Molecular Microbiology (2014) 94(6), 1330–1342 ■
doi:10.1111/mmi.12832
First published online 6 November 2014

Novel proteins for homocysteine biosynthesis in anaerobic microorganisms

Benjamin Julius Rauch,[1] Andrew Gustafson[2] and John J. Perona[1,2]*
[1]*Department of Biochemistry and Molecular Biology, Oregon Health and Science University, 3181 S.W. Sam Jackson Park Road Portland, OR 97239-3098, USA.*
[2]*Department of Chemistry, Portland State University, 1719 SW 10th Avenue, Portland, OR 97201, USA.*

Summary

The metabolic network for sulfide assimilation and trafficking in methanogens is largely unknown. To discover novel proteins required for these processes, we used bioinformatics to identify genes co-occurring with the protein biosynthesis enzyme SepCysS, which converts phosphoseryl-tRNA^{Cys} to cysteinyl-tRNA^{Cys} in nearly all methanogens. Exhaustive analysis revealed three conserved protein families, each containing molecular signatures predicting function in sulfur metabolism. One of these families, classified within clusters of orthologous groups (COG) 1900, possesses two conserved cysteine residues and is often found in genomic contexts together with known sulfur metabolic genes. A second protein family is predicted to bind two 4Fe-4S clusters. All three genes were also identified in more than 50 strictly anaerobic bacterial genera from nine distinct phyla. Gene-deletion and growth experiments in *Methanosarcina acetivorans*, using sulfide as the sole sulfur source, demonstrate that two of the proteins (MA1821 and MA1822) are essential to homocysteine biosynthesis in a background lacking an additional gene for sulfur insertion into homocysteine. Mutational analysis confirms the importance of several structural elements, including a conserved cysteine residue and the predicted 4Fe-4S cluster-binding domain.

Introduction

Sulfur assimilation in aerobic organisms occurs by reduction of sulfate, which is taken up into cells and converted to sulfide via a highly conserved, ATP-dependent four-step pathway (Sekowska *et al.*, 2000). The sulfide is then incorporated into homocysteine (Hcy) for methionine (Met) and *S*-adenosylmethionine (SAM) biosynthesis (in aerobes possessing the OAHS enzyme; Fig. 1A), and into cysteine (Cys) for dispersal into a variety of other metabolites. This trafficking proceeds by persulfide (R-S-SH) relay initiated by cysteine desulfurases (CD), (Hidese *et al.*, 2011) which remove the sulfur from free Cys, yielding alanine and a persulfided Cys residue on the enzyme. The terminal sulfur of the CD-bound persulfide group is then relayed as sulfane (S^0) to conserved Cys residues on other proteins via a series of nucleophilic substitution reactions. By this process sulfur is ultimately inserted into a broad variety of compounds, including enzyme cofactors, ubiquitin-like proteins, and RNA modifications (Fig. 1A) (Mueller, 2006). In many cases the mobile sulfane is reduced to sulfide prior to insertion.

By contrast, sulfate levels in the ancient anaerobic biosphere were much lower than they are today, reflecting the suppressed weathering of sulfides in the absence of high oxygen levels (Habicht *et al.*, 2002). Therefore, in the Archaean Earth sulfur was probably assimilated into biosynthetic pathways by quite different processes (Liu *et al.*, 2012a). These processes may persist in contemporary methanogens, which inhabit anaerobic environments and utilize sulfide rather than sulfate as a sulfur source. The key enzymes of the methanogenesis pathway are of a single origin, were vertically inherited from an ancestral euryarchaeote, and require unique sulfur-containing cofactors (Thauer, 1998; Blank, 2009), suggesting that ancient mechanisms of sulfur assimilation and trafficking may also be preserved.

In methanogens, the canonical enzymes responsible for the formation of free Cys, Hcy, cysteinylated tRNA (Cys-tRNA^{Cys}) and persulfide are often absent, implying the existence of novel alternative pathways (Fig. 1A). A clear first example of such a pathway was provided by the discovery of a unique mechanism for cysteinylation of tRNA, which is conserved in nearly all methanogens and in the related strictly anaerobic *Archaeoglobi* (Fig. 1A and B) (Sauerwald *et al.*, 2005). In these organisms, tRNA is first aminoacylated with phosphoserine (Sep) by phosphoseryl-tRNA synthetase (SepRS). Next, the resulting Sep-tRNA^{Cys} intermediate is converted to Cys-tRNA by Sep-tRNA:Cys-tRNA synthase (SepCysS). The canonical

Accepted 12 October, 2014. *For correspondence. E-mail perona@pdx.edu; Tel. (+1) 503 725 2426; Fax (+1) 503 725 2426.

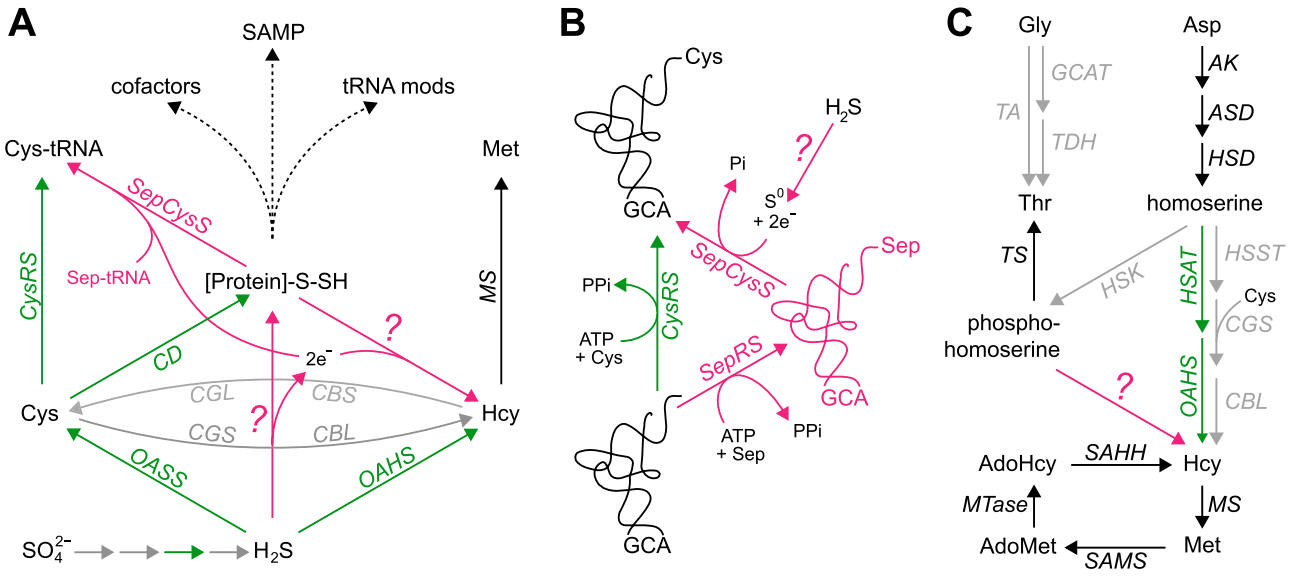

Fig. 1. A. Sulfur assimilation and trafficking pathways in *Methanosarcina acetivorans*. Known and proposed reactions representative of the ancient euryarchaeal metabolic strategy (pink) are presumed to occur in all genomes also encoding SepCysS. These include Sep-tRNA:Cys-tRNA synthase (SepCysS) and proposed sulfide-dependent persulfide and Hcy biosynthesis enzymes (question marks indicate that the existence of the pathways is not established). Known reactions representative of contemporary metabolic strategies, present in *M. acetivorans* (green), are often absent from other SepCysS-encoding genomes. These include cysteinyl-tRNA synthetase (CysRS), cysteine desulfurase (CD), *O*-acetylserine sulfhydrylase (OASS), and *O*-acetylhomoserine sulfhydrylase (OAHS). Known reactions representative of contemporary metabolic strategies, absent from *M. acetivorans* and some other SepCysS-encoding genomes, are indicated in gray. These include the transulfuration pathways for interconversion of Cys and Hcy, comprising the enzymes cystathionine gamma-synthase (CGS), cystathionine beta-lyase (CBL), cystathionine beta-synthase (CBS) and cystathionine gamma-lyase (CGL). Known enzymes common to both strategies are depicted in black (methionine synthase; MS). Protein-bound persulfide ([Protein]-SSH) likely provides sulfur for the biosyntheses of cofactors, tRNA modifications, and the archaeal modifying protein and sulfur carrier (SAMP).
B. Pathways for synthesis of Cys-tRNA^{Cys}. All SepCysS-encoding genomes also include the gene for SepRS, while CysRS is often absent. The identity of a presumed protein sulfur donor generating a persulfide group on SepCysS is unknown. Color-coding is as specified in (A).
C. Pathways of carbon flow relevant to sulfide assimilation. Additional enzymes depicted include aspartate kinase (AK), aspartate semialdehyde dehydrogenase (ASD), homoserine dehydrogenase (HSD), homoserine succinyltransferase (HSST), homoserine kinase (HSK), homoserine acetyltransferase (HSAT), *S*-adenosylmethionine synthetase (SAMS), AdoMet-dependent methyltransferases (MTase), *S*-adenosylhomocysteine hydrolase (SAHH), threonine aldolase (TA), glycine C-acetyltransferase (GCAT), threonine dehydrogenase (TDH) and threonine synthase (TS). Color coding is as depicted in panel (A). An HSK homolog cannot be detected in the *M. acetivorans* genome, yet the functioning of TS with a phosphohomoserine substrate may represent the only route to Thr in this organism. Therefore, a highly diverged HSK-like activity may be present.

cysteinyl-tRNA synthetase (CysRS) is absent in many methanogens, so this two-step pathway represents the only route to Cys-tRNA^{Cys}. Further, deletion of the gene encoding SepRS from *Methanococcus maripaludis* (which also encodes CysRS) results in cysteine auxotrophy. This indicates that the SepRS–SepCysS pathway, in addition to providing Cys for protein synthesis, is also the only means for *de novo* Cys biosynthesis in this organism (Sauerwald *et al.*, 2005).

The pyridoxal phosphate (PLP)-dependent mechanism for conversion of Sep-tRNA^{Cys} to Cys-tRNA^{Cys} by SepCysS requires attack of nucleophilic sulfide on the β-carbon of the phosphoserine side-chain, with release of inorganic phosphate. *In vitro* reconstitution of the reaction with sodium sulfide as sulfur donor is highly inefficient, suggesting that dissolved sulfide is unlikely to be the native substrate (Sauerwald *et al.*, 2005; Hauenstein and Perona, 2008). When expressed in *Escherichia coli*, SepCysS is able to accept relayed persulfide as a sulfur donor (Helgadóttir *et al.*, 2012; Liu *et al.*, 2012b), but the mechanism of persulfide synthesis and delivery to SepCysS in methanogens remain unknown (Sauerwald *et al.*, 2005; Hauenstein and Perona, 2008; Yuan *et al.*, 2010; Helgadóttir *et al.*, 2012; Liu *et al.*, 2012b). Since sulfur insertion reactions typically require nucleophilic sulfide instead of electrophilic sulfane (Mueller, 2006), SepCysS may use additional conserved Cys residues in its active site to form a disulfide bond with the persulfided Cys residue, thereby liberating the sulfide required for catalysis (Hauenstein and Perona, 2008; Liu *et al.*, 2012b).

The IscS-type CD from *E. coli* relays persulfide to SepCysS *in vitro* (Liu *et al.*, 2012b). However, a CD is unlikely to be the ancestral sulfur source for SepCysS because (i) known CD genes are often completely absent from the genomes of methanogens (Fig. 1), (ii) genes responsible for the synthesis of free Cys are also often absent from methanogen genomes, consistent with the finding that the SepRS–SepCysS pathway provides the

sole *de novo* Cys biosynthesis pathway in *M. maripaludis* (Sauerwald *et al*., 2005), (iii) the intracellular pool of free Cys in *M. maripaludis* is five to 10-fold lower than in aerobic bacteria (Liu *et al*., 2010), (iv) Cys is not the sulfur source for either iron-sulfur cluster or methionine biosynthesis in *M. maripaludis* (Liu *et al*., 2010), and (v) sulfur donation from CD to SepCysS would represent a fruitless cycle because it would require Cys as a source for its own biosynthesis (Mueller, 2006). Unlike in aerobes, the low concentration of free Cys in methanogens suggests that this metabolite does not play a central role as a metabolic hub for anabolic sulfur distribution. Nonetheless, the presence of SepCysS suggests that an alternative means for initiating persulfide relays exists. Persulfide groups have been identified in several *M. maripaludis* proteins involved in the biosynthesis of thiolated tRNA nucleosides (Liu *et al*., 2012c; 2014).

The ancestral methanogen pathway for Hcy biosynthesis is also unknown. Neither the transsulfuration pathway from Cys nor the OAHS enzyme for Hcy formation from sulfide are present in many of the contemporary methanogens, including *M. maripaludis* (Fig. 1A and C; Table S1) (Liu *et al*., 2010). Interestingly, neither sulfide nor cysteine serves as the immediate sulfur donor for Met biosynthesis in *M. maripaludis* (Liu *et al*., 2010). Similarly, neither sulfide nor Fe-S clusters are the immediate sulfur donor for Hcy biosynthesis in *Methanocaldococcus jannaschii* (White, 2003). Provocatively, however, the latter study showed that the cellular persulfide pool was sufficiently large to support Hcy biosynthesis for the experiments reported, although persulfide-dependent Hcy biosynthesis has never been described.

These considerations strongly suggest that novel processes for the synthesis of both persulfide and Hcy await discovery in methanogens. To identify new proteins associated with sulfide assimilation we employed a bioinformatics approach to discover candidate genes present in anaerobes containing the SepRS–SepCysS pathway for $tRNA^{Cys}$ aminoacylation. Genetic analyses performed in *Methanosarcina acetivorans* then revealed that two of the corresponding proteins are required for Hcy biosynthesis. Phylogenetic reconstruction of the identified protein sequences are congruent with respect to presently accepted phylogenies, suggesting that these proteins were vertically inherited, together with the methanogenesis pathway, from the ancestral euryarchaeote.

Results

Bioinformatics identifies three gene families co-occurring with SepCysS in Archaea

To identify novel genes involved in anaerobic sulfide assimilation, we employed occurrence profiling using SepCysS as the focal point, taking a comprehensive approach not dependent on predetermined protein classifications such as the clusters of orthologous groups (COGs) (Tatusov *et al*., 1997). We focused on SepCysS because it is the only known sulfur assimilatory enzyme that is specific to anaerobic sulfur metabolism, and because its limited occurrence in methanogens and the related *Archaeoglobi* made the *in silico* approach more practicable. To proceed, query sequences consisting of proteins with known roles in aspects of sulfur metabolism, and those encoded in the genomic neighborhoods of their homologs in methanogens, were used to search all non-redundant protein sequences in all archaeal genomes using BLASTp (Fig. 2; Table S2). Phylogenetic reconstructions were then performed on aligned protein sequences to infer orthology and observe conservation patterns. This approach identified three gene families, represented within the *M. acetivorans* C2A genome (NC_003552.1) by *ma1821*, *ma1822* and *ma1715*. Homologs of each gene were found in all SepCysS-encoding genomes (Table S3) and were absent in nearly all archaeal genomes lacking SepCysS. The presence of several unusual methanogen genomes in the out-group lacking SepCysS, representing organisms found in the human intestinal tract, was valuable because it allowed exclusion of unrelated taxon-specific genes such as those encoding proteins participating in methanogenesis. Comprehensive transposon mutagenesis performed in *M. maripaludis* showed that

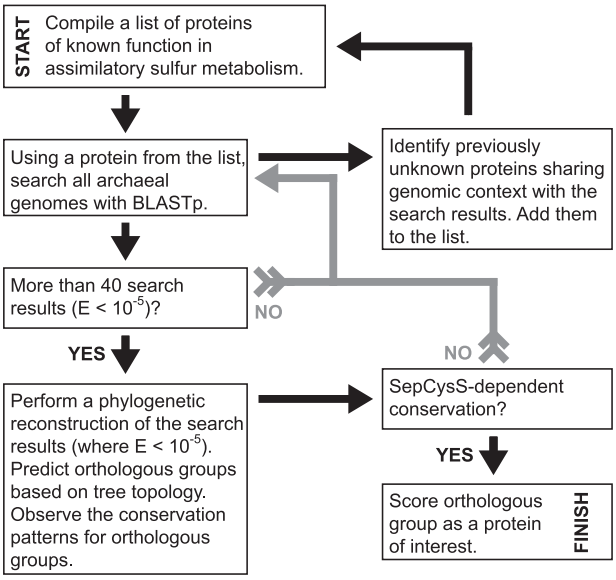

Fig. 2. Depiction of the occurrence profiling strategy used to identify protein-coding genes present in archaeal genomes encoding SepCysS. See text and Supplementary experimental procedures for details. Over 100 archaeal genomes were used in this strategy at the time this study was completed. Of these, 41 genomes encoded SepCysS.

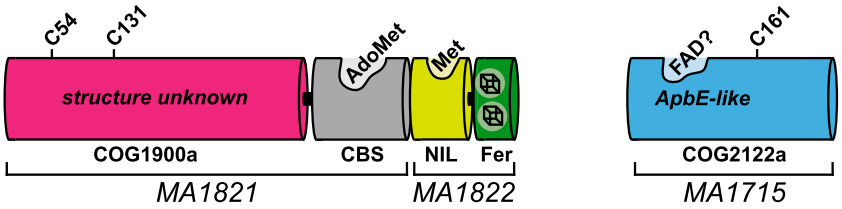

Fig. 3. Cartoon representations for the proteins sequences encoded by *MA1821* (pink/gray), MA1822 (green/yellow) and MA1715 (blue) are drawn to scale. COG1900a (DUF39; pink) is of unknown structure and function, but contains two conserved cysteine residues. The C-terminal CBS domains (gray) of MA1821 are proposed to bind *S*-adenosylmethionine. The NIL domain (yellow) of MA1822 may bind methionine based on homology to the methionine transporter. The Fer_4_7 domain (Fer; green) of MA1822 has eight conserved cysteine residues thought to bind two [4Fe-4S] clusters. MA1715 is classified into COG2122a, and has a single conserved cysteine residue. Its bacterial paralog, ApbE, binds FAD and is involved in thiamine biosynthesis. It is not known if MA1715 binds FAD.

the homologs of all three of these genes are essential in this hydrogenotrophic methanogen, which lacks CD enzymes and does not possess the OASS or OAHS genes for formation of Cys or Hcy using sulfide as the direct sulfur source (Fig. 1A) (Sarmiento *et al.*, 2013).

Description and classification of gene families

The 500-amino-acid *MA1821* protein from *M. acetivorans* contains two distinguishable regions. The Conserved Domain Database (CDD) identifies a large N-terminal domain (amino acids 4–362) as COG1900 (DUF39 superfamily; cl14897), followed by a smaller C-terminal region (amino acids 380–494) assigned to COG0517 (CBS_pair superfamily; cl15354) (Fig. 3) (Marchler-Bauer and Bryant, 2004; Marchler-Bauer *et al.*, 2011). A BLASTp search of all genomes using the N-terminal domain identified 373 unique protein sequences with significant alignments (e-value $< 10^{-8}$), all within COG1900. Using phylogenetic reconstruction, we further classified the proteins into four subdivisions, COG1900A-D (Fig. S1; see below for further elaboration).

MA1821 is associated with COG1900A, which contains 182 unique protein sequences. Of these, 64 sequences belong to SepCysS-encoding *Euryarchaeota* and two sequences are from the *Methanoplasmatales*, a very recently described taxon that lacks SepCysS (Borrel *et al.*, 2013). The remaining 118 sequences belong to 57 bacterial genera from nine different phyla. All species found in the bacterial genera encoding COG1900A are anaerobic, suggesting that the function of the protein may be oxygen-sensitive. Further, COG1900A is often encoded in close genomic proximity with genes predicted to play a role in sulfur metabolism, especially those required for *de novo* biosynthesis of methionine and sulfur-containing cofactors (Fig. 4). The large N-terminal domain of COG1900A contains two highly conserved Cys residues (Fig. 3). Cys54 in MA1821 is conserved in all COG1900A proteins, while Cys131 is strictly conserved in all methanogens and is rarely substituted with aspartate or glutamate in bacterial orthologs. COG1900A also lacks a conserved lysine residue, suggesting that the proteins do not bind pyridoxal phosphate (PLP).

In all SepCysS-encoding genomes, two CBS domains (named for the protein cystathione-β-synthase) are located C-terminal to the DUF39 region. These CBS domains are absent from MA1821 orthologs in genomes lacking SepCysS, suggesting a non-essential function. CBS domains are thought to function in a regulatory capacity in response to binding of adenosine derivatives (Baykov *et al.*, 2011). This is consistent with the recent finding that the CBS domains found in the *M. jannaschii* MA1821 ortholog, MJ0100, undergo a conformational change in response to binding of *S*-adenosylmethionine (AdoMet) (Lucas *et al.*, 2010). CBS domains are typically found in proteins involved in energy and sulfur metabolism (Baykov *et al.*, 2011). Their presence exclusively in MA1821 orthologs from genomes containing SepCysS suggests that the functions of MA1821 and SepCysS may be linked. Interestingly, a CBS domain with 70% sequence similarity to that found in MA1821 is predicted in the gene encoding *M. acetivorans* homoserine acetyltransferase (HSAT) (Fig. 1C). HSAT catalyzes the synthesis of *O*-acetylhomoserine from homoserine and acetyl-CoA; *O*-acetylhomoserine is a precursor of Hcy in many *Bacteria* and in some *Eukarya* and *Archaea* (Miyajima and Shiio, 1973). The presence of similar CBS domains in MA1821 and HSAT may indicate a shared regulatory strategy.

The 128-amino-acid MA1822 protein contains two distinct domains recognizable within the CDD. The N-terminal domain (amino acids 5–67) belongs to the *NIL* superfamily (cl09633; pfam09383; smart00930), a proposed substrate-binding domain found in the C-terminal, intracellular region of the MetN subunit of the methionine ABC transporter (Kadaba *et al.*, 2008). The C-terminal portion of MA1822 (amino acids 81–124) is of the Fer4_7 superfamily (pfam12838). It contains eight conserved Cys residues and is predicted to bind two iron sulfur clusters [4Fe-4S]. All SepCysS-encoding genomes encode this *NIL-Fer4_7* fusion protein (Table S1). The protein is almost always

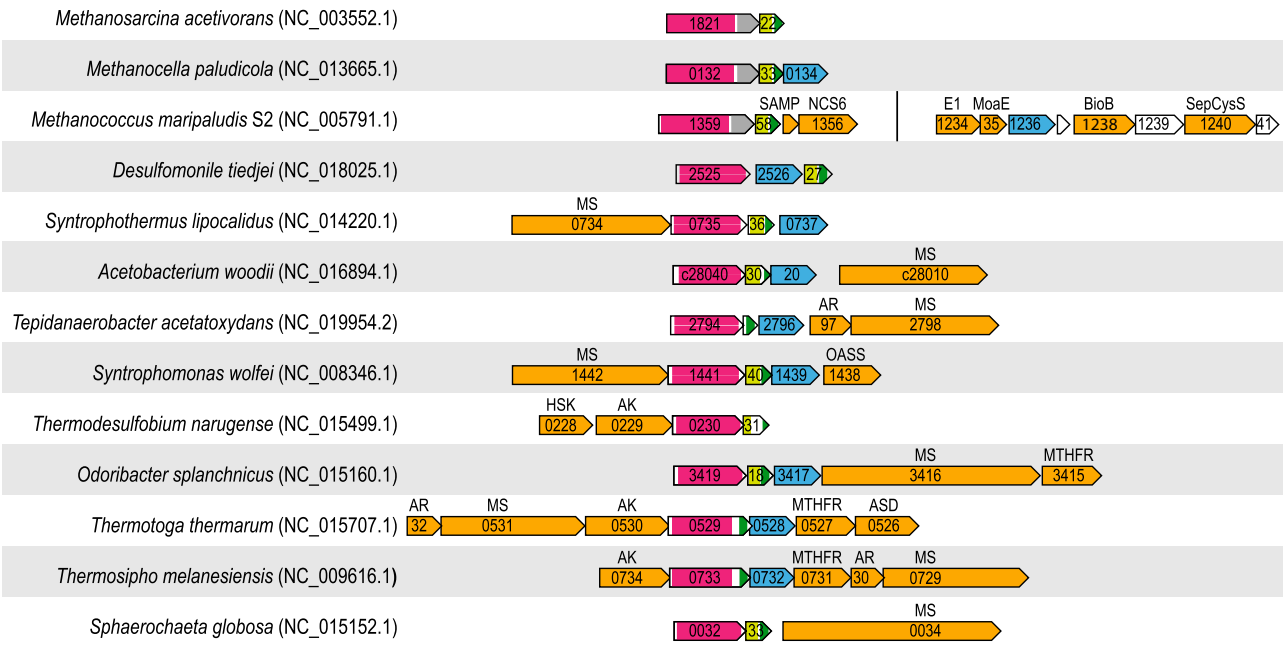

Fig. 4. Orthologs of MA1821, MA1822 and MA1715 from *Bacteria* and *Archaea* occur in genomic neighborhoods with genes relevant to sulfur metabolism (orange). These include the small archaeal modifying protein and sulfur carrier (SAMP), the s^2U34 biosynthesis protein (Ncs6), the SAMP adenylating protein *E1*, (MoaE), (BioB), Sep-tRNA:Cys-tRNA synthase (SepCysS), methionine synthase (MS), the methionine synthase activating region (AR), *O*-acetylserine sulfhydrylase (OASS), homoserine kinase (HK), aspartate kinase (AK), methylenetetrahydrofolate reductase (MTHFR) and aspartatesemialdehyde dehydrogenase (ASD). Predicted open reading frames are labeled by NCBI locus number. Orthologs of MA1821, defined by the presence of COG1900a (pink), are fused to CBS domains (gray) in SepCysS-encoding genomes. Orthologs of MA1822 are defined by the presence of the ferredoxin domain (green) occurring downstream from orthologs of MA1821. Usually, the ferredoxin domain occurs in a separate open reading frame fused to NIL (yellow-green). However, in examples where NIL is absent, the ferredoxin either is fused to the C-terminus of COG1900a or occurs as a free-standing open reading frame downstream from MA1821 orthologs. Orthologs of MA1715 are defined by the presence of COG2122a (blue), and are often encoded downstream from MA1822 orthologs. Cartoons were constructed to scale with the assistance of the Microbial Genomic context Viewer (MGcV; http://mgcv.cmbi.ru.nl/) (Overmars *et al.*, 2013).

encoded immediately downstream of MA1821 orthologs, although a variation exists in *Thermotoga*, *Thermosipho*, *Fervidobacterium* and *Coprothermobacter* spp., in which the Fer4_7 iron-sulfur clusters from the MA1822 homolog are fused to the C-terminus of the MA1821 homolog protein (Figs 3 and 4, Fig. S1). There are also several genomes that possess an MA1822 ortholog lacking the NIL coding region (Figs 3 and 4). These observations suggest that *NIL* may be dispensable, that *Fer4_7* is the essential functional element of MA1822-like proteins, and that the functions of MA1821 and MA1822 orthologs are linked.

Finally, the 253-amino-acid *MA1715* protein belongs to COG2122 (cl17892; PRK04334). COG2122 is uncharacterized but is classified in the ApbE superfamily (cl18387), which features FAD-binding enzymes implicated in thiamine biosynthesis and iron-sulfur cluster maintenance (Beck and Downs, 1998; Skovran and Downs, 2003). The protein contains a single conserved Cys residue, and no conserved lysine residues. A BLASTp search of all genomes using *MA1715* identifies 171 unique protein sequences (on 01/16/14; 76 in *Archaea*) yielding significant alignments (< 1e-3), all classified within COG2122. Three subdivisions were found and were designated COG2122A-C (Fig. S2), with MA1715 falling in COG 2122A. The occurrence of COG2122A in *Archaea* and *Bacteria* is nearly identical to that of COG1900A; the two families co-occur in over 98% of genomes (Table S2). Furthermore, COG2122A is often encoded in close genomic proximity to COG1900A, and is also often found immediately downstream from proteins resembling MA1822 (Fig. 4).

Evolutionary history of MA1821, MA1822 and MA1715 orthologs

Phylogenetic reconstructions of COG1900A and COG2122A proteins families show that each is congruent with respect to currently accepted organismal phylogenies (data not shown). To provide a further assessment, we repeated the analysis with a single sequence corresponding to the concatenated protein sequences for orthologs of MA1821, MA1822 and MA1715. This reconstruction again yielded a tree that is nearly identical to those observed for methanogen genome phylogeny and

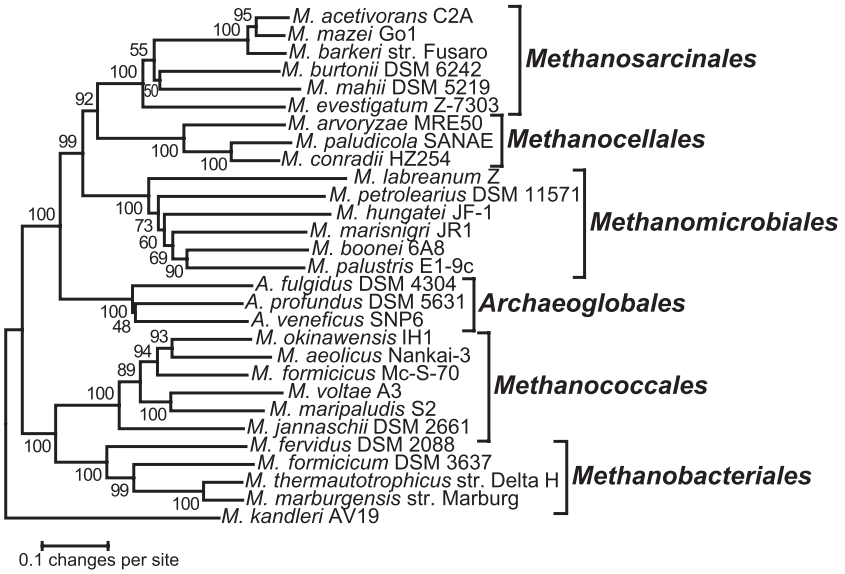

Fig. 5. Phylogenetic reconstruction of concatenated protein sequences for MA1821, MA1822 and MA1715 orthologs from selected SepCysS-encoding *Archaea*. Sequences were downloaded from GenBank and aligned with CLUSTALW, using default settings. A neighbor-joining tree was constructed in MEGA6 using the Poisson model, uniform rates and complete deletion of gaps. Numbers reflect bootstrapping support as a percentage of 1000 replications. Taxonomic orders are grouped under brackets.

methanogenesis enzymes (Fig. 5 and Blank, 2009). This suggests that the three gene families we have identified are of similar ancestry and originated early in the evolution of the *Euryarchaeota* (Bapteste *et al*., 2005). This is consistent with the notion that SepCysS is also an ancient gene (O'Donoghue *et al*., 2005), and that the mechanism for sulfide assimilation in methanogens is linked to the origins of methanogenesis.

The congruence of this phylogenetic tree with respect to methanogen genome phylogenies suggests a history of vertical inheritance of ma1821, ma1822 and ma1715 in SepCysS-encoding organisms. To extend this analysis to the Bacteria, we further examined the evolutionary histories of MA1821, MA1822 and MA1715 at the domain-domain level using ratios of evolutionary distances (RED) analysis (Farahi *et al*., 2004). For both MA1821 and MA1822, intra-phylum and -domain distances separating orthologs scaled proportionately to the corresponding control distances, suggesting widespread vertical inheritance within *Bacteria* and *Archaea* (Fig. S10). Interestingly, for both MA1821 and MA1715 orthologs, interdomain distances did not extrapolate linearly from the intradomain data. This non-linearity likely reflects an ancient interdomain lateral gene transfer event that predated the divergence of bacterial phyla. Alternatively, non-linearity might be the result of a change in the evolutionary rates for orthologs of both MA1821 and MA1715. Our analysis also detected the existence of lateral gene transfer between some bacterial taxa.

Identification of COG1900 paralogous groups

Phylogenetic reconstruction reveals that the 373 proteins recovered in BLASTp searches with the MA1821 protein are classifiable into four subdivisions, which we have termed COG1900A–D (Figs S1 and S3). The orthologous group corresponding to the MA1821 protein is assigned as COG1900A, and is the largest of the four. None of the proteins in the other three groups possess the CBS domain. COGs 1900B and 1900C are each found exclusively in cyanobacteria, and consist of 108 and six sequences respectively. Of the two Cys residues conserved in all COG1900A proteins (Cys54 and Cys131 in MA1821), the proteins in COG1900B contain Cys131 only, while those in COG1900C and COG1900D possess neither. Cys54 is substituted as Glu, Asp or Ser in COGs 1900B, 1900C and 1900D respectively (Fig. S3).

COG1900D proteins possess the Fer4_16 (pfam 13484) domain containing two 4Fe-4S clusters fused to their C-termini, and are found almost exclusively in methanogens. These proteins possess Ser54 instead of Cys54, and lack a consensus TDCYP motif that surrounds Cys131 of MA1821 (Fig. S3). The lack of conservation of these Cys residues in COG1900D, together with the conserved fusion of the 4Fe-4S clusters at the C-termini, suggests a function related to the methanogenesis lifestyle that is distinct from COG1900A. The comprehensive transposon mutagenesis suggests that COG1900D proteins are not essential in *M. maripaludis* (Sarmiento *et al*., 2013).

Functional roles of MA1821 and MA1822

We took a genetic approach to examine the functions of *ma1821* and *ma1822*. *M. acetivorans* was chosen as the experimental organism based on the availability of genetic techniques for manipulation of the chromosome (Buan *et al*., 2011). *M. acetivorans* is a mesophile that is efficiently grown on methanol. As observed for other species in the

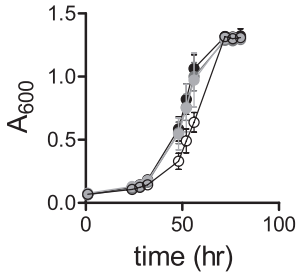

Fig. 6. Growth curves demonstrating that genes encoding OASS and OAHS are not essential to sulfide assimilation in *Methanosarcina acetivorans*. Strains Δ*oass* (dark grey), Δ*oahs* (light gray), Δ*oass*Δ*oahs* (black open circles), and pseudo-wild-type strain WWM75 (black closed circles) were grown on a medium containing sulfide as the sole sulfur source (HS_{DTT} medium supplemented with 0.80 mM sodium sulfide; see *Experimental procedures*). Data points represent the mean values for six independent experiments. Error bars depict standard deviations.

Methanosarcina (such as *M. barkeri* and *M. mazei)*, *M. acetivorans* has acquired a large number of bacterial genes by horizontal transfer (Galagan *et al*., 2002). Although detailed evolutionary histories have not been examined, among these genes may be those encoding enzymes participating in sulfur metabolism. In particular, *M. acetivorans* possesses the genes encoding OAHS and O-acetylserine sulfhydrylase (OASS) for synthesis of Hcy and Cys, respectively, from sulfide, and CD for the relay and synthesis of persulfide from cysteine (Fig. 1A). It also encodes CysRS and several CD homologs. The redundant sulfur assimilation and trafficking pathways is advantageous in gene knockout studies, since otherwise essential genes are predicted not to result in lethality when deleted. This characteristic may be particularly important for identifying ancient genes involved in the primary uptake of sulfide (see *Discussion*).

To demonstrate the existence of a new sulfide assimilation pathway in *M. acetivorans*, the genes encoding OASS and OAHS were deleted from the chromosome both individually and in combination, using a markerless approach (Buan *et al*., 2011; Kohler and Metcalf, 2012). Each deletion was confirmed by both PCR analysis and Southern blotting (Figs S4 and S5). The three resulting strains – Δ*oass*, Δ*oahs*, and Δ*oass*Δ*oahs* – were each capable of growth on a medium containing sulfide as the sole sulfur source (Fig. 6). All three mutant strains grew comparably to the wild type. The viability of the Δ*oass*Δ*oahs* strain on sulfide-only medium demonstrates that an alternative pathway for sulfide assimilation must exist in *M. acetivorans*.

We next investigated the function of the *ma1821* and *ma1822* genes. A double-deletion strain in which the adjacent *ma1821* and *ma1822* genes were excised together was first created using the markerless deletion approach (Fig. 7). A triple chromosomal deletion consisting of the further removal of the gene encoding OAHS was then constructed in a subsequent experiment (Fig. S4). The resulting strains – Δ*ma1821-22* and Δ*ma1821-22*Δ*oahs* – were grown on a medium containing sulfide as the sole sulfur source. Under these conditions, Δ*ma1821-22* grew comparably to wild type, whereas the triple mutant Δ*ma1821-22*Δ*oahs* exhibited no growth (Fig. 8A). The robust growth of the Δ*ma1821-22* strain demonstrates that OAHS and OASS function efficiently in sulfide uptake.

Growth of the Δ*ma1821-22*Δ*oahs* triple mutant was recovered by addition of either Met or Hcy to the growth

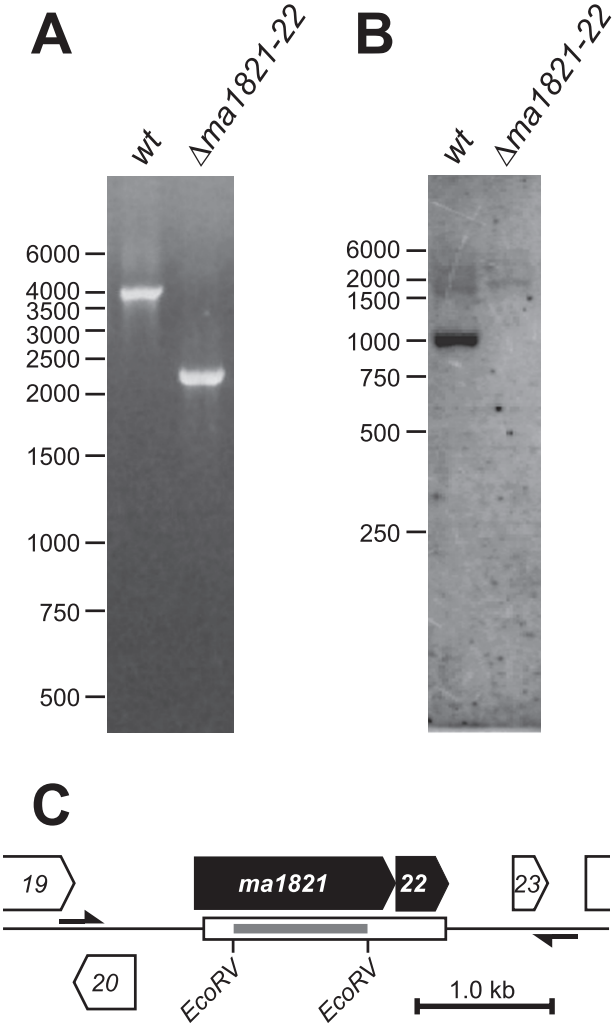

Fig. 7. Markerless chromosomal deletion of *ma1821-22* constructed in the pseudo-wild-type background of WWM75 (wt). PCR (A) and Southern blot (B) analyses are consistent with the desired manipulation (C), which removes a 1790 nt segment (white rectangle) of *ma1821-22* from the chromosome. Primers for PCR analysis (black arrows) were used to amplify the genomic neighborhood of *ma1821-22*, which is predicted to measure 3841 nt in the wild-type strain and 2051 nt in the Δ*ma1821-22* strain. The probe used for Southern blot analysis (grey line) measures 990 nt and hybridizes to the 997 nt fragment that lies between two endogenous EcoRV restriction sites within *ma1821*.

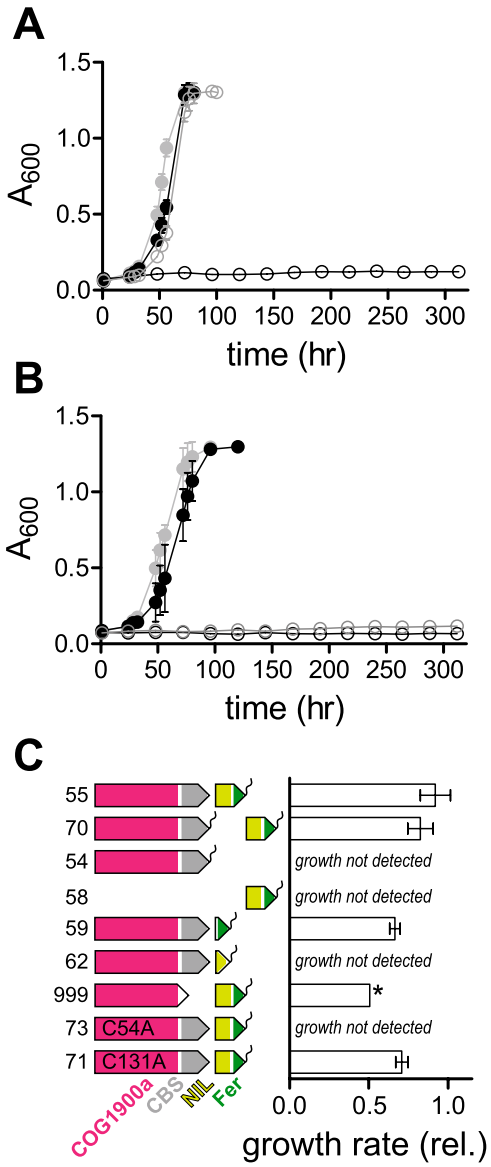

Fig. 8. Growth curves demonstrating that *ma1821* and *ma1822* are essential to Hcy biosynthesis in an engineered *Methanosarcina acetivorans* strain lacking OAHS.
A. Strains Δ*ma1821-22* (gray filled circles), Δ*ma1821-22* with the *ma1821-22* genes added back in plasmid pBR031 (open gray circles; construct 55, panel C and Fig. S7), Δ*ma1821-22*Δ*oahs* (open black circles) and Δ*ma1821-22*Δ*oass* (black filled circles) were grown on a medium containing sulfide as the sole sulfur source (HS_{DTT} medium supplemented with 0.80 mM sodium sulfide; see *Experimental procedures*). Only marginal growth of Δ*ma1821-22* with the *ma1821-22* genes added back in plasmid pBR031 is observed when the tetracycline inducer is omitted (data not shown).
B. Growth of the Δ*ma1821-22*Δ*oahs* strain on a medium containing 0.8 mM sodium sulfide supplemented with Cys (3 mM; open circles), Met (3 mM; filled gray circles) cystathionine (3 mM; open gray circles) or Hcy (3 mM; filled black circles).
C. Add-back experiments for identification of the essential structural elements of *ma1821-22*. The identities of the color-coded domains are indicated at bottom left. Variants were added back to the Δ*ma1821-22*Δ*oahs* strain using the shuttle vector pBR031 (Fig. S7) and subsequently grown on a medium containing 0.8 mM sodium sulfide as the sole sulfur source. Growths were conducted in the presence of tetracycline (50 μg ml^{-1}) and puromycin (2 μg ml^{-1}). The curved lines at the C-termini of the constructs represent poly-histidine tags. MA1821 and MA1822 are separately expressed from independent ribosomal binding sites. The intergenic spacing of two nucleotides in the *M. acetivorans* chromosome is preserved in constructs 55, 59, 62, 73 and 71. Construct 999 (CBS domain deletion) possesses a larger interdomain spacer, as does construct 70 containing a second His tag added to the C-terminus of MA1821. The corresponding plasmids are numbered at left as in Table S4. Growth curves for these strains are shown in Fig. S8. Statistical analysis was not performed for the strain indicated by the asterisk.

medium (Fig. 8B). This finding indicates that the strain is an Hcy auxotroph, thereby implicating the MA1821-22-coding region in a novel, OAHS-independent biosynthetic route to Hcy. To confirm the properties of this strain, we constructed a multi-copy, tetracycline-dependent expression plasmid (Fig. S7) and inserted the MA1821-22-coding region for expression of the proteins *in vivo*. The resulting transformed triple chromosomal mutant strain contained plasmid-borne *ma1821* and *ma1822* genes, and grew comparably to wild type on sulfide-only medium supplemented with tetracycline (Fig. 8C and Fig. S8), signifying the restoration of OAHS-independent Hcy biosynthesis.

Unlike the ability of Met or Hcy to restore growth to the Δ*ma1821-22*Δ*oahs* strain, growth could not be recovered by adding Cys to a growth medium containing sulfide (Fig. 8B). Therefore, it appears that the metabolic defect in the Δ*ma1821-22*Δ*oahs* strain is the inability of the cell to synthesize Hcy from sulfide. The inability of Cys to complement the deletions indicates that the transsulfuration pathway by which Cys is converted to Hcy is not present in *M. acetivorans*. This is consistent with our inability to detect homologs of cystathione γ-synthase (CGS) or cystathionine β-lyase (CBL) in the genome (Fig. 1A), and with metabolite labeling experiments in *M. maripaludis* demonstrating that Cys is not the sulfur source for Met biosynthesis (Liu *et al.*, 2010). To further examine the possibility of transulfuration despite the apparent absence of CGS and CBL, we attempted to recover growth of the Δ*ma1821-22*Δ*oahs* strain by adding 3 mM cystathionine to a growth medium containing sulfide (Fig. 8B). Like Cys, cystathionine also fails to complement the deletions, providing further evidence that MA1821 and MA1822 are involved in direct sulfhydrylation of homoserine derivatives (Fig. 1C; see *Discussion*).

To examine whether MA1821 and MA1822 may also play a role in cysteine biosynthesis, we constructed a triple deletion strain in which the *ma1821* and *ma1822* genes were excised together with the gene encoding OASS (Fig. S6) (Δ*ma1821-22*Δ*oass*). This mutant strain grows comparably to wild type on sulfide-only medium (Fig. 8A),

and does not require the addition of exogenous cysteine. This finding suggests that MA1821 and MA1822 do not participate in the SepCysS-dependent pathway for Cys biosynthesis (although they may possibly function together with OASS in the redundant SepCysS-independent Cys biosynthesis pathway in *M. acetivorans*). Both Δ*ma1821-22*Δ*oass* and pseudo wild-type *M. acetivorans* (strain WWM75) are also capable of growth on either Cys or Met alone as a sulfur source, and do not require exogenous sulfide (Fig. S9; data not shown). The ability of Met to support growth as the sole sulfur source, albeit weakly, suggests that a route exists for Hcy-dependent Cys biosynthesis, despite the apparent absence of the reverse pathway by which Hcy can be biosynthesized from exogenously added Cys (since Δ*ma1821-22*Δ*oahs* cannot grow with either Cys and sulfide or cystathionine and sulfide; Fig. 8B). However, no genes encoding homologs of transsulfuration enzymes that interconvert Cys and Hcy are detected in *M. acetivorans*. The ability of Cys to support growth as the sole sulfur source further suggests that a pathway may exist for the generation of sulfide from Cys (see *Discussion*).

Structure–function analysis of the MA1821 and MA1822 proteins

To identify protein structural elements essential to OAHS-independent Hcy biosynthesis, a number of MA1821-22-coding region variants were constructed in pBR031 and transformed into the Δ*ma1821-22*Δ*oahs* strain (Fig. 8C, Figs S7 and S8). The resulting strains were then assayed for the capacity to biosynthesize Hcy by monitoring growth in the presence of tetracycline, in a medium containing sodium sulfide as the sole sulfur source (Fig. 8 and Fig. S8). As an initial experiment, we first separately expressed the *ma1821* and *ma1822* genes in the Hcy auxotroph triple chromosomal deletion strain. Neither gene alone was capable of restoring growth [doubling times (T_d) > 300 h], indicating that structural elements from both proteins are required for OAHS-independent Hcy biosynthesis (Fig. 8C and Fig. S8).

To explore the roles of particular subdomains, we constructed and studied the following proteins (Fig. 3): (i) full-length MA1821 together with the separately expressed NIL domain of MA1822, (ii) full-length MA1821 together with the separately expressed Fer domain of MA1822, and (iii) the N-terminal COG1900a portion of MA1821 together with separately expressed full-length MA1822. The truncated MA1822 protein consisting of the NIL domain alone was unable to restore a capacity for biosynthesis of Hcy to the triple deletion mutant strain (Fig. 8C and Fig. S8), demonstrating that the predicted iron-sulfur clusters in Fer are essential to catalytic function. However, the truncated MA1822 protein consisting of the Fer domain alone remained capable of restoring robust growth in the presence of sodium sulfide, suggesting that NIL does not confer an essential function in Hcy biosynthesis. Finally, the truncated MA1821 protein lacking both CBS domains also generated a strain capable of growth with sodium sulfide as the sole sulfur source, when coexpressed with full-length MA1822. However, in this case the doubling time of the strain was increased from 9 h to 16 h (Fig. 8C and Fig. S8), indicating that these domains contribute to the efficiency of OAHS-independent Hcy biosynthesis.

The expendability of the CBS domains implies that the N-terminal COG1900A region of MA1821 has an essential role in Hcy biosynthesis. To explore the roles of the conserved C54 and C131 residues in this domain, alanine mutations were separately introduced at each position. Surprisingly, the C131A mutant strain is capable of growth on sulfide-only medium, indicating that this Cys residue is not functionally significant in this context. However, the strain bearing C54A was incapable of growth under these conditions (T_d > 300 h), suggesting that the thiol of C54 has an essential role in OAHS-independent Hcy biosynthesis (Fig. 8C and Fig. S8).

Discussion

We have demonstrated that two highly conserved and previously unexamined genes are essential to Hcy biosynthesis in *M. acetivorans* when present in a genetic background lacking the *oahs* gene for Hcy formation (Fig. 1A). Homologs of the identified *M. acetivorans ma1821* and *ma1822* genes are found in all *Archaea* possessing the SepCysS enzyme, which inserts sulfur into Sep-tRNA^{Cys} and is essential to a novel pathway for Cys biosynthesis in methanogens and *Archaeoglobi* (Sauerwald *et al.*, 2005). Other homologs were then identified in species belonging to over 50 distinct bacterial genera, all of which are anaerobic (Fig. S1 and Table S3). Phylogenetic reconstruction of concatenated protein sequences for MA1821, MA1822 and MA1715 orthologs, which are often linked on the chromosome, shows congruence with accepted organismal phylogenies within the *Archaea* (Fig. 5), and RED analysis suggests the corresponding genes may have been subject to an ancient lateral gene transfer event (Fig. S10). Therefore, these two genes function in a novel and broadly conserved process for incorporation of sulfide into Hcy in anaerobes, that may have originated in microbial communities in the Archaean Eon. A role for MA1821 and MA1822 in Hcy biosynthesis is consistent with genomic colocalization of the corresponding genes with those for Met biosynthesis in many organisms (Fig. 4). The absence of these genes in aerobes is consistent with function in sulfide assimilation, since sulfide is spontaneously oxidized in aerobic environments. The function of the third highly conserved gene that we have identified by bioinfor-

matic techniques, which presumptively encodes the *M. acetivorans* MA1715 protein, is under investigation in our laboratory. Although the precise biochemical activities of the proteins are not yet known, this discovery of the physiological function of MA1821 and MA1822 offers key insight into the question of how Hcy is synthesized in anaerobes that lack both OAHS and transsulfuration enzymes.

The co-occurrence of COG1900A and NIL-Fer with SepCysS in methanogens, their apparent vertical inheritance (O'Donoghue *et al*., 2005), and the role of each protein in biosynthesis of an essential sulfur-containing amino acid suggests common participation in an ancient, anaerobic process for sulfide assimilation and sulfur trafficking that was present in the ancestral euryarchaeote. In parallel with the role of COG1900A and NIL-Fer in methanogen Hcy biosynthesis, SepCysS is instead essential to the other key branch of methanogen sulfur assimilation: formation of Cys in those organisms that lack both OASS and transulfuration enzymes for synthesis of Cys from Hcy (Fig. 1A and C and Sauerwald *et al*., 2005). Free Cys is then presumably generated by protein turnover and by deacylation of Cys-tRNA^{Cys} before incorporation. Intracellular concentrations of Cys in methanogens are 5- to 10-fold lower than found in *E. coli*, consistent with a diminished role for Cys in cellular metabolism and with the apparent inefficiency of this biosynthetic route (Liu *et al*., 2010). Hydrogenotrophic methanogens such as *M. maripaludis* and *M. jannaschii* also lack CD enzymes, so that Cys is not the sulfur source for iron-cluster biosynthesis or perhaps for biosynthesis of other sulfur-containing cofactors and tRNAs via persulfide relay pathways. The origin of persulfide synthesis in these methanogens, and the mechanism for delivery of sulfane sulfur to SepCysS, remain as key unknowns.

Our findings suggest several possibilities for the biochemical functions of MA1821 and MA1822. Since each protein possesses a conserved and essential molecular signature consistent with participation in sulfur trafficking, we suggest that their most likely role is in the direct catalysis of Hcy formation. A mechanism requiring just one essential Cys residue in MA1821 would be consistent with the experimental data showing that Cys54, but not Cys131, is required to rescue the Δ*ma1821-22*Δ*oahs* triple deletion strain when supplied *in trans* (Fig. 8). We speculate that MA1821 may catalyze the direct attachment of sulfide to its conserved and essential C54 sulfhydryl, forming a persulfide group with concomitant transfer of two electrons to the 4Fe-4S clusters on MA1822. Alternatively, MA1821 may accept sulfane sulfur from an as-yet unidentified protein donor. In either case, the sulfane sulfur could then function as a nucleophile to attack Cγ of an activated homoserine compound, generating an enzyme-substrate disulfide bond that could be resolved following electron donation from the 4Fe-4S clusters (Fig. 1A and C). If the sulfane sulfur on MA1821 is accepted from a protein donor, this mechanism would require another electron donor to regenerate the 4Fe-4S cluster. These considerations are consistent with a single conserved and essential Cys residue in all COG1900A proteins, and with the essential function of the 4Fe-4S clusters (Fig. 8C and Fig. S8). The function of the peripheral sulfur of a persulfide group as a nucleophile is well-established (Mueller, 2006), and this proposed mechanism is also consistent with the showing that free sulfide is not the immediate sulfur source for Hcy biosynthesis in either *M. jannaschii* or *M. maripaludis* (White, 2003; Liu *et al*., 2010).

Another possibility for MA1821–MA1822 function that is formally consistent with our genetic findings could be participation in the biosynthesis of an activated homoserine compound that would in turn serve as the precursor to Hcy via a direct sulfhydrylation pathway. In *M. jannaschii*, Hcy is biosynthesized from phosphohomoserine via the action of homoserine kinase, and the phosphohomoserine in turn functions as the only known precursor for threonine via the conserved threonine synthase enzyme (Fig. 1C and White, 2003). Although we cannot detect the gene for homoserine kinase in the *M. acetivorans* genome, the apparent requirement for phosphohomoserine for threonine biosynthesis suggests that it may be present in a modified form (Fig. 1C). Alternatively, acetylhomoserine may provide the Hcy precursor, since the homoserine acetyltransferase enzyme for biosynthesis of acetylhomoserine from homoserine is present in *M. acetivorans* (Fig. 1C and Table S1). There is precedent for Hcy biosynthesis from either precursor, each of which provides a suitable leaving group. Enzymes necessary for the biosynthesis of homoserine from aspartate are conserved in all SepCysS-containing *Archaea* (Fig. 1C and Table S1), consistent with the notion that an activated homoserine derivative is required for Hcy biosynthesis from a scaffold metabolite that lacks sulfur (White, 2003).

While the Hcy auxotrophy of Δ*ma1821-22*Δ*oahs* is consistent with biochemical function of MA1821-22 at any step of the Hcy biosynthesis pathway, the possibility that the proteins are involved in biosynthesis of an activated homoserine compound appears less compelling because catalysis of that reaction would not necessarily require conserved a Cys residue or 4Fe-4S cluster. Further, the Δ*ma1821-22* strains are not auxotrophic for threonine, as would be expected if MA1821-22 function in early stages of a pathway up to and including formation of an activated homoserine precursor to Hcy (Fig. 1C). The genetic data are also consistent with the notion that MA1821-22 are essential for methionine recycling, although the presence in *M. acetivorans* of all four known genes involved in this process, and the conservation of these genes in aerobes as well as anaerobes, does not favor this hypothesis. Finally, it is also possible that the genes may play a

regulatory role – although in this case the conservation of an essential Cys and 4Fe-4S clusters is again more difficult to rationalize.

The data also suggest that MA1821 and MA1822 are limited to Hcy biosynthesis and are not participants in a more general pathway of sulfide assimilation or persulfide synthesis. Deletion of an upstream general persulfide synthesis enzyme together with OASS would be expected to result in a phenotypic response resembling Cys auxotrophy, since exogenously added Cys would then be required to reactivate the CD enzymes in *M. acetivorans* (Fig. 1A). However, Cys auxotrophy was not observed in the Δ*ma1821-22*Δ*oass* strain, which grows comparably to the pseudo-wild-type strain WWM75 on a medium containing sulfide as the sole sulfur source (Fig. 8A). Therefore, it appears likely that MA1821 and MA1822 are not required for providing sulfane sulfur to biosynthetic pathways that might otherwise acquire it from CD, including pathways delivering sulfur to Cys-tRNA^{Cys}, cofactors and modified ribonucleotides. Elucidation of the ancestral pathway for persulfide synthesis is of great importance, as it would provide the essential basis for exploring downstream sulfur anabolic pathways in anaerobes, and would also offer a salient contrast with pathways in aerobic bacteria and eukaryotes.

The pseudo wild-type strain WWM75 is capable of very slow growth on a medium containing Met as the sole sulfur source (Fig. S9), indicating that some means for recycling sulfur from Met (and Hcy) to Cys must exist despite the fact that a transsulfuration pathway is not detected by homology with known enzymes from other organisms. It is possible that a weakly active pathway is nonetheless present; if so, the Δ*ma1821-22*Δ*oahs* strain might represent a pseudoauxotroph for Hcy if such a pathway exists and operates reversibly. Further genetic and biochemical experiments would be required to evaluate this possibility. The very weak growth on Met alone suggests that operation of this pathway could not account for the rapid growth of Δ*ma1821-22*Δ*oass* on sulfide-only medium (where Hcy is synthesized by OAHS).

Finally, it is provocative that the pseudo-WT WWM75 strain is able to grow relatively efficiently on Cys alone (Fig. S9). It seems unlikely that this level of growth could be provided by a transsulfuration pathway for Hcy biosynthesis, since growth of WWM75 on Met alone is very weak (Fig. S9), and neither Cys nor cystathionine can rescue the Δ*ma1821-22*Δ*oahs* strain (Fig. 8B). Growth on Cys alone should, however, allow for the efficient generation of protein persulfides on the multiple CD enzymes that are present. Therefore, we speculate that persulfide transfer to MA1821 for Hcy biosynthesis could readily occur under these conditions, and that the MA1821/MA1822-dependent pathway, rather than OAHS, may be providing the main route to Hcy.

Experimental procedures

Strains and media

Escherichia coli strains were grown aerobically in LB medium under the appropriate antibiotic selection. DH5α was used as the host for pET-22b(+), DH5α/λ-pir was used for the *oriR6K*-dependent plasmids pMP44 and pWM321 (Kolter *et al*., 1978), and WM3118 for the *oriV*-dependent pJK031A (Guss *et al*., 2008). Rhamnose (10 mM) was supplemented into the WM3118 cultures prior to plasmid purification, to permit plasmid replication at high copy number.

Methanosarcina acetivorans strain WWM75 and derived strains were grown in high-salt (HS) medium to promote single-cell morphology (Table S4; see Supplementary experimental procedures) (Sowers *et al*., 1993; Metcalf *et al*., 1996; Guss *et al*., 2008). All growths were carried out at 37°C, under 2 psi N_2:CO_2 (80:20), with methanol (125 mM) as the methanogenesis substrate. HS medium supplemented with 3 mM methionine (HS_{Met}) was used for routine strain propagation and genetic manipulations. Growth-monitoring experiments were conducted in HS medium lacking cysteine, sodium sulfide, thiamine, biotin and α-lipoic acid, with 1.5 mM dithiothreitol added as a reducing agent (HS_{DTT}). Sulfur sources were added to HS_{DTT} after autoclaving, immediately prior to use. Dithiothreitol alone does not support growth of the *M. acetivorans* strains used in this study (data not shown).

Construction of markerless deletions in M. acetivorans

Upstream and downstream flanking sequences, extending 1000–1250 bp from the gene to be deleted, were amplified separately by PCR from WWM75 genomic DNA. Opposing flanking sequences were then fused by overlap extension PCR, yielding mutant genotypes suitable for insertion into pMP44 using SpeI and KpnI sites (Tables S3–S5).

Liposome-mediated transfection was used to construct mero-diploid intermediate strains as described (Buan *et al*., 2011). Purified intermediate colonies were grown in 10 ml liquid medium without selection. After reaching quiescence, the strain was propagated for 15–20 additional generations. Counter-selective plating was performed by resuspending mature cultures (A_{600} ~ 1.0) in 0.1 ml HS_{Met} medium, and spreading 30 µl of the cell suspension at dilutions of 1:1, 1:10 and 1:100 on agar-solidified media containing 200 µg ml^{-1} 8-aza-2,6-diaminopurine sulfate (8ADP; Santa Cruz Biotechnology). Under these conditions, 8ADP-resistant colonies emerged after 20 days. Typically eight colonies were picked and grown in 10 ml liquid medium without selection. After assessing genotype by PCR, confirmed knockout strains were restreaked for purification on agar-solidified medium containing 200 µg ml^{-1} 8ADP. Purified strains were grown to quiescence without selection in 10 ml liquid medium and analyzed further by DNA hybridization via Southern blotting analysis, which was performed as described (Alphey and Parry, 1995; Parry and Alphey, 1995).

Complementation of the Δma1821-22Δoahs strain with pBR031

To construct pBR031 (Fig. S7), P_{mcrB}(tetO1) was first amplified from pJK031A and subsequently cloned into pWM321

using the SpeI and SphI sites (Tables S4, S5 and S6). Prior to insertion into pBR031, the coding regions for *ma1821, ma1821-1822* and *ma1822* were first cloned into pET22b(+) (Novagen), via NdeI and XhoI restriction sites, yielding pBR004, pBR038 and pBR05 respectively (Tables S4 and S5). Using site-directed mutagenesis, two endogenous SphI restriction sites occurring in *ma1821* were removed from pBR004 and pBR038, resulting in pBR051 and pBR053 (Tables S4 and S5). Coding regions from pBR051, pBR053 and pBR05 were amplified and cloned into pBR031 using SphI and ApaI sites, resulting in pBR054, pBR055 and pBR058 (Tables S4 and S5). Cysteine mutants were constructed in *ma1821* by site-directed mutagenesis on template plasmid pBR055 (Tables S4 and S5).

The Δ*ma1821-22*Δ*Fer* mutant was constructed from pBR055 in two steps. First, using site-directed mutagenesis, an XhoI restriction site was introduced immediately downstream from the NIL-coding region of *ma1822*, yielding pBR056. Second, the Fer-coding region was removed from pBR056 by digestion with XhoI, followed by intramolecular ligation, yielding pBR062.

The $ma1821\text{-}22_{\Delta NIL}$ mutant was constructed by amplification of *ma1821* and $ma1822_{Fer}$ from pBR053. *Ma1821* and $ma1822_{Fer}$ were then fused by overlap extension and cloned into pBR031 using SphI and ApaI, yielding pBR059 (Tables S4 and S5).

Plasmids derived from pBR031 were introduced into the Δ*ma1821-22*Δ*oahs* strain by liposome-mediated transfection, as described (Buan *et al*., 2011). Puromycin-resistant colonies were used in subsequent growth experiments without further purification.

Growth experiments

All growth experiments were conducted at 37°C in Balch tubes containing 10 ml HS_{DTT} medium supplemented with various sulfur sources. Starter cultures were grown to late exponential phase ($0.50 < A_{600} < 0.70$) and diluted into the experimental growth medium such that $A_{600(initial)} = 0.010$. To minimize carry-over of sulfur sources from HS_{Met} medium, starter cultures were grown in the experimental growth medium for at least seven generations when possible, such that carry-over concentrations for cysteine, methionine and sulfide were less than 1.2 μM, 1.2 μM and 0.3 μM respectively. Of necessity, in the case of strains auxotrophic for homocysteine, starter cultures were carried out with HS_{Met} medium.

Measurement of A_{600} was performed using a Turner 330 spectrophotometer. The instrument was modified to accommodate 18 mm Balch tubes, so that culture tubes could be inserted directly in the light path.

Doubling times were calculated by fitting data points ($A_{600} < 1.1$) to the exponential growth equation:

$$y = ae^{kx}$$

where y is time-dependent A_{600}, a is the initial A_{600}, k is the rate constant for growth and x is time. Prior to fitting, individual time-courses were normalized by subtracting a single value from all time points such that A_{600} of the 1 h time point = 0.01. Data were fit in GraphPad Prism without fixing a to 0.01. Reported growth rates and growth curves reflect the mean of five independent experiments. Error bars represent standard deviations.

Acknowledgements

We thank Bill Metcalf and Junkai Zhang for *Methanosarcina acetivorans* strains and relevant plasmids, and for invaluable assistance with procedures for gene deletion. We also thank Naomi Iverson for assistance with construction of plasmid vectors.

References

Alphey, L., and Parry, H.D. (1995) Making nucleic acid probes. In *DNA Cloning 1*. Glover, D.M., and Hames, B.D. (eds). New York.: Oxford University Press, pp. 121–123.

Bapteste, E., Brochier, C., and Boucher, Y. (2005) Higher-level classification of the Archaea: evolution of methanogenesis and methanogens. *Archaea* **1:** 353–363.

Baykov, A.A., Tuominen, H.K., and Lahti, R. (2011) The CBS domain: a protein module with an emerging prominent role in regulation. *ACS Chem Biol* **6:** 1156–1163.

Beck, B.J., and Downs, D.M. (1998) The *apbE* gene encodes a lipoprotein involved in thiamine synthesis in *Salmonella typhimurium*. *J Bacteriol* **180:** 885–891.

Blank, C.E. (2009) Phylogenomic dating – the relative antiquity of archaeal metabolic and physiological traits. *Astrobiology* **9:** 193–219.

Borrel, G., O'Toole, P.W., Harris, H.M.B., Peyret, P., Brugère, J.-F., and Gribaldo, S. (2013) Phylogenomic data support a seventh order of Methylotrophic methanogens and provide insights into the evolution of Methanogenesis. *Genome Biol Evol* **5:** 1769–1780.

Buan, N., Kulkarni, G., and Metcalf, W. (2011) Genetic methods for methanosarcina species. *Methods Enzymol* **494:** 23–42.

Farahi, K., Pusch, G.D., Overbeek, R., and Whitman, W.B. (2004) Detection of lateral gene transfer events in the prokaryotic tRNA synthetases by the ratios of evolutionary distances method. *J Mol Evol* **58:** 615–631.

Galagan, J.E., Nusbaum, C., Roy, A., Endrizzi, M.G., Macdonald, P., FitzHugh, W., *et al*. (2002) The genome of *M. acetivorans* reveals extensive metabolic and physiological diversity. *Genome Res* **12:** 532–542.

Guss, A.M., Rother, M., Zhang, J.K., Kulkarni, G., and Metcalf, W.W. (2008) New methods for tightly regulated gene expression and highly efficient chromosomal integration of cloned genes for *Methanosarcina* species. *Archaea* **2:** 193–203.

Habicht, K.S., Gade, M., Thamdrup, B., Berg, P., and Canfield, D.E. (2002) Calibration of sulfate levels in the archean ocean. *Science* **298:** 2372–2374.

Hauenstein, S.I., and Perona, J.J. (2008) Redundant synthesis of cysteinyl-tRNACys in *Methanosarcina mazei*. *J Biol Chem* **283:** 22007–22017.

Helgadóttir, S., Sinapah, S., Söll, D., and Ling, J. (2012) Mutational analysis of Sep-tRNA:Cys-tRNA synthase reveals critical residues for tRNA-dependent cysteine formation. *FEBS Lett* **586:** 60–63.

Hidese, R., Mihara, H., and Esaki, N. (2011) Bacterial

cysteine desulfurases: versatile key players in biosynthetic pathways of sulfur-containing biofactors. *Appl Microbiol Biotechnol* **91:** 47–61.
Kadaba, N.S., Kaiser, J.T., Johnson, E., Lee, A., and Rees, D.C. (2008) The high-affinity *E. coli* methionine ABC transporter: structure and allosteric regulation. *Science* **321:** 250–253.
Kohler, P.R.A., and Metcalf, W.W. (2012) Genetic manipulation of *Methanosarcina* spp. *Front Microbiol* **3:** 259.
Kolter, R., Inuzuka, M., and Helinski, D.R. (1978) Trans-complementation-dependent replication of a low molecular weight origin fragment from plasmid R6K. *Cell* **15:** 1199–1208.
Liu, Y., Sieprawska-Lupa, M., Whitman, W.B., and White, R.H. (2010) Cysteine is not the sulfur source for iron-sulfur cluster and methionine biosynthesis in the methanogenic archaeon *Methanococcus maripaludis*. *J Biol Chem* **285:** 31923–31929.
Liu, Y., Beer, L.L., and Whitman, W.B. (2012a) Methanogens: a window into ancient sulfur metabolism. *Trends Microbiol* **20:** 251–258.
Liu, Y., Dos Santos, P.C., Zhu, X., Orlando, R., Dean, D.R., Söll, D., and Yuan, J. (2012b) Catalytic mechanism of Sep-tRNA:Cys-tRNA synthase: sulfur transfer is mediated by disulfide and persulfide. *J Biol Chem* **287:** 5426–5433.
Liu, Y., Zhu, X., Nakamura, A., Orlando, R., Söll, D., and Whitman, W.B. (2012c) Biosynthesis of 4-thiouridine in tRNA in the methanogenic archaeon *Methanococcus maripaludis*. *J Biol Chem* **287:** 36683–36692.
Liu, Y., Long, F., Wang, L., Söll, D., and Whitman, W.B. (2014) The putative tRNA 2-thiouridine synthetase Ncs6 is an essential sulfur carrier in *Methanococcus maripaludis*. *FEBS Lett* **588:** 873–877.
Lucas, M., Encinar, J.A., Arribas, E.A., Oyenarte, I., García, I.G., Kortazar, D., *et al*. (2010) Binding of S-methyl-5′-thioadenosine and S-adenosyl-L-methionine to protein MJ0100 triggers an open-to-closed conformational change in its CBS motif pair. *J Mol Biol* **396:** 800–820.
Marchler-Bauer, A., and Bryant, S.H. (2004) CD-Search: protein domain annotations on the fly. *Nucleic Acids Res* **32:** W327–W331.
Marchler-Bauer, A., Lu, S., Anderson, J.B., Chitsaz, F., Derbyshire, M.K., DeWeese-Scott, C., *et al*. (2011) CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res* **39:** D225–D229.
Metcalf, W.W., Zhang, J.K., Shi, X., and Wolfe, R.S. (1996) Molecular, genetic, and biochemical characterization of the *serC* gene of *Methanosarcina barkeri* Fusaro. *J Bacteriol* **178:** 5797–5802.
Miyajima, R., and Shiio, I. (1973) Regulation of aspartate family amino acid biosynthesis in *Brevibacterium flavum*. VII. Properties of homoserine O-transacetylase. *J Biochem (Tokyo)* **73:** 1061–1068.
Mueller, E.G. (2006) Trafficking in persulfides: delivering sulfur in biosynthetic pathways. *Nat Chem Biol* **2:** 185–194.
O'Donoghue, P., Sethi, A., Woese, C.R., and Luthey-Schulten, Z.A. (2005) The evolutionary history of Cys-tRNACys formation. *Proc Natl Acad Sci USA* **102:** 19003–19008.
Overmars, L., Kerkhoven, R., Siezen, R.J., and Francke, C. (2013) MGcV: the microbial genomic context viewer for comparative genome analysis. *BMC Genomics* **14:** 209.
Parry, H.D., and Alphey, L. (1995) The utilization of cloned DNAs to study gene organization and expression. In *DNA Cloning 1*. Glover, D.M., and Hames, B.D. (eds). New York.: Oxford University Press, pp. 143–158.
Sarmiento, F., Mrázek, J., and Whitman, W.B. (2013) Genome-scale analysis of gene function in the hydrogenotrophic methanogenic archaeon *Methanococcus maripaludis*. *Proc Natl Acad Sci USA* **110:** 4726–4731.
Sauerwald, A., Zhu, W., Major, T.A., Roy, H., Palioura, S., Jahn, D., *et al*. (2005) RNA-dependent cysteine biosynthesis in archaea. *Science* **307:** 1969–1972.
Sekowska, A., Kung, H.F., and Danchin, A. (2000) Sulfur metabolism in *Escherichia coli* and related bacteria: facts and fiction. *J Mol Microbiol Biotechnol* **2:** 145–177.
Skovran, E., and Downs, D.M. (2003) Lack of the ApbC or ApbE protein results in a defect in Fe-S cluster metabolism in *Salmonella enterica* serovar Typhimurium. *J Bacteriol* **185:** 98–106.
Sowers, K.R., Boone, J.E., and Gunsalus, R.P. (1993) Disaggregation of *Methanosarcina* spp. and growth as single cells at elevated osmolarity. *Appl Environ Microbiol* **59:** 3832–3839.
Tatusov, R.L., Koonin, E.V., and Lipman, D.J. (1997) A genomic perspective on protein families. *Science* **278:** 631–637.
Thauer, R.K. (1998) Biochemistry of methanogenesis: a tribute to Marjory Stephenson. 1998 Marjory Stephenson Prize Lecture. *Microbiology* **144** (Part 9): 2377–2406.
White, R.H. (2003) The biosynthesis of cysteine and homocysteine in *Methanococcus jannaschii*. *Biochim Biophys Acta* **1624:** 46–53.
Yuan, J., Hohn, M.J., Sherrer, R.L., Palioura, S., Su, D., and Söll, D. (2010) A tRNA-dependent cysteine biosynthesis enzyme recognizes the selenocysteine-specific tRNA in *Escherichia coli*. *FEBS Lett* **584:** 2857–2861.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's web-site.