environmental microbiology

Environmental Microbiology (2016) **18**(9), 2843–2855 doi:10.1111/1462-2920.12973

Unravelling the one-carbon metabolism of the acetogen *Sporomusa* strain An4 by genome and proteome analysis

Michael Visser,[1]* Mervin M. Pieterse,[2,3] Martijn W. H. Pinkse,[2,3] Bart Nijsse,[4] Peter D. E. M. Verhaert,[2,3] Willem M. de Vos,[1,5,6] Peter J. Schaap[4] and Alfons J. M. Stams[1,7]

[1]*Laboratory of Microbiology and [4]Laboratory of Systems and Synthetic Biology, Wageningen University, Stippeneng 4, 6708 WE Wageningen, The Netherlands.*

[2]*Department of Biotechnology, Technische Universiteit Delft, Julianalaan 67, 2628 BC Delft, The Netherlands.*

[3]*Netherlands Proteomics Centre, Julianalaan 67, 2628 BC Delft, The Netherlands.*

Departments of [5]Veterinary Biosciences and [6]Bacteriology and Immunology, Helsinki University, Helsinki, Finland.

[7]*CEB-Centre of Biological Engineering, University of Minho, Campus de Gualtar, 4710-057 Braga, Portugal.*

Summary

The *Sporomusa* genus comprises anaerobic spore-forming acetogenic bacteria that stain Gram-negative. *Sporomusa* species typically grow with one-carbon substrates and N-methylated compounds. In the degradation of these compounds methyltransferases are involved. In addition, *Sporomusa* species can grow autotrophically with H_2 and CO_2, and use a variety of sugars for acetogenic growth. Here we describe a genome analysis of *Sporomusa* strain An4 and a proteome analysis of cells grown under five different conditions. Comparison of the genomes of *Sporomusa* strain An4 and *Sporomusa ovata* strain H1 indicated that An4 is a *S. ovata* strain. Proteome analysis showed a high abundance of several methyltransferases, predominantly trimethylamine methyltransferases, during growth with betaine, whereas trimethylamine is one of the main end-products of betaine degradation. In methanol degradation methyltransferases are also involved. In methanol-utilizing methanogens, two methyltransferases catalyse methanol conversion, methyltransferase 1 composed of subunits MtaB and MtaC and methyltransferase 2, also called MtaA. The two methyltransferase 1 subunits MtaB and MtaC were highly abundant when strain An4 was grown with methanol. However, instead of MtaA a methyltetrahydrofolate methyltransferase was synthesized. We propose a novel methanol degradation pathway in *Sporomusa* strain An4 that uses a methyltetrahydrofolate methyltransferase instead of MtaA.

Received 2 April, 2015; revised 19 June, 2015; accepted 26 June, 2015. *For correspondence. E-mail: michael.visser@wur.nl; Tel. 0317 483739.

Introduction

The genus *Sporomusa* was described in 1984 (Möller *et al.*, 1984). It is a genus of motile spore-forming acetogenic bacteria that stain Gram-negative. *Sporomusa* strains have been isolated from soils and sediments (Möller *et al.*, 1984; Hermann *et al.*, 1987; Dehning *et al.*, 1989; Sass *et al.*, 1998), wastewater (Möller *et al.*, 1984; Ollivier *et al.*, 1989) and the gut and faeces of animals (Möller *et al.*, 1984; Breznak and Switzer, 1986; Breznak *et al.*, 1988; Boga *et al.*, 2003). Currently, the genus *Sporomusa* consists of nine validated species. The characteristic substrates for *Sporomusa* species are one-carbon compounds, such as methanol, and N-methylated compounds such as betaine. In addition, they can grow autotrophically with H_2 and CO_2, and use a variety of sugars for acetogenic growth. Like other acetogenic bacteria, they employ the acetyl-coenzyme A (CoA) pathway for energy conservation and CO_2 fixation (Drake *et al.*, 1997; 2007).

Sporomusa strain An4 was isolated from an underground gas storage in Russia with methanol and perchlorate as substrates. Besides perchlorate, the strain was also able to use nitrate as electron acceptor. Unfortunately, in the course of this study, the strain lost its ability to use perchlorate as electron acceptor. The ability to respire with inorganic electron acceptors is not very common among *Sporomusa* species; only *Sporomusa ovata* and *Sporomusa* strain An4 have been described to reduce nitrate to ammonium (Möller *et al.*, 1984; Balk *et al.*, 2010). *Sporomusa* strain An4 is, according to phylogenetic analysis based on 16S rRNA gene

sequences, closely related to *Sporomusa* strain DR5 and *S. ovata* strain H1 (99% and 98% sequence similarity respectively) (Balk *et al*., 2010).

Sporomusa strains are known to produce atypical corrinoids. More than 90% of the corrinoids in *S. ovata* consist of two synthesized coenzyme B_{12} analogues: *p*-cresolyl cobamide and phenolyl cobamide (Stupperich *et al*., 1988; 1989; 1990). Some other studies provided insight into the transfer of the methyl group from methylated substrates to tetrahydrofolate (THF). For *S. ovata* the involvement of a cobamide-containing protein in the formation of methyl-tetrahydrofolate (CH_3-THF) was shown (Stupperich and Konle, 1993). The synthesis of this protein was induced when *S. ovata* was grown on methanol. The cobamide-containing methyltransferase (MT) of *S. ovata* was purified and characterized (Stupperich *et al*., 1992; Wagner *et al*., 1994). Moreover, it was shown that different MTs are involved in degradation of different methylated-substrates (Stupperich and Konle, 1993) and cytochromes appeared to play a role in autotrophic growth and methyl group oxidation (Kamlage and Blaut, 1993). These and other studies (Dobrindt and Blaut, 1996; Wagner *et al*., 2000) gave insight into the physiology of *Sporomusa.*

Lack of genome information hampered further understanding of the metabolism of *Sporomusa* species. Only recently genomic information became available for *S. ovata* strain H1 (Poehlein *et al*., 2013). Here, we describe a comparison of the recently sequenced genome of *S. ovata* strain H1 and *Sporomusa* strain An4. Shotgun proteomics was applied to get insight into the catabolic pathways and the regulatory aspects of the metabolism of *Sporomusa* strain An4. Five different conditions were selected: acetogenic growth with H_2 and CO_2, methanol, betaine, and fructose; and respiratory growth with methanol and nitrate.

Results

Genome comparison of Sporomusa strain An4 and S. ovata strain H1

Whole genome shotgun sequencing combined with mate-pair sequencing of *Sporomusa* strain An4 resulted in 4.9 Mbp assembled sequence data with an average scaffold size of 308 kbp. The draft genome was automatically annotated using the RAST Web service (Aziz *et al*., 2008). The genomes of *Sporomusa* strain An4 and *S. ovata* strain H1 (Poehlein *et al*., 2013) were compared at DNA level. The average nucleotide identity (ANI) was above 99%. This ANI value is significantly higher than the 95–96% ANI values shown to correspond to the 70% DNA-DNA hybridization level (Richter and Rossello-Mora, 2009). This indicates that *Sporomusa* strain An4 is a *S. ovata* strain. The genome of *Sporomusa* strain An4 and strain H1 contained 5262 and 5110 protein-coding genes respectively. A domain analysis was performed to get insight into functional differences. The genome of the *S. ovata* strain H1 contained six domains that could not be found in the genome of strain An4. They are the domains of genes that code for two transposases, a protein of unknown function, an isopentenyl-diphosphate delta-isomerase and a toxin–antitoxin protein. The genome of strain An4 contained two domains that could not be found in the genome of *S. ovata* strain H1. The genes that contain these domains code for a multicopper oxidase (SpAn4DRAFT_0241) and a peptidase (SpAn4DRAFT_0978). The function of these proteins in strain An4 is currently not known.

Acetogenesis and the acetyl-CoA pathway

All genes coding for enzymes of the acetyl-CoA pathway were found to be present in the genome of *Sporomusa* strain An4 (Fig. 1). Many of these genes are situated in close proximity with each other in scaffold 4, such as a methenyl-THF cyclohydrolase, methylene-THF dehydrogenase, the acetyl-CoA synthase operon and formate dehydrogenase (FDH) genes. The genes coding for the acetate kinase and the phosphate acetyltransferase are in other scaffolds. The genome of strain An4 contains several genes that code for the same enzymes of the acetyl-CoA pathway, including formate-THF ligase (SpAn4DRAFT_2921, SpAn4DRAFT_4515), methenyl-THF cyclohydrolase (SpAn4DRAFT_2056, SpAn4DRAFT_2917), methylene-THF reductase (metF, SpAn4DRAFT_2800, SpAn4DRAFT_2934, SpAn4DRAFT_5207), acetyl-CoA synthase catalytic subunit (acsB, SpAn4DRAFT_2715, SpAn4DRAFT_2923 and SpAn4DRAFT_3609) and carbon monoxide dehydrogenase (CODH) catalytic subunit (cooS, SpAn4DRAFT_0020, SpAn4DRAFT_2714, SpAn4DRAFT_2922). The proteomic analysis showed peptide abundance of all enzymes of the acetyl-CoA pathway in all growth conditions (Fig. 2).

One-carbon metabolism

One-carbon compounds, for example carbon monoxide, formate and methanol, are characteristic growth substrates of *Sporomusa* species. Genes putatively coding for enzymes involved in one-carbon metabolism could be found in the genome of strain An4. Growth with CO requires a CODH. The genome of strain An4 contains three genes that putatively code for a CODH catalytic subunit (cooS, SpAn4DRAFT_0020, 2714, 2922). However, not all products of these genes can be found in the proteomics data (File S1). The proteomics data only

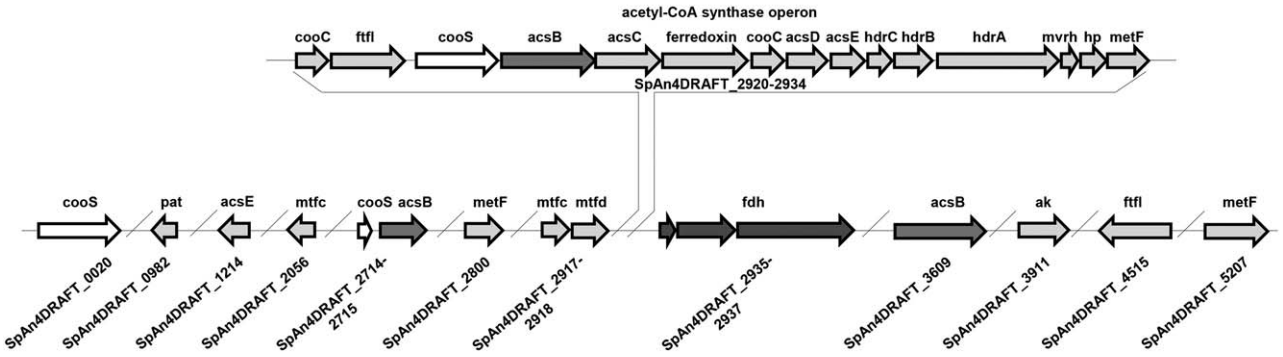

Fig. 1. Relative position of the acetyl-CoA synthase operon and genes of the acetyl-CoA pathway outside the operon. Scaffold and base pair numbers are located below the genes. Abbreviations are acsE: CH_3-THF methyltransferase, ak: acetate kinase, fdh: formate dehydrogenase, ftfl: formate-THF ligase, hdr: heterodisulfide reductase, hp: hypothetical protein, mtfc: methenyl-THF cyclohydrolase, mtfd: methylene-THF dehydrogenase, mvrh: methyl viologen-reducing hydrogenase, pat: phosphate acetyltransferase.

show abundance of the gene product of the *cooS* that is part of the acetyl-CoA synthase operon (SpAn4DRAFT_2922). This cooS is abundantly present in all growth conditions (Fig. 2).

Sporomusa strain An4 can grow with formate, and formate is also an intermediate of the acetyl-CoA pathway. The genome of strain An4 contains four putative FDHs, two dimeric cytoplasmic FDHs (SpAn4DRAFT_1877–1878, SpAn4DRAFT_5070–5071) and one dimeric extracytoplasmic FDH (SpAn4DRAFT_1889–1890); the small subunit contains a twin arginine consensus motif of Tat signal peptides involved in Sec-independent protein translocation, which suggests transport of the FDH complex across the cytoplasmic membrane. One trimeric cytoplasmic FDH (SpAn4DRAFT_2935–2937) contains a predicted ferredoxin and a nicotinamide adenine dinucleotide (NADH)-binding site, suggesting a possible bifurcating function as described by Wang and colleagues (2013). The latter FDH genes are situated upstream of the acetyl-CoA synthase operon (Fig. 1), and the products of these FDH genes were present in high abundance in the proteome analysis of all growth conditions (Fig. 2). The other two cytoplasmic FDHs were not detected in the proteome analysis, and the extracytoplasmic FDH (SpAn4DRAFT_1889–1890) only showed abundance in one of the five growth conditions (File S1). This suggests that the trimeric cytoplasmic FDH (SpAn4DRAFT_2935–2937) in strain An4 is the sole FDH involved in the acetyl-CoA pathway.

The genome of strain An4 contains a gene coding for an MT that is predicted to be methanol specific (SpAn4DRAFT_1215). The amino acid sequence of this putative MT was used for BLAST analysis to find homologues. BLAST results showed 99% coverage and 62% amino acid identity with the methanol: corrinoid MT, MtaB, of *Moorella thermoacetica* (Pierce *et al*., 2008). Another MT gene (SpAn4DRAFT_1216) is situated downstream of the methanol MT in the genome of *Sporomusa* strain An4. This gene is homologous to the MT MtaC of *M. thermoacetica*, with 98% coverage and 45% amino acid identity. Situated upstream of the *mtaB* homologue is a CH_3-THF MT gene (SpAn4DRAFT_1214) and a gene coding for a cobalamin synthesis protein (SpAn4DRAFT_1213, Fig. 3).

In *M. thermoacetica*, the two genes *mtaB* and *mtaC* are known to code for two subunits of a methanol-specific MT (Zhou *et al*., 2005; Das *et al*., 2007). Another MT thought to be involved in the methanol metabolism in *M. thermoacetica* is encoded by a *mtaA* gene (Das *et al*., 2007; Pierce *et al*., 2008). Six possible *mtaA* genes can be found in the genome of *Sporomusa* strain An4 (SpAn4_0190, 0193, 2181, 4129, 4142, 4143). BLASTX analysis showed that these six genes had 93–97% coverage and 24–31% amino acid identity with *MtaA* in *M. thermoacetica*.

Proteome analysis revealed that the products of the two MT genes, *mtaB* and *mtaC*, had a high protein abundance in strain An4 when grown with methanol (Fig. 2). However, none of the six possible *mtaA* homologues showed any abundance in the proteome of cells grown with methanol (File S1). Moreover, the gene product of the CH_3-THF MT (SpAn4DRAFT_1214) gene upstream of the *mtaB* showed high abundance when grown with methanol, suggesting its involvement in methanol metabolism.

The MtaC is a cobamide-binding protein. Cobamide biosynthesis and transport proteins in strain An4 were found in the proteome (File S1). More genes coding for MTs that function in a similar manner as the methanol MTs were found in the genome of strain An4. These include genes putatively coding for tri-, di- and monomethylamine (TMA, DMA and MMA respectively) MTs, which can be involved in the utilization of methylamines or the methyl groups of betaine (Möller *et al*., 1984; Andreesen, 1994; Krätzer *et al*., 2009).

Predicted function	Locus tag	H_2	CO_2	MeOH	NO_3^-	B	F
acetyl-CoA pathway							
phosphate acetyltransferase	SpAn4DRAFT_0982						
AcsB	SpAn4DRAFT_2715						
MetF	SpAn4DRAFT_2800						
methenyl-THF cyclohydrolase	SpAn4DRAFT_2917						
methylene-THF dehydrogenase	SpAn4DRAFT_2918						
CooC I	SpAn4DRAFT_2920						
formate-THF ligase	SpAn4DRAFT_2921						
CooS	SpAn4DRAFT_2922						
AcsB	SpAn4DRAFT_2923						
AcsC	SpAn4DRAFT_2924						
ferredoxin	SpAn4DRAFT_2925						
CooC II	SpAn4DRAFT_2926						
AcsD	SpAn4DRAFT_2927						
AcsE (CH3-THF methyltransferase)	SpAn4DRAFT_2928						
heterodisulfide reductase	SpAn4DRAFT_2929-31						
methyl-viologen-reducing hydrogenase, delta subunit	SpAn4DRAFT_2932						
zinc-finger protein	SpAn4DRAFT_2933						
MetF	SpAn4DRAFT_2934						
formate dehydrogenase	SpAn4DRAFT_2935-37						
acsB	SpAn4DRAFT_3609						
acetate kinase	SpAn4DRAFT_3911						
formate-THF ligase	SpAn4DRAFT_4515						
MetF	SpAn4DRAFT_5207						
methanol metabolism							
CH3-THF MT	SpAn4DRAFT_1214						
MtaB	SpAn4DRAFT_1215						
MtaC	SpAn4DRAFT_1216						
betaine metabolism							
MttA/MtbA/MtmA	SpAn4DRAFT_0194						
TMA MT	SpAn4DRAFT_0203-05						
betaine reductase component B	SpAn4DRAFT_2117-18						
L-seryl-tRNA(Sec) selenium transferase	SpAn4DRAFT_2119						
glycine/sarcosine/betaine reductase component C	SpAn4DRAFT_2120-21						
selenocysteine-specific translation elongation factor	SpAn4DRAFT_2123						
selenide,water dikinase	SpAn4DRAFT_2124						
glycine betaine transporter OpuD	SpAn4DRAFT_2125						
glycine/sarcosine/betaine reductase protein A	SpAn4DRAFT_2126						
glycine/sarcosine/betaine reductase component C	SpAn4DRAFT_2128-29						
thioredoxin	SpAn4DRAFT_2131						
thioredoxin reductase	SpAn4DRAFT_2132						
TMA MT	SpAn4DRAFT_2138-40						
glycine/sarcosine reductase component B	SpAn4DRAFT_2141-42						
5-CH3THF homocysteine MT	SpAn4DRAFT_2145						
MttC	SpAn4DRAFT_2180						
MttA/MtbA/MtmA	SpAn4DRAFT_2181						
TMA MT	SpAn4DRAFT_2185-86						
N5-CH3-tetrahydromethanopterin coM MT subunit H	SpAn4DRAFT_2187						
MttB	SpAn4DRAFT_2283						
fructose metabolism							
fructose-bisphosphate aldolase	SpAn4DRAFT_0436						
fructose-bisphosphate aldolase	SpAn4DRAFT_0573						
pyruvate kinase	SpAn4DRAFT_0761						
enolase	SpAn4DRAFT_1336						
phosphoglycerate mutase	SpAn4DRAFT_1337						
triosephosphate isomerase	SpAn4DRAFT_1338						
phosphoglycerate kinase	SpAn4DRAFT_1339						
NAD-dependent glyceraldehyde-3-phosphate dehydrogenase	SpAn4DRAFT_1340						
fructose specific PTS system	SpAn4DRAFT_2294-97						
1-phosphofructokinase	SpAn4DRAFT_2298						
transcriptional repressor of the fructose operon	SpAn4DRAFT_2299						
pyruvate-phosphate dikinase	SpAn4DRAFT_5113						
Nitrate respiration							
Nitrate reductase	SpAn4DRAFT_3643-46						
CcmA	SpAn4DRAFT_3648						
CcmC	SpAn4DRAFT_3650						
CcmF	SpAn4DRAFT_3652						
heme biosynthesis protein	SpAn4DRAFT_3653						
nitrite reductase	SpAn4DRAFT_3654						
hydroxylamine reductase	SpAn4DRAFT_3657						
hydroxylamine reductase	SpAn4DRAFT_4403						
NrfA	SpAn4DRAFT_4404						

Legend:	scale
	< 0.05
	0.05 - 0.1
	0.1 - 0.2
	0.2 - 0.4
	0.4 - 0.8
	> 0.8

Fig. 2. Heatmap visualization of the proteomic data of proteins involved in the acetyl-CoA pathway, methanol metabolism, betaine metabolism and the selenocysteine incorporation system, fructose metabolism, and nitrate reduction, nitrite reduction and the biosynthesis of cytochrome c. The figure shows the predicted function of the proteins, the reference to the genome, and their related fractionation values in the five different growth conditions: hydrogen and carbon dioxide (H_2 CO_2), methanol (meoh), methanol and nitrate (NO_3^-), betaine (B), and fructose (F). The fractionation values range between 0 and 1 and are calculated by comparing the peptide abundance of one protein of a condition to the total peptide count of all conditions.

Betaine metabolism

In *Sporomusa* species, betaine is reduced to TMA and acetyl phosphate. The electrons required for this reaction can be generated in the oxidation of methyl groups (Möller *et al*., 1984; Andreesen, 1994). The genome of *Sporomusa* strain An4 contains genes putatively coding for glycine/sarcosine/betaine reductase components A, B and C. The substrate specificity of the three different reductases of *Eubacterium acidaminophilum* is determined by the three different B components (Meyer *et al*., 1995). B components consist of an alpha and a beta subunit in an $\alpha_2\beta_2$ structure. Subunits GrdEB are described to be glycine specific, GrdGF sarcosine specific and GrdIH betaine specific (Meyer *et al*., 1995). We aligned the amino acid sequence of the B components of *E. acidaminophilum* with those of *Sporomusa* strain An4 and constructed a neighbour-joining tree (File S2). The genome of An4 encodes two B components with an alpha and beta subunits. The two subunits of one B component (SpAn4DRAFT_2118 and SpAn4DRAFT_2117) cluster together with the GrdI and GrdH of *E. acidaminophilum*, suggesting specificity for betaine. The other two genes of strain An4 (SpAn4DRAFT_2142 and SpAn4DRAFT_2141) do not cluster with either GrdB and GrdE or GrdF and GrdG, leaving the substrate specificity indistinct.

The proteome results indicated the synthesis of the glycine/sarcosine/betaine reductase components A, B and C by strain An4 especially when grown with betaine (Fig. 2). This included both above-described substrate-specific B components, suggesting that a betaine reductase and a sarcosine or glycine reductase are synthesized by strain An4. Additionally, a thioredoxin involved in these reductase-catalysed reactions was also detected in the proteome (Fig. 2).

The putative A component and the beta subunit of the B components in strain An4 are selenocysteine containing proteins. Selenocysteine incorporation into a protein requires the presence of a selenocysteine incorporation system (Fischer *et al*., 2007). Genes encoding such a system are present in the genome of strain An4. Additionally, the proteins involved in this system are synthesized by strain An4 (Fig. 2).

Many MTs showed high abundance during growth of strain An4 with betaine (Fig. 2). These MTs were predominantly TMA MTs. Two other MTs showed high abundance in betaine-grown cells, tetrahydromethanopterin S-MT subunit H (SpAn4DRAFT_2187) and a CH_3THF MT (SpAn4DRAFT_2145), indicating that they are involved in the betaine metabolism of *Sporomusa* strain An4.

Hydrogen metabolism

The genome of *Sporomusa* strain An4 contains four [FeFe] hydrogenases and three [NiFe] hydrogenases, divided in the following groups: one trimeric NADH-binding (possibly bifurcating as described by Schut and Adams, 2009) [FeFe] hydrogenase (SpAn4DRAFT_2115-2113); one trimeric membrane associated [FeFe] hydrogenase (SpAn4DRAFT_1380-1378), with a cytochrome b as the membrane-associated subunit; two single [FeFe] hydrogenases, one is predicted to be cytoplasmic (SpAn4DRAFT_0657) and the other membrane associated (SpAn4DRAFT_2844); one dimeric membrane-linked [NiFe] hydrogenase (SpAn4DRAFT_2046-2045); one dimeric cytoplasmic [NiFe] hydrogenase (SpAn4DRAFT_0843–0844); and one trimeric membrane-associated hydrogenase (SpAn4DRAFT_3047–3049). The small subunit (SpAn4DRAFT_3049) of this hydrogenase contains a tat signal motif, which suggests transport of the complex across the cytoplasmic membrane. Moreover, the membrane-associated subunit is a putative four transmembrane helix containing cytochrome b. The proteome results showed protein abundance of the NiFe hydrogenase

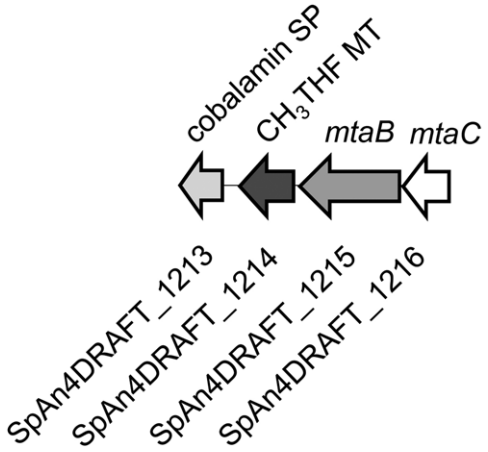

Fig. 3. Methyltransferase genes, mtaB and mtaC, predicted to be involved in methanol-specific methyl transfer of *Sporomusa* strain An4, including a CH_3-THF methyltransferase (CH_3-THF MT) and a cobalamin synthesis protein (cobalamin SP) upstream.

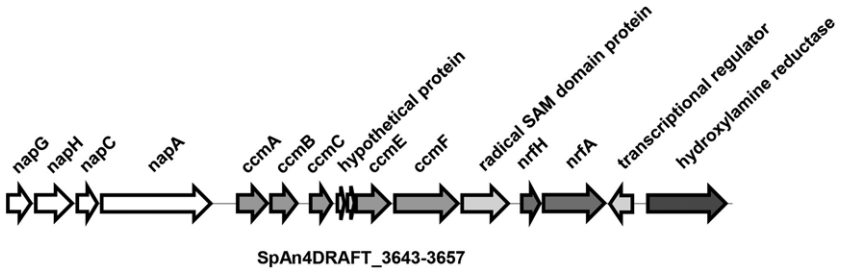

Fig. 4. Relative location of nitrate reduction genes, nitrite reduction genes, c type cytochrome biosynthesis genes and neighbouring genes in the *Sporomusa* strain An4 draft genome.

(SpAn4DRAFT_3047–3049) and NADH-binding [FeFe] hydrogenase (SpAn4DRAFT_2115-2113) in all five growth conditions (File S1). No other hydrogenases were detected in the proteome.

Sugar metabolism

Sporomusa strain An4 was described, like *S. ovata*, to grow with fructose, but not with glucose (Balk *et al*., 2010). The genome of *Sporomusa* strain An4 contains all genes encoding the enzymes required for the conversion of fructose-6-phosphate to acetate. Additionally, the phosphotransferase system (PTS) necessary to transport fructose into the cell and simultaneously phosphorylate the sugar, to fructose-6-phosphate, is also encoded in the genome of strain An4. PTS complexes are also described for sugars other than fructose, for example glucose (Postma *et al*., 1993). The substrate specificity of these complexes is due to the PTS EII proteins. On the basis of sequence alignment, EIIs can be grouped into four classes: the glucose, mannitol/fructose, lactose and mannose classes (Postma *et al*., 1993). The genome of *Sporomusa* strain An4 contains genes encoding EII proteins of the mannitol/fructose and lactose classes. Hence explains why strain An4 cannot grow with glucose. In addition, growth with mannitol or lactose is not possible because of the lack of genes in the genome coding for the enzymes degrading these compounds. Furthermore, no other sugar transporters or degradation enzymes are encoded in the genome of strain An4.

All enzymes necessary for fructose degradation were detected in the proteome analysis (Fig. 2).

Electron acceptor metabolism

Sporomusa strain An4, like *S. ovata*, reduces nitrate via nitrite to ammonium (Balk *et al*., 2010). In the genome of strain An4, genes coding for enzymes of nitrate reduction are present (Fig. 4). The *napGHCA* genes are positioned next to each other. Additionally, the *nrfH* (SpAn4DRAFT_3654) and *nrfA* (SpAn4DRAFT_3655) genes are near the napGHCA cluster. Between the *napGHCA* cluster and the *nrfHA* nitrite reduction genes, the cytochrome c biosynthesis genes, *ccmA* (SpAn4DRAFT_3648), *ccmB* (SpAn4DRAFT_3649), *ccmC* (SpAn4DRAFT_3650), *ccmE* (SpAn4DRAFT_3651), and a *ccmF* gene (SpAn4DRAFT_3652) are positioned. Moreover, a hydroxylamine reductase gene (SpAn4DRAFT_3657) is upstream the *nrfH* and *nrfA* genes.

Another gene known to be involved in cytochrome c biosynthesis, *ccdA* gene, is also present in the genome of strain An4 (SpAn4DRAFT_3036). In addition, gene copies of the *ccmA* (SpAn4DRAFT_0162, SpAn4DRAFT_4410) *ccmB* (SpAn4DRAFT_0161, SpAn4DRAFT_4409), *ccmC* (SpAn4DRAFT_0160, SpAn4DRAFT_4408), *ccmE* (SpAn4DRAFT_0159, SpAn4DRAFT_4407), *ccmF* (SpAn4DRAFT_4406), *nrfH* (SpAn4DRAFT_4405), *nrfA* (SpAn4DRAFT_4404) and hydroxylamine reductase (SpAn4DRAFT_4403) are present in the genome.

Nitrate reductase, nitrite reductase, hydroxylamine reductase and cytochrome c biosynthesis proteins were all detected when strain An4 was grown with nitrate as electron acceptor (Fig. 2, File S1). These proteins are mainly encoded by the genes illustrated in Fig. 4. The function of hydroxylamine reductase in nitrate ammonification is not clear. However, hydroxylamine could be a product of nitrite reduction and be reduced to ammonium by hydroxylamine reductase as described in *Nautilia profundicola* AmH (Hanson *et al*., 2013).

Enzyme activity measurements

Additional enzyme activity measurements were performed on some of the key enzymes to relate the abundances of proteins determined by proteomics to protein function. Enzymes measured were CODH, FDH, nitrate reductase and nitrite reductase. Enzyme activity was measured in cell-free extract from cells grown with H_2 and CO_2, methanol and methanol plus nitrate. Enzyme activities were measured in Table 1, confirming their presence and functionality. Moreover, the nitrate reductase and the nitrite reductase activities were higher in cell-free extracts of cells grown with methanol and nitrate. This is in accordance with the higher peptide abundance of these proteins in the methanol with nitrate condition. Furthermore, the trimeric cytoplasmic FDH (SpAn4DRAFT_2935–2937) is the only FDH present in the proteome of the growth con-

Table 1. Enzyme activity (in U/mg and U = μmol min^{-1}) measured from cell-free extracts of *Sporomusa* strain An4 cells grown with H_2 and CO_2, methanol (meoh), and methanol and nitrate (NO_3^-).

Enzyme name	Reference to proteome/genome	H_2 CO_2	SD	MeOH	SD	NO3–	SD
CODH	CooS/SpAn4DRAFT_2922	1.47 (718)	0.2	1.37 (619)	0.15	0.45 (232)	0.06
FDH	FDH alpha subunit/ SpAn4DRAFT_2937	2.44 (756)	0.18	2.2 (751)	0.2	1.86 (574)	0.15
NO_3^- reductase	NapA/SpAn4DRAFT_3646	0.55 (0)	0.11	0.77 (0)	0.13	5.88 (170)	0.18
NO_2^- reductase	NrfA/SpAn4DRAFT_3655, SpAn4DRAFT_4404	0.37 (5)	0.15	0.51 (1)	0.07	8.1 (177)	0.2

Enzymes measured included carbon monoxide dehydrogenase (CODH), formate dehydrogenase (FDH), nitrate reductase (NO3– reductase) and nitrite reductase (NO2– reductase). The activity values are averages of biological replicates (n3). Standard deviation (SD) between replicates are included. Moreover, reference to the proteome/genome is given, and the proteome data values are included in parentheses.

ditions used for enzyme activity measurements. It is, therefore, likely that the FDH activity is due to the presence of this enzyme (Table 1).

Discussion

Strain An4 was originally isolated with methanol and perchlorate. (Per)chlorate reduction was one of the physiological differences between strain An4 and *S. ovata* strain H1. Unfortunately, the strain had lost its ability to use perchlorate as electron acceptor before we could analyse the proteins involved. In *Alicycliphilus denitrificans* strain BC the chlorate reduction genes are located on a plasmid (Oosterkamp *et al*., 2013). Moreover, Clark and colleagues (2013) describe that chlorate reduction genes are flanked by insertion sequences. This suggests that the ability to reduce perchlorate is not necessarily genetically stable. Apparently, *Sporomusa* strain An4 lost genes involved in perchlorate reduction before the start of our experiments. The deposited DSM strain also was not able to reduce (per)chlorate.

Genome comparison of Sporomusa *strain An4 and* S. ovata *strain H1*

Sporomusa strain An4 and *S. ovata* strain H1 are isogenic and the genome comparison showed only a few differences. The ANI value was above 99% and therefore it was concluded that An4 is a *S. ovata* strain. One of the differences is a multicopper oxidase, encoded by the genome of *s*train An4. Multicopper oxidases can couple the reduction of O_2 to H_2O with substrate oxidation (Rosenzweig and Sazinsky, 2006), but other enzymes involved in superoxide dismutation (MacPherson and Murphy, 2007), manganese oxidation (Butterfield *et al*., 2013), copper oxidation (in copper resistance) (Altimira *et al*., 2012) and nitrite reduction (MacPherson and Murphy, 2007) are also multicopper oxidases. However, none of these known enzymes shows high similarity to the multicopper oxidase of strain An4. Moreover, in strain An4, the proteome indicates the involvement of a cytochrome nitrite reductase and not a copper nitrite reductase. Furthermore, the multicopper oxidase is not detected in the proteome in any of the five tested conditions. Therefore, the function of the multicopper oxidase and why it is present in strain An4 and not in *S. ovata* strain H1 cannot be deduced from our analysis.

Proteogenomic analysis

The proteome analysis resulted in 2280 proteins with 2 or more unique peptides. A total of 1162 of these proteins had a peptide abundance of at least 5. High peptide abundance of a protein verifies synthesis by strain An4. The detected proteins included the enzymes necessary for growth with the selected growth conditions. Insight in the physiology of *Sporomusa* strain An4 was obtained by combining genome analysis with shotgun proteomics data. For example, the results indicate that one FDH (SpAn4DRAFT_2935–2937) is involved in the acetyl-CoA pathway in strain An4. The other FDHs present in the genome might be involved in for example growth with formate as a substrate.

Not all proteins that we expected to detect are present in the proteomics data. This can be due to the difference in the extraction of proteins; cytoplasmic proteins are easier to extract compared with membrane proteins, or to the size difference of proteins. Large proteins in general generate more peptides upon trypsin digestion and are, therefore, easier to detect. The catalytic subunit of nitrate reductase, NapA, obtains its electrons from the cytoplasmic membrane quinol pool by electron transport of either NapC or NapGH. Both NapC and NapGH are quinol dehydrogenases (Brondijk *et al*., 2002, 2004; Kern and Simon, 2008; 2009; Simpson *et al*., 2010) and are present in the genome of strain An4. Although *napC* is present in the genome, its product was not found in the proteomics data. This could mean that either strain An4 only uses the NapGH quinol dehydrogenase complex to transport electrons to nitrate reductase NapA, or the quinol dehydrogenase NapC was not detected because it is a small membrane protein.

In nitrate reduction, c-type cytochromes are important for the translocation of electrons. The biosynthesis of c-type cytochromes, also called the Ccm system, consists

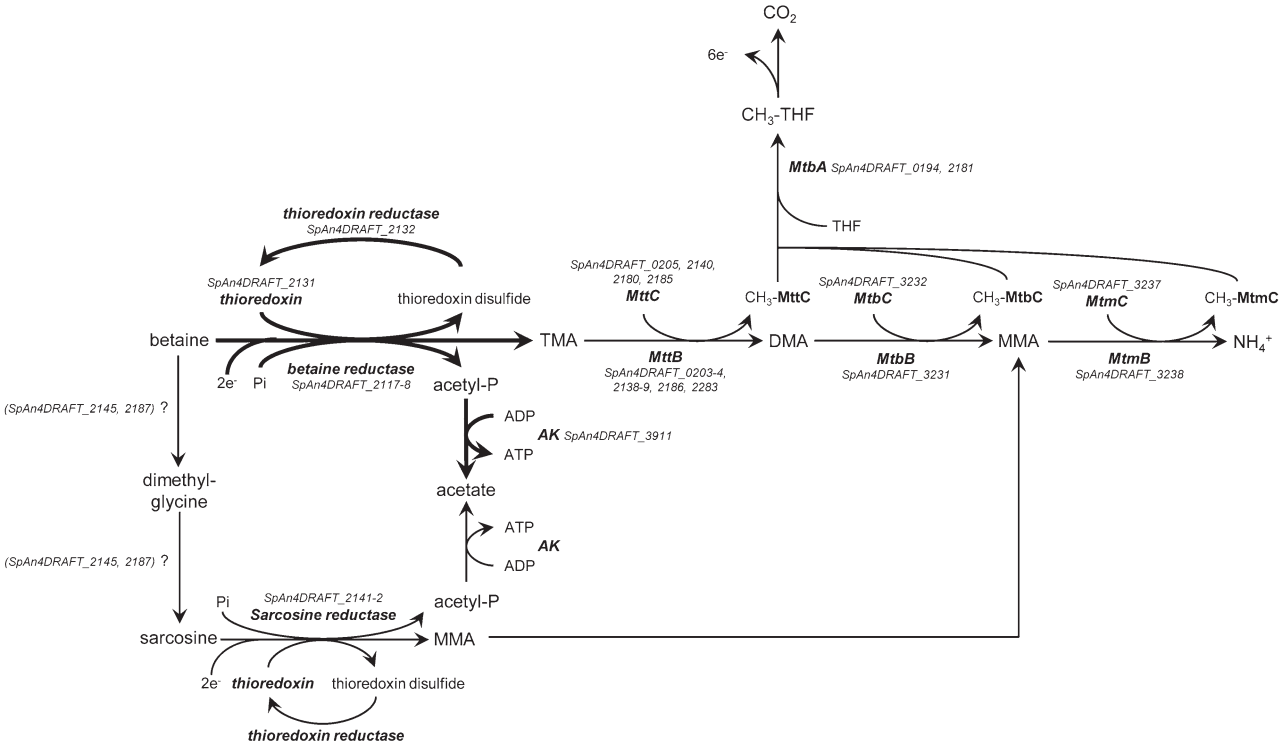

Fig. 5. Proposed betaine degradation pathway in *Sporomusa* strain An4. The thick arrows indicate the main degradation route, while the thin arrows represent reactions that generate electrons for the betaine reductase via the oxidation of methyl groups. Proteins are in bold italic and locus tags are in italic. Abbreviations: AK, acetate kinase; DMA, dimethylamine; e^-, electron; MMA, monomethylamine; Pi, phosphate; TMA, trimethylamine.

of up to 10 components. These components are CcmA to CcmI and CcdA or DsbD (Sanders *et al*., 2010). In *Sporomusa* strain An4, only CcmA, CcmC and CcmF protein levels were detected among the five conditions tested. CcmB, CcmE and CcdA are membrane proteins and therefore might have escaped detection.

Sporomusa strain An4 can grow with both sarcosine and glycine (Balk *et al*., 2010). Therefore, the substrate specificity of one of the highly abundant reductases (SpAn4DRAFT_2142-2141) during growth with betaine is unclear. *Sporomusa ovata* can only utilize sarcosine, not glycine, and produces methylamine, acetate and CO_2, which suggest the involvement of a reductive cleavage reaction. Moreover, Möller and colleagues (1984) suggested that during growth with betaine, a small amount of betaine is demethylated to N,N dimethyl-glycine. Subsequently, this N,N dimethyl-glycine was presumably demethylated further to sarcosine. This, together with the fact that the genome comparison between the two *Sporomusa* strains did not indicate that strain An4 uses a different reaction for sarcosine utilization, makes it likely that strain An4 uses a sarcosine reductase for betaine degradation and not a glycine reductase. Figure 5 shows the possible betaine degradation pathway in strain An4.

The proteome analysis shows high abundance of many MTs, predominantly TMA MTs, during growth with betaine. However, TMA is one of the main end-products of betaine degradation in *Sporomusa* species, including strain An4 (data not shown). The betaine reductase requires electrons that can be generated via the oxidation of methyl groups (Möller *et al*., 1984; Andreesen, 1994). The TMA MTs, the DMA MT and the MMA MT might be involved in the oxidation of small amounts of TMA to NH_4 (Fig. 5). However, this does not explain the many TMA MTs that are synthesized during growth with betaine. We hypothesize that these TMA MTs bind different cobamides. Mok and Taga (2013) showed that in addition to *p*-cresolyl cobamide and phenolyl cobamide *S. ovata* can also produce other cobamides. Moreover, both Chan and Escalante-Semerena (2011) and Mok and Taga (2013) suggested that in *S. ovata*, there could be substrates that require enzymes that bind other cobamides than *p*-cresolyl cobamide and phenolyl cobamide. *Sporomusa ovata* is able to use externally supplied benzimidazoles to form benzimidazolyl cobamides. This inhibits methanol degradation because it interferes with the production of *p*-cresolyl cobamides (Mok and Taga, 2013). However, betaine degradation was not strongly affected by adding benzimidazoles, indicating less dependence on *p*-cresolyl cobamides, which could be explained if the different TMA MTs in *S. ovata* bind different cobamides.

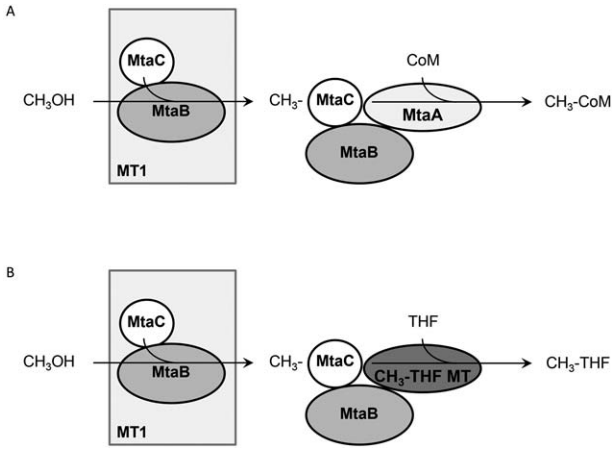

Fig. 6. Methanol conversion as described for a methanol utilizing methanogen (A) and proposed methanol conversion in Sporomusa strain An4 (B). The AcsB subunit of methanol methyltransferase one (MT1) cleaves the C–O bond of methanol and transfers the CH_3 to the MtaC subunit. Subsequently, in the methanogenic pathway MtaA catalyses the transfer of the CH_3 group to coenzyme M (CoM, A). In *Sporomusa* strain An4 the CH_3-THF MT catalyses the transfer of the CH_3 group to THF, creating CH_3-THF (B).

Two other MTs, tetrahydromethanopterin S-MT subunit H and a CH_3THF MT, also showed high abundances in the proteomics results during growth with betaine. The gene coding for the tetrahydromethanopterin S-MT subunit H is also present in the genome of *S. ovata* strain H1, but is designated as an MT 2. MtaA, MtmA, MtbA and MttA MTs are also of the MT 2 type. The tetrahydromethanopterin S-MT subunit H could, therefore, be involved in the methyl transfer from a corrinoid bound TMA MT to THF, or maybe these two MTs that have high abundance in strain An4 during growth with betaine are somehow involved in demethylation of betaine to sarcosine (Fig. 5). Clearly, growth with betaine leads to the synthesis of many MTs in *Sporomusa* strain An4 for which the exact function needs to be elucidated.

Degradation of methanol by *S. ovata* was described before. Stupperich and colleagues (1992) showed that a cobamide-binding methanol-specific MT was expressed by *S. ovata* and they hypothesized that the enzyme was involved in the cleavage of the C–O bond of methanol, the transfer of the CH_3 residue of methanol metabolism or both. Later, the CH_3 group from methanol was found to be transferred to THF, creating CH_3-THF and an entrance into the acetyl-CoA pathway (Stupperich and Konle, 1993). Transfer of the CH_3 residue of methanol in methanogens was shown to be due to three enzyme subunits of the methanogenic methanol : coenzyme M MT system. This system is composed of two MTs (van der Meijden *et al.*, 1983; 1984). Methanol: 5-hydroxybenzimidazolylcobamide MT is the first MT and consists of two subunits, MtaB and MtaC, also known as methanol MT 1 (Sauer *et al.*, 1997). MtaB catalyses the transfer of the CH_3 group from methanol to a corrinoid bound to MtaC. The second MT consists of only one subunit (MtaA), which catalyses the transfer of the CH_3 residue bound to MtaC to coenzyme M (Fig. 6A, van der Meijden *et al.*, 1983; 1984, Harms and Thauer, 1996, LeClerc and Grahame, 1996).

The proteomics results indicate no involvement of a MtaA homologue in methanol degradation by strain An4. Therefore, we propose a novel methanol metabolism without the involvement of an MtaA homologue. Instead of MtaA, the CH_3-THF MT transfers the CH_3 group bound to MtaC to THF (Fig. 6B).

Coenzyme M and THF are structurally different cofactors, hence it is logical that there is a different MT involved in methanogens compared with acetogens, MtaA for binding coenzyme M and CH_3-THF MT for binding THF. Therefore, the methanol metabolism we propose for *Sporomusa* strain An4 could also be the methanol metabolism for other methanol utilizing acetogenic bacteria.

Experimental procedures

Source of inoculum and culture medium

Sporomusa strain An4 (=DSM 21435, =JSM 15643) was isolated from a water sample of an underground gas storage in Russia (Balk *et al.*, 2010). The strain was enriched and isolated with methanol and perchlorate in a basal bicarbonate buffered medium described by Stams and colleagues (1993). The same medium but with the addition of 0.1 g l^{-1} yeast extract was used in this study. The electron donors were added in 15 mM concentrations and electron acceptors in 12 mM concentration, from concentrated stock solutions (sterilized by autoclaving). Cultivation of strain An4 was done at neutral pH and 30°C in 117 ml of glass serum vials with butyl rubber stoppers and aluminium crimp seals. The vials contained 50 ml of basal medium and a gas phase of 1.7 bar N_2/CO_2 (80%/20%, v/v). When hydrogen was the electron donor, the gas phase was 1.7 bar H_2/CO_2 (80%/20%, v/v). In all experiments, the inoculum size was 10%(v/v).

Experimental design

To obtain insight into the physiology of *Sporomusa* strain An4, a combined genome and proteome analysis was performed. For genome analysis, cells were grown with methanol. For proteome analysis, strain An4 was grown at five conditions with the following substrates: (1) hydrogen and carbon dioxide (H_2 and CO_2), (2) methanol, (3) methanol and nitrate, (4) betaine and (5) fructose. Cells were adapted to the different growth conditions by transferring three times. Subsequently, 1.2 l of bottles containing 500 ml of medium were inoculated. Protein extraction of all growth conditions was performed at the late exponential phase.

Analytical methods

Organic compounds were quantified by high pressure liquid chromatography. Methanol, betaine, fructose and acetate

were analysed using a MetaCarb 67H 5 x 300 mm column (Varian, Middelburg, The Netherlands) connected to a SpectraSYSTEM RI-150 detector (Thermo Electron Corporation, Waltham, MA, USA). The 5 mM H_2SO_4 mobile phase had a flow of 0.8 ml min^{-1}. The temperature was controlled at 30°C during the analysis. Crotonate was used as an internal standard.

Nitrate and nitrite were measured using an ED 40 electrochemical detector (Dionex, Breda, The Netherlands) after separation on an Ionpac AS9-SC 4 × 50 mm column (Dionex). Potassium fluoride (2 mM) was used as internal standard. The analysis was conducted at a temperature of 35°C with a flow rate of 1.2 ml min^{-1}. The mobile phase used consisted of 1.29 g l^{-1} Na_2CO_3 10 H_2O and 0.12 g l^{-1} $NaHCO_3$.

Hydrogen was analysed by gas chromatography with a Shimadzu GC-14B (Shimadzu, Kyoto, Japan) equipped with a packed column (Molsieve 13X, 60–80 mesh, 2 m length, 3 mm internal diameter; Varian) and a thermal conductivity detector set at 70 mA. The injector temperature and the oven temperature were both 100°C. The detector temperature was 150°C. Argon was used as the carrier gas at a flow rate of 30 ml min^{-1}.

Growth was measured by following the optical density of cultures at 600 nm (OD_{600}) until the late exponential phase. Uninoculated medium served as a reference.

DNA isolation, genome sequencing and genome annotation

Genomic DNA from *Sporomusa* strain An4 was isolated using the standard Department of Energy Joint Genome Institute (DOE JGI) hexadecyltrimethylammonium bromide (CTAB) method. Sequencing the genome of strain An4 was done by using the 454 pyrosequencing technique and additional mate-pair sequencing (Droege and Hill, 2008). Shot-gun sequencing was performed to a redundancy of 44× (539893 reads). The assembly of these reads was performed using the NEWBLER ASSEMBLER software (454 Life Sciences) resulting in 163 contigs of more than 500 bp. The additional Illumina mate-pair reads were used to extend the contigs and to create scaffolds by using SSPACE BASIC v2.0 (Boetzer *et al*., 2011) with a size insert derived from PICARDTOOLS. Gaps were filled with the mate-pair reads by using SSPACE GAPFILLING v1.11 (Boetzer and Pirovano, 2012). The final assembly was improved with PILON V1.4 using both 454 and mate-pair reads, which resulted in 16 scaffolds of more than 42 kbp. Scaffolds were reordered according to the *S. ovata* strain H1 draft genome. The 16 scaffolds were submitted to RAST server service (Aziz *et al*., 2008) for automatic annotation yielding 5264 protein encoding genes.

Genome comparison

The genome of *Sporomusa* strain An4 was compared with the genome of *S. ovata* strain H1 by analysing their functional domain profiles, which were obtained by first processing both genome assemblies with PRODIGAL (Hyatt *et al*., 2010) and subsequently using INTERPROSCAN 5 (version 5RC7, 27 January 2014). Differences are discussed in the text. The ANI of shared genes was used to determine the phylogenetic positioning of *Sporomusa* strain An4.

Comparison of selected *Sporomusa* strain An4 genes and putative proteins to other microorganisms than *S. ovata* was performed by BLAST analysis. Nucleotide sequences were compared with genes in the NCBI nucleotide collection database. The comparison was optimized for highly similar sequences. Amino acid sequences of selected strain An4 proteins were compared with proteins for homology, using the NCBI non-redundant protein sequence database, with default settings. For alignment, only amino acid sequences were used when the genome of the microorganism was fully sequenced. Additionally, the sequences required a minimum coverage of 80%, and for alignment purposes, the minimum identity was at least 40%. Alignment was performed using MUSCLE (Edgar, 2004).

To identify cofactor-binding motifs, transmembrane helices and twin-arginine translocation motifs in the N-terminus, we used PFAM 27.0 (March 2013) (Punta *et al*., 2012), TMHMM SERVER v. 2.0 (Moller *et al*., 2001) and the TATP 1.0 SERVER (Bendtsen *et al*., 2005) respectively. Sequences with similarity to iron-only or [FeFe]-hydrogenases, were manually analysed for the presence of conserved H-cluster residues (Vignais and Billoud, 2007). RNA loop prediction with MFOLD version 3.2. was used to predict incorporation of selenocysteine (Mathews *et al*., 1999; Zuker, 2003). We compared the predicted RNA loop in the 50–100 bp region downstream of the UGA-codon with the consensus loop described (Zhang and Gladyshev, 2005).

Protein extraction

For the preparation of protein samples 500 ml of each growth condition was centrifuged at 26 000 × *g*. The pellets were resuspended in sodium dodecyl sulfate dithiothreitol Tris (SDT) lysis buffer [100 mM Tris/HCl pH 7.6 + 4% SDS + 0.1M dithiothreitol (DTT)] and sonicated to disrupt the bacterial cell wall. Unbroken cells and debris were removed by centrifugation at 16 000 × *g* for 10 min. Protein concentration in the samples was measured by using Bradford reagent (Sigma Aldrich, St. Louis, MO, USA) using bovine serum albumin as a standard.

Proteome analysis

For each growth condition, an equal amount of total protein was separated by SDS-PAGE on a 10-well PAGE® Novex 4–12% Bis-Tris gel (Invitrogen) and ran using 2-(N-morpholino)ethanesulfonic acid (MES)-SDS as running buffer for 30 min at a constant voltage of 200 V. The gel was stained with Coomassie Blue (Colloidal Blue Staining Kit, Invitrogen) after which, the entire gel lanes were cut into 25 equal slices using a gridcutter (Gelcompany, SA, USA). Gel pieces were reduced with DTT (10 mM, 30 min, room temperature), alkylated with iodoacetamide (20 mM, 60 min, room temperature in the dark) and digested with trypsin overnight at 37°C. After digestion, formic acid and dimethylsulphoxide were added (both 5% v/v) to increase peptide recovery. Protein digests were analysed on a reversed phase nano-liquid chromatography (LC) coupled to an LTQ Orbitrap Velos (Thermo Fisher Scientific, Bremen, Germany). An Agilent 1200 series high-performance liquid chromatography system was equipped with in-house packed trapping column (100 μm ID and 20 mm length) and analytical column (50 μm ID and 250 mm length) filled with Reprosil

Pur 120 C18-AQ (Dr. Maisch, Ammerbuch-Entringen, Germany) essentially as described by Meiring and colleagues (2002). Trapping was performed at 5 μl/min for 10 min in solvent A (0.1 M acetic acid), and elution was achieved with a gradient from 0% to 40% solvent B (0.1 M acetic acid in 8:2 v/v acetonitrile : water) for 40 min. The LTQ Orbitrap Velos was operated in data-dependent mode automatically switching between mass spectrometry (MS) and MS/MS. Survey full scan MS spectra were acquired from *m/z* 400 to 1500 in the Orbitrap with a resolution of 30 000 at *m/z* 400 after accumulation to a target value of 1×10^6 in the linear ion trap. The 10 most intense, multiply charged ions at a threshold of above 1000 were fragmented in the linear ion trap using collision-induced dissociation at a target value of 1×10^4. All raw data files were processed into peaklists using PROTEOME DISCOVERER 1.1.

For each condition approximately 80 000 to 100 000 MS/MS spectra were obtained from trypsin-digested protein mixtures of cell-free extracts. The MS/MS spectra were submitted to a local implementation of the OMSSA search engine (Geer *et al*., 2004) and searched against a peptide database derived from the predicted proteins of the *Sporomusa* strain An4 draft genome RAST annotation. The following search parameters were used: a precursor ion tolerance of 0.03 Da, fragment ion tolerance of 0.5 Da, a miss cleavage allowance of up to and including 2. Cysteines were considered to be carboxyamidomethylated, oxidation of methionine, and deamination of glutamine and asparagine were treated as variable modifications. The set E-value threshold was determined iteratively from the false discovery rate (FDR) and was set to 0.01. With this setting, an FDR of <2% was obtained for all samples. The FDR was calculated from top hit spectral matches to peptides in the reversed database as described by Elias and Gygi (2007). The proteome analysis resulted in the identification of 2280 different proteins (with at least 2 unique peptides identified). Spectral counting (Liu *et al*., 2004) was used to study protein abundance. Higher counts lower the chance of false positives, making it more reliable that the protein is synthesized. A total of 1162 predicted proteins showed >4 counts and are listed in File S1. Proteins that possibly play an important role under different growth conditions are discussed in more detail in the text.

Enzyme activity assays

Enzyme activities of CODH, FDH, nitrate reductase and nitrite reductase were determined spectrophotometrically by measuring the colour change of methyl viologen (MV) at 578 nm and 37°C, due to reduction or oxidation. Glass cuvettes were anaerobically prepared by flushing with N_2, except for the CODH assay, where the glass cuvette was flushed with CO. The CODH assay mixture contained 50 mM 3-(N-morpholino)propanesulfonic acid-KOH (pH 7), 1 mM MV and 0.1 mM DTT. The FDH assay mixture contained Tris-HCl (pH 8), 1 mM MV, 0.1 mM DTT and 10 mM formate. In the CODH and FDH assay, the absorbance increased.

The nitrate and nitrite reductase assay mixture both contained 50 mM Tris-HCl (pH 7.5), 0.5 mM MV, a small amount of dithionite solution to reach a starting absorbance of approximately 1.5, and 4 mM nitrate or 0.25 mM nitrite respectively.

Acknowledgements

Research was funded by grants of the Division of Chemical Sciences (CW-TOP 700.55.343) and Earth and Life Sciences (ALW 819.02.014) and the 2008 Spinoza Award of The Netherlands Organisation for Scientific Research (NWO), and of The Netherlands Proteomics Centre embedded in The Netherlands Genomics Initiative. A.J.M. Stams and W.M. de Vos acknowledge the financial support of ERC grants (323009 and 250172) and the Gravitation grant (024.002.002) of the Netherlands Ministry of Education, Culture and Science.

References

Altimira, F., Yáñez, C., Bravo, G., González, M., Rojas, L.A., and Seeger, M. (2012) Characterization of copper-resistant bacteria and bacterial communities from copper-polluted agricultural soils of central Chile. *BMC Microbiol* **12:** 193–204.

Andreesen, J.R. (1994) Glycine metabolism in anaerobes. *Antonie Van Leeuwenhoek* **66:** 223–237.

Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., *et al*. (2008) The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* **9:** 75.

Balk, M., Mehboob, F., van Gelder, A.H., Rijpstra, I.C., Sinninghe Damsté, J.S., and Stams, A.J.M. (2010) (Per)chlorate reduction by an acetogenic bacterium, *Sporomusa* sp., isolated from an underground gas storage. *Appl Microbiol Biotechnol* **88:** 595–603.

Bendtsen, J.D., Nielsen, H., Widdick, D., Palmer, T., and Brunak, S. (2005) Prediction of twin-arginine signal peptides BMC. *Bioinformatics* **6:** 167.

Boetzer, M., and Pirovano, W. (2012) Toward almost closed genomes with GapFiller. *Genome Biol* **13:** R56.

Boetzer, M., Henkel, C.V., Jansen, H.J., Butler, D., and Pirovano, W. (2011) Scaffolding preassembled contigs using SSPACE. *Bioinformatics* **27:** 578–579.

Boga, H.I., Ludwig, W., and Brune, A. (2003) *Sporomusa aerivorans* sp. nov., an oxygen-reducing homoacetogenic bacterium from the gut of a soil-feeding termite. *Int J Syst Evol Microbiol* **53:** 1397–1404.

Breznak, J.A., and Switzer, J.M. (1986) Acetate synthesis from H_2 plus CO_2 by termite gut microbes. *Appl Environ Microbiol* **52:** 623–630.

Breznak, J.A., Switzer, J.M., and Seitz, H.J. (1988) *Sporomusa termitida* sp. nov., an H_2/CO_2-utilizing acetogen isolated from termites. *Arch Microbiol* **150:** 282–288.

Brondijk, T.H., Fiegen, D., Richardson, D.J., and Cole, J.A. (2002) Roles of NapF, NapG and NapH, subunits of the *Escherichia coli* periplasmic nitrate reductase, in ubiquinol oxidation. *Mol Microbiol* **44:** 245–255.

Brondijk, T.H., Nilavongse, A., Filenko, N., Richardson, D.J., and Cole, J.A. (2004) NapGH components of the periplasmic nitrate reductase of *Escherichia coli* K-12: location, topology and physiological roles in quinol oxidation and redox balancing. *Biochem J* **379:** 47–55.

Butterfield, C.N., Soldatova, A.V., Lee, S.W., Spiro, T.G., and Tebo, B.M. (2013) Mn(II,III) oxidation and MnO_2 mineralization by an expressed bacterial multicopper oxidase. *Proc Natl Acad Sci USA* **110:** 11731–11735.

Chan, C.H., and Escalante-Semerena, J.C. (2011) ArsAB, a novel enzyme from *Sporomusa ovata* activates phenolic bases for adenosylcobamide biosynthesis. *Mol Microbiol* **81:** 952–967.

Clark, I.C., Melnyk, R.A., Engelbrektson, A., and Coates, J.D. (2013) Structure and evolution of chlorate reduction composite transposons. *MBio* **4:** e00379-13.

Das, A., Fu, Z.Q., Tempel, W., Liu, Z.J., Chang, J., Chen, L., *et al*. (2007) Characterization of a corrinoid protein involved in the C1 metabolism of strict anaerobic bacterium *Moorella thermoacetica*. *Proteins* **67:** 167–176.

Dehning, I., Stieb, M., and Schink, B. (1989) *Sporomusa malonica* sp. nov., a homoacetogenic bacterium growing by decarboxylation of malonate or succinate. *Arch Microbiol* **151:** 421–426.

Dobrindt, U., and Blaut, M. (1996) Purification and characterization of a membrane-bound hydrogenase from *Sporomusa sphaeroides* involved in energy-transducing electron transport. *Arch Microbiol* **165:** 141–147.

Drake, H.L., Daniel, S.L., Küsel, K., Matthies, C., Kuhner, C., and Braus-Stromeyer, S. (1997) Acetogenic bacteria: what are the in situ consequences of their diverse metabolic versatilities? *Biofactors* **6:** 13–24.

Drake, H.L., Gößner, A.S., and Daniel, S.L. (2007) Old acetogens, new light. *Ann N Y Acad Sci* **1125:** 100–128.

Droege, M., and Hill, B. (2008) The Genome Sequencer FLX System-longer reads, more applications, straight forward bioinformatics and more complete data sets. *J Biotechnol* **136:** 3–10.

Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32:** 1792–1797.

Elias, J.E., and Gygi, S.P. (2007) Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat Methods* **4:** 207–214.

Fischer, N., Paleskava, A., Gromadski, K.B., Konevega, A.L., Wahl, M.C., Stark, H., and Rodnina, M.V. (2007) Towards understanding selenocysteine incorporation into bacterial proteins. *Biol Chem* **388:** 1061–1067.

Geer, L.Y., Markey, S.P., Kowalak, J.A., Wagner, L., Xu, M., Maynard, D.M., *et al*. (2004) Open mass spectrometry search algorithm. *J Proteome Res* **3:** 958–964.

Hanson, T.E., Campbell, B.J., Kalis, K.M., Campbell, M.A., and Klotz, M.G. (2013) Nitrate ammonification by *Nautilia profundicola* AmH: experimental evidence consistent with a free hydroxylamine intermediate. *Front Microbiol* **4:** 180.

Harms, U., and Thauer, R.K. (1996) Methylcobalamin: coenzyme M methyltransferase isoenzymes MtaA and MtbA from *Methanosarcina barkeri*. Cloning, sequencing and differential transcription of the encoding genes, and functional overexpression of the mtaA gene in *Escherichia coli*. *FEBS J* **235:** 653–659.

Hermann, M., Popoff, M.R., and Sebald, M. (1987) *Sporomusa paucivorans* sp. nov., a methylotrophic bacterium that forms acetic acid from hydrogen and carbon dioxide. *Int J Syst Bacteriol* **37:** 93–101.

Hyatt, D., Chen, G.L., Locascio, P.F., Land, M.L., Larimer, F.W., and Hauser, L.J. (2010) Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11:** 119.

Kamlage, B., and Blaut, M. (1993) Isolation of a cytochrome-deficient mutant strain of *Sporomusa sphaeroides* not capable of oxidizing methyl groups. *J Bacteriol* **175:** 3043–3050.

Kern, M., and Simon, J. (2008) Characterization of the NapGH quinol dehydrogenase complex involved in *Wolinella succinogenes* nitrate respiration. *Mol Microbiol* **69:** 1137–1152.

Kern, M., and Simon, J. (2009) Periplasmic nitrate reduction in *Wolinella succinogenes*: cytoplasmic NapF facilitates NapA maturation and requires the menaquinol dehydrogenase NapH for membrane attachment. *Microbiology* **155:** 2784–2794.

Krätzer, C., Carini, P., Hovey, R., and Deppenmeier, U. (2009) Transcriptional profiling of methyltransferase genes during growth of *Methanosarcina mazei* on trimethylamine. *J Bacteriol* **191:** 5108–5115.

LeClerc, G.M., and Grahame, D.A. (1996) Methylcobamide : coenzyme M methyltransferase isozymes from *Methanosarcina barkeri*. Physicochemical characterization, cloning, sequence analysis, and heterologous gene expression. *Eur J Biochem* **271:** 18725–18731.

Liu, H., Sadygov, R.G., and Yates, J.R. (2004) A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Anal Chem* **76:** 4193–4201.

MacPherson, I.S., and Murphy, M.E.P. (2007) Type-2 copper-containing enzymes. *Cell Mol Life Sci* **64:** 2887–2899.

Mathews, D.H., Sabina, J., Zuker, M., and Turner, D.H. (1999) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J Mol Biol* **288:** 911–940.

van der Meijden, P., Heythuysen, H.J., Pouwels, A., Houwen, F., van der Drift, C., and Vogels, G.D. (1983) Methyltransferases involved in methanol conversion by *Methanosarcina barkeri*. *Arch Microbiol* **134:** 238–242.

van der Meijden, P., te Brömmelstroet, B.W., Poirot, C.M., van der Drift, C., and Vogels, G.D. (1984) Purification and properties of methanol: 5-hydroxybenzimidazolylcobamide methyltransferase from *Methanosarcina barkeri*. *J Bacteriol* **160:** 629–635.

Meiring, H.D., van der Heeft, E., and ten Hove, G.J. (2002) Nanoscale LC-MS(n): technical design and applications to peptide and protein analysis. *J Sep Sci* **25:** 557–568.

Meyer, M., Granderath, K., and Andreesen, J.R. (1995) Purification and characterization of protein P, of betaine reductase and its relationship to the corresponding proteins glycine reductase and sarcosine reductase from *Eubacterium acidaminophilum*. *Eur J Biochem* **234:** 184–191.

Mok, K.C., and Taga, M.E. (2013) Growth inhibition of *Sporomusa ovata* by incorporation of benzimidazole bases into cobamides. *J Bacteriol* **195:** 1902–1911.

Moller, S., Croning, M.D.R., and Apweiler, R. (2001) Evaluation of methods for the prediction of membrane spanning regions. *Bioinformatics* **17:** 646–653.

Möller, B., Oβmer, R., Howard, B.H., Gottschalk, G., and Hippe, H. (1984) *Sporomusa*, a new genus of gram-negative anaerobic bacteria including *Sporomusa sphaeroides* spec. nov. and *Sporomusa ovata* spec. nov. *Arch Microbiol* **139:** 388–396.

Ollivier, B., Cordruwisch, R., Lombardo, A., and Garcia, J.L. (1989) Isolation and characterization of *Sporomusa acidovorans* sp. nov., a methylotrophic homoacetogenic bacterium. *Arch Microbiol* **142:** 307–310.

Oosterkamp, M.J., Veuskens, T., Talarico Saia, F., Weelink, S.A., Goodwin, L.A., Daligault, H.E., *et al*. (2013) Genome analysis and physiological comparison of *Alicycliphilus denitrificans* strains BC and K601T. *PLoS ONE* **8:** e66971.

Pierce, E., Xie, G., Barabote, R.D., Saunders, E., Han, C.S., Detter, J.C., *et al*. (2008) The complete genome sequence of *Moorella thermoacetica* (f. *Clostridium thermoaceticum*). *Environ Microbiol* **10:** 2550–2573.

Poehlein, A., Gottschalk, G., and Daniel, R. (2013) First Insights into the genome of the Gram-negative, endospore-forming organism *Sporomusa ovata* strain H1 DSM 2662. *Genome Announc* **1:** e00734-13.

Postma, P.W., Lengeler, J.W., and Jacobson, G.R. (1993) Phosphoenolpyruvate: carbohydrate phosphotransferase systems of bacteria. *Microbiol Rev* **57:** 543–594.

Punta, M., Coggill, P.C., Eberhardt, R.Y., Mistry, J., Tate, J., Boursnell, C., *et al*. (2012) The Pfam protein families database. *Nucleic Acids Res* **40:** D290–D301.

Richter, M., and Rossello-Mora, R. (2009) Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci USA* **106:** 19126–19131.

Rosenzweig, A.C., and Sazinsky, M.H. (2006) Structural insights into dioxygen-activating copper enzymes. *Curr Opin Struct Biol* **16:** 729–735.

Sanders, C., Turkarslan, S., Lee, D.W., and Daldal, F. (2010) Cytochrome c biogenesis: the Ccm system. *Trends Microbiol* **18:** 266–274.

Sass, H., Wieringa, E., Cypionka, H., Babenzien, H.D., and Overmann, J. (1998) High genetic and physiological diversity of sulfate-reducing bacteria isolated from an oligotrophic lake sediment. *Arch Microbiol* **170:** 243–251.

Sauer, K., Harms, U., and Thauer, R.K. (1997) Methanol:coenzyme M methyltransferase from *Methanosarcina barkeri*. Purification, properties and encoding genes of the corrinoid protein MT1. *Eur J Biochem* **243:** 670–677.

Schut, G.J., and Adams, M.W. (2009) The iron-hydrogenase of *Thermotoga maritima* utilizes ferredoxin and NADH synergistically: a new perspective on anaerobic hydrogen production. *J Bacteriol* **191:** 4451–4457.

Simpson, P.J., Richardson, D.J., and Codd, R. (2010) The periplasmic nitrate reductase in *Shewanella*: the resolution, distribution and functional implications of two NAP isoforms, NapEDABC and NapDAGHB. *Microbiology* **156:** 302–312.

Stams, A.J.M., Van Dijk, J.B., Dijkema, C., and Plugge, C.M. (1993) Growth of syntrophic propionate-oxidizing bacteria with fumarate in the absence of methanogenic bacteria. *Appl Environ Microbiol* **59:** 1114–1119.

Stupperich, E., and Konle, R. (1993) Corrinoid-dependent methyl transfer reactions are involved in methanol and 3,4-Dimethoxybenzoate metabolism by *Sporomusa ovata*. *Appl Environ Microbiol* **59:** 3110–3116.

Stupperich, E., Eisinger, H.J., and Krautler, B. (1988) Diversity of corrinoids in acetogenic bacteria. p-cresolylcobamide from *Sporomusa ovata*, 5-methoxy-6-methylbenzimidazolylcobamide from *Clostridium formicoaceticum* and vitamin B12 from *Acetobacterium woodii*. *Eur J Biochem* **172:** 459–464.

Stupperich, E., Eisinger, H.J., and Krautler, B. (1989) Identification of phenolyl cobamide from the homoacetogenic bacterium *Sporomusa ovata*. *Eur J Biochem* **186:** 657–661.

Stupperich, E., Eisinger, H.J., and Albracht, S.P. (1990) Evidence for a super-reduced cobamide as the major corrinoid fraction in vivo and a histidine residue as a cobalt ligand of the p-cresolyl cobamide in the acetogenic bacterium *Sporomusa ovata*. *Eur J Biochem* **193:** 105–109.

Stupperich, E., Aulkemeyer, P., and Eckerskorn, C. (1992) Purification and characterization of a methanol-induced cobamide-containing protein from *Sporomusa ovata*. *Arch Microbiol* **158:** 370–373.

Vignais, P.M., and Billoud, B. (2007) Occurrence, classification, and biological function of hydrogenases: an overview. *Chem Rev* **107:** 4206–4272.

Wagner, U.G., Stupperich, E., Aulkemeyer, P., and Kratky, C. (1994) Crystallization and preliminary X-ray diffraction studies of a corrinoid protein from *Sporomusa ovata*. *J Mol Biol* **236:** 388–389.

Wagner, U.G., Stupperich, E., and Kratky, C. (2000) Structure of the molybdate/tungstate binding protein mop from *Sporomusa ovata*. *Structure* **8:** 1127–1136.

Wang, S., Huang, H., Kahnt, J., and Thauer, R.K. (2013) *Clostridium acidurici* electron-bifurcating formate dehydrogenase. *Appl Environ Microbiol* **79:** 6176–6179.

Zhang, Y., and Gladyshev, V.N. (2005) An algorithm for identification of bacterial selenocysteine insertion sequence elements and selenoprotein genes. *Bioinformatics* **21:** 2580–2589.

Zhou, W., Das, A., Habel, J.E., Liu, Z.J., Chang, J., Chen, L., *et al*. (2005) Isolation, crystallization and preliminary X-ray analysis of a methanol-induced corrinoid protein from *Moorella thermoacetica*. *Acta Crystallogr F Struct Biol Cryst Commun* **61:** 537–540.

Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* **31:** 3406–3415.

Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

File S1. Proteomic data of the proteins with peptide counts >4 of Sporomusa strain An4. The supporting information shows the predicted function of the proteins, the reference to the genome, and their related peptide abundance in the five different growth conditions: hydrogen and carbon dioxide (H_2 CO_2), methanol (MeOH), methanol and nitrate (NO3-), betaine (B), and fructose (F).

File S2. Neighbor joining tree constructed by using the amino acid sequences of the glycine/sarcosine/betaine reductase B components of Eubacterium acidaminophilum and Sporomusa strain An4. The B components have an alpha and a beta subunit. The numbers at the branches are bootstrap percentage values. The scale bar represents 20% sequence difference.