D528. doi: 10.1093/nar/gkw1101
- Vallenet, D., Labarre, L., Rouy, Z., Barbe, V., Bocs, S., Cruveiller, S., et al. (2006). MaGe: a microbial genome annotation system supported by synteny results. *Nucleic Acids Res.* 34, 53–65. doi: 10.1093/nar/gkj406
- van den Berg, W. A. M., Hagen, W. R., and van Dongen, W. M. A. M. (2000). The hybrid-cluster protein (‘prismane protein’) from *Escherichia coli*. *Eur. J. Biochem.* 267, 666–676. doi: 10.1046/j.1432-1327.2000.01032.x
- van Maris, A. J. A., Abbott, D. A., Bellissimi, E., van den Brink, J., Kuyper, M., Luttik, M. A. H., et al. (2006). Alcoholic fermentation of carbon sources in biomass hydrolysates by *Saccharomyces cerevisiae*: current status. *Antonie van Leeuwenhoek* 90, 391–418. doi: 10.1007/s10482-006-9085-7
- Weghoff, M. C., and Müller, V. (2016). CO metabolism in the thermophilic acetogen *Thermoanaerobacter kivui*. *Appl. Environ. Microbiol.* 82, 2312–2319. doi: 10.1128/AEM.00122-16
- Wolfe, M. T., Heo, J., Garavelli, J. S., and Ludden, P. W. (2002). Hydroxylamine reductase activity of the hybrid cluster protein from *Escherichia coli*. *J. Bacteriol.* 184, 5898–5902. doi: 10.1128/JB.184.21.5898-5902.2002
- Wood, H. G., and Harris, D. L. (1952). A study of carbon dioxide fixation by mass determination of the types of C₁₃-acetate. *J. Biol. Chem.* 194, 905–931.
- Zhuang, W., Yi, S., Bill, M., Brisson, V. L., Feng, X., Men, Y., et al. (2014). Incomplete Wood–Ljungdahl pathway facilitates one-carbon metabolism in organohalide-respiring *Dehalococcoides mccartyi*. *Proc. Natl. Acad. Sci. U.S.A.* 111, 6419–6424. doi: 10.1073/pnas.1321542111

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Valk, Diender, Stouten, Petersen, Nielsen, Dueholm, Pronk and van Loosdrecht. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.