JOURNAL OF BACTERIOLOGY, Dec. 1996, p. 6849–6856 Vol. 178, No. 23
0021-9193/96/$04.00+0

Analysis of the CO Dehydrogenase/Acetyl-Coenzyme A Synthase Operon of *Methanosarcina thermophila*

JULIE A. MAUPIN-FURLOW[1,2]† AND JAMES G. FERRY[1]*

Department of Biochemistry and Molecular Biology, Pennsylvania State University, University Park, Pennsylvania 16802-4500,[1] and Department of Microbiology and Cell Science, University of Florida, Gainesville, Florida 32611[2]

Received 3 June 1996/Accepted 23 September 1996

The *cdhABC* genes encoding the respective α, ε, and β subunits of the five-subunit (α, β, γ, δ, and ε) CO dehydrogenase/acetyl-coenzyme synthase (CODH/ACS) complex from *Methanosarcina thermophila* were cloned and sequenced. Northern (RNA) blot analyses indicated that the *cdh* genes encoding the five subunits and an open reading frame (ORF1) with unknown function are cotranscribed during growth on acetate. Northern blot and primer extension analyses suggested that mRNA processing and multiple promoters may be involved in *cdh* transcript synthesis. The putative CdhA (α subunit) and CdhB (ε subunit) proteins each have 40% identity to CdhA and CdhB of the CODH/ACS complex from *Methanosaeta soehngenii*. The *cdhC* gene encodes the β subunit (CdhC) of the CODH/ACS complex from *M. thermophila*. The N-terminal 397 amino acids of CdhC are 42% identical to the C-terminal half of the α subunit of CODH/ACS from the acetogenic anaerobe *Clostridium thermoaceticum*. Sequence analysis suggested potential structures and functions for the previously uncharacterized β subunit from *M. thermophila*. The deduced protein sequence of ORF1, located between the *cdhC* and *cdhD* genes, has 29% identity to NifH2 from *Methanobacterium ivanovii*.

The five-subunit (α, β, γ, δ, and ε) CO dehydrogenase/acetyl-coenzyme A (CoA) synthase (CODH/ACS) complex is central to the pathway for the fermentation of acetate to CH_4 and CO_2 in the methanoarchaeon *Methanosarcina thermophila* (reviewed in reference 8), where it functions to cleave the C-C and C-S bonds in the acetyl moiety of acetyl-CoA, oxidize the carbonyl group to CO_2, and transfer the methyl group to tetrahydrosarcinapterin (THSPt). The complex also catalyzes CO oxidation-CO_2 reduction (CO dehydrogenase activity) and synthesis of acetyl-CoA with CO, a methyl group, and CoA; however, the synthesis activity is not important in the fermentation pathway. A membrane-bound methyltransferase transfers the methyl group from methyl-THSPt to CoM. The methyl-CoM is reductively demethylated to CH_4 with electrons that originate from oxidation of the carbonyl group of acetyl-CoA to CO_2 by the CODH/ACS enzyme complex. Detergent treatment resolves the five-subunit CODH/ACS complex into two enzyme components: the nickel–iron-sulfur (αε) component contains the α and ε subunits, and the corrinoid–iron-sulfur (γδ) component contains the γ and δ subunits of the complex (1). The β subunit is also resolved, but it is unstable and has not been characterized. The resolved αε component has CO dehydrogenase activity and, therefore, is proposed to oxidize the carbonyl group of acetyl-CoA to CO_2 (1). Electron paramagnetic resonance (EPR) spectroscopy identifies three metal clusters (A, B, and C) in the αε component (21) which have EPR spectroscopic properties indistinguishable from those of clusters A, B, and C in the well-characterized $\alpha_2\beta_2$ CODH/ACS from *Clostridium thermoaceticum* (reviewed in reference 29). The CODH/ACS from *C. thermoaceticum* functions to synthesize acetyl-CoA from CoA, CO, and a methyl group donated by a corrinoid–iron-sulfur protein. Cluster A, the proposed site for synthesis or cleavage of the C-C and C-S bonds of acetyl-CoA, is a novel Ni-X-[Fe_4-S_4] cluster, where X is an unknown bridging atom (27). Cluster C, the proposed site for CO dehydrogenase activity, is also a bimetallic Ni-X-[Fe_4-S_4] cluster (16). Cluster B is a conventional Fe_4S_4 center thought to shuttle electrons to and from cluster C. On the basis of EPR evidence for metal clusters A, B, and C in the αε component of the *M. thermophila* CODH/ACS complex, it is proposed that the αε component cleaves acetyl-CoA and oxidizes the carbonyl group to CO_2 (21). Biochemical characterization of the γδ component (17, 24) strongly suggests that the αε component transfers the methyl group to the corrinoid-containing δ subunit of the γδ component, which in turn methylates THSPt.

Methanosarcina barkeri also contains a five-subunit CODH/ACS enzyme complex with properties nearly identical to those of the complex from *M. thermophila* (10–14). *Methanosaeta* (formerly *Methanothrix*) is the only other genus capable of fermenting acetate to CH_4 and CO_2. The CODH/ACS complex from *Methanosaeta soehngenii* contains an $\alpha_2\beta_2$ component similar to the αε component of the CODH/ACS complex from *M. thermophila* (18). The $\alpha_2\beta_2$ component from *M. soehngenii* has CO dehydrogenase and acetyl-CoA cleavage activities; however, EPR spectroscopy identifies only clusters B and C (18). Although the ability to cleave acetyl-CoA suggests that cluster A is present in the $\alpha_2\beta_2$ component from *M. soehngenii*, EPR spectroscopic evidence for cluster A has not been obtained; however, EPR spectroscopy identifies a high-spin signal ascribed to an Fe_6S_6 prismane cluster. Recently, an enzyme with CODH activity from acetate-grown *Methanosarcina frisia* which has a subunit composition identical to the $\alpha_2\beta_2$ component from the *M. soehngenii* CODH/ACS complex and contains Ni, Fe, and acid-labile S was described (7). The amino acid sequences deduced from the genes encoding both subunits have 60% identity and 50% similarity with the sequences deduced for subunits of the *M. soehngenii* $\alpha_2\beta_2$ component. EPR spectroscopy indicates two interacting Fe_4S_4 clusters for the *M. frisia* enzyme (7). No evidence for a cluster A was reported;

* Corresponding author. Mailing address: Department of Biochemistry and Molecular Biology, Pennsylvania State University, University Park, PA 16802-4500. Phone: (814) 863-5721. Fax: (814) 863-7024. Electronic mail address: jgf3@psuvm.psu.edu.

† Present address: Department of Microbiology and Cell Science, University of Florida, Gainesville, FLA 32611.

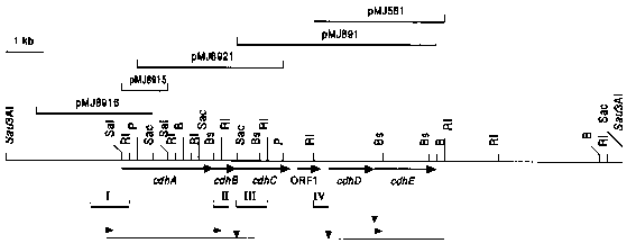

FIG. 1. A partial map of endonuclease restriction sites for a 16-kb *Sau*3A1 genomic DNA fragment from *M. thermophila*. DNA fragments used to construct plasmids pMJ581, pMJ891, pMJ8921, pMJ8915, and pMJ8916 (see Materials and Methods) are shown directly above the physical map. The locations of the *cdhABCDE* genes and ORF1, comprising the CODH/ACS operon, are shown directly below the physical map. The arrows indicate the direction of transcription. The locations of DNA fragments used as probes in Northern blotting experiments are indicated (I through IV). The 8.7-kb mRNA species, represented by a line below the probes, is located relative to complementary DNA. Symbols and abbreviations: ▶, putative transcriptional start site; ▼, putative mRNA processing site(s); Bs, *Bst*EII; RI, *Eco*RI; B, *Bam*HI; P, *Pst*I; Sac, *Sac*I; Sal, *Sal*I.

however, a high-spin EPR signal with properties similar to the proposed Fe_6S_6 prismane cluster present in the *M. soehngenii* component was described. Clusters B and C are present in the monomeric CODH from *Rhodospirillum rubrum* (36). This enzyme has CO dehydrogenase activity but does not have a cluster A and, therefore, is incompetent to synthesize or cleave acetyl-CoA (3).

The genes encoding the αβ subunits of CODH/ACS (25) and the αβ subunits of the corrinoid–iron-sulfur protein (22) from *C. thermoaceticum*, the gene encoding the CODH from *R. rubrum* (19), and the genes encoding the αβ subunits of the $\alpha_2\beta_2$ CODH enzymes from *M. soehngenii* (6) and *M. frisia* (7) have been cloned and sequenced. The genes encoding the γδ subunits (corrinoid–iron-sulfur component) of the CODH/ACS complex from *M. thermophila* have been cloned and sequenced (24); however, the sequences of the genes encoding the ε and β subunits have not been reported. Only a partial sequence of the gene encoding the α subunit of the *M. thermophila* αε component has been published (35). Primer extension analysis indicates a transcriptional start site upstream of the *cdhA* gene encoding the α subunit (35). Here we present the complete sequence of the genes encoding the α, ε, and β subunits from *M. thermophila* and show that all five subunits of the CODH/ACS complex are cotranscribed with an open reading frame (ORF) with unknown function. Sequence analysis suggests potential structures and functions for the previously uncharacterized β subunit. Northern (RNA) blot and primer extension analysis suggest that mRNA processing and multiple promoters may be involved in *cdh* transcript synthesis.

MATERIALS AND METHODS

DNA sequence determination and protein analysis. A 16-kb *Sau*3A1 *M. thermophila* genomic fragment in λ phage vector GEM-11 (λMJ583), containing the complete *cdh* operon, was isolated using a *cdhD*-specific synthetic oligonucleotide as described previously (24). The DNA sequence of both strands of the *cdh* region was determined according to the dideoxy chain termination method using Sequenase version 2.0 T7 DNA polymerase (US Biochemical Corp., Cleveland, Ohio). λ MJ583 DNA was digested with restriction endonucleases, and *cdh*-specific DNA fragments were isolated and ligated into plasmid vector pUC19. The resulting plasmid constructs include pMJ581, pMJ8916, pMJ8915, pMJ8921, and pMJ891, which contain 3.4-kb *Eco*RI, 3.1-kb *Sac*I, 1.3-kb *Sal*I, 3.6-kb *Pst*I, and 5.3-kb *Sac*I-to-*Bam*HI genomic DNA fragments, respectively (Fig. 1). These plasmids, restriction endonuclease subclones, and exonuclease III deletion derivatives of plasmids pMJ581, pMJ8921, and pMJ891 were used as templates for DNA sequence analysis. Synthetic oligonucleotides (University of Florida Interdisciplinary Center for Biotechnology Research and Department of Microbiology and Cell Science, Gainesville, Fla.) based on the *cdh* gene sequence were used to complete the sequences of both DNA strands. The subunits of purified CO dehydrogenase (37) were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), and the N-terminal sequences were determined by automated Edman degradation as previously described (24).

Deduced amino acid sequences were compared to protein sequences available in the GenBank, EMBL, and SwissProt databases at the National Center for Biotechnology Information (Bethesda, Md.) by using the network servers BLAST (Genetics Computer Group [GCG] Program, Madison, Wis.) and BLITZ (MPsrch version 1.3, Biocomputing Research Unit, University of Edinburgh, United Kingdom). BESTFIT (GCG Program) and Genpro 5.0 (Hoefer Scientific, San Francisco, Calif.) were used for protein sequence alignment. ISOELECTRIC was used for estimation of pI values (GCG Program).

Northern hybridization and primer extension analyses. Total RNA was isolated from mid-log-phase *M. thermophila* cells grown in marine medium supplemented with 100 mM acetate, fractionated by electrophoresis on formaldehyde-agarose gels, and transferred to membranes as previously described (24). Northern hybridization probes included an 892-bp *Ssp*I-to-*Eco*RI DNA fragment spanning nucleotides −575 to +317 of *cdhA* (I), a 349-bp *Hin*cII-to-*Sca*I DNA fragment spanning nucleotides +7 to +355 of *cdhB* (II), a 712-bp *Sac*I-to-*Eco*RI DNA fragment spanning nucleotides +74 to +785 of *cdhC* (III), and a 371-bp *Eco*RI-to-*Bgl*II fragment spanning nucleotides +489 to +163 downstream of ORF1 (IV) (Fig. 1) which were labeled with [^{32}P]α-dATP (DuPont-New England Nuclear Research Products, Boston, Mass.) using random hexanucleotides and the Klenow fragment of DNA polymerase (Boehringer Mannheim Biochemical, Indianapolis, Ind.). Membrane equilibration, RNA-DNA hybridization, and preparation of wash stringencies were as previously described (24).

The 5′ ends of *cdhB*-, *cdhC*-, and *cdhE*-specific mRNA were mapped by avian myeloblastosis virus reverse transcriptase primer extension analysis as described elsewhere (24). Synthetic oligonucleotides (Pennsylvania State University Biotechnology Institute, University Park, Pa.) complementary to nucleotides +40 to +64 of *cdhB* (5′-TTGCCTTTGAGGTGGTCACCCCGTA-3′), complementary to nucleotides +40 to +64 of *cdhC* (5′-TTCCTTCCTTTCTTACTCTTTCTCC-3′), complementary to nucleotides +40 to +64 of ORF1 (5′-TAGTATCTGC ATCTGCATCAACTGC-3′), and complementary to nucleotides +41 to +65 of *cdhE* (5′-CACATTCTCCACAGTTGGTCTGGGG-3′) were annealed to total RNA isolated from acetate-grown *M. thermophila.*

Materials. Biochemicals were purchased from Sigma Chemical Co. (St. Louis, Mo.). Other organic and inorganic chemicals were from Fisher Scientific (Pittsburgh, Pa.) and were analytical grade. Restriction endonucleases and DNA-modifying enzymes were from New England Biolabs (Beverly, Mass.), Promega (Madison, Wis.), or US Biochemical.

Nucleotide sequence accession number. The GenBank accession number U66032 has been assigned to an 8,321-bp region containing the *cdh* operon.

RESULTS AND DISCUSSION

Genetic organization and transcriptional analysis of the *cdh* (carbon monoxide dehydrogenase/acetyl-CoA synthase) operon. It was previously reported that the genes (*cdhDE*) encoding the δ and γ subunits of the CODH/ACS complex are located on a 16-kb *Sau*3A1 genomic fragment of *M. thermophila* (24) (Fig. 1). Previously reported Northern hybridizations and primer extension analysis suggest that an 8.7-kb primary transcript is processed, yielding a 3.0-kb *cdhDE*-specific mRNA species (24). Other features of the primary transcript were not reported. Sequence analysis of DNA upstream of *cdhD* revealed four additional ORFs, each preceded by an apparent ribosome binding site (Fig. 1 and 2). The N-terminal amino acid sequences of the α, ε, and β subunits of the CODH/ACS complex corresponded to the deduced amino acid sequences of three ORFs which identified the *cdhABC* genes encoding the α, ε, and β subunits, respectively (Fig. 2). The DNA sequence of 1,656 nucleotides, including nucleotides +1 to +318 of the *cdhA* gene from *M. thermophila*, was reported previously (35) and is identical to the DNA sequence presented here. The fourth ORF with unknown function (ORF1) is located between *cdhC* and *cdhD* (Fig. 1 and 2).

Genomic organization of *cdhABCDE* and ORF1 (Fig. 1 and 2) suggested that the six genes are cotranscribed; therefore, DNA fragments (I to IV) (Fig. 1) containing portions of the coding regions for *cdhA*, *cdhB*, *cdhC*, and ORF1 were used as probes in Northern hybridizations. All four probes hybridized to an 8.7-kb mRNA species (Fig. 3), consistent with the 8.2-kb coding region of *cdhABCDE* and ORF1 (Fig. 1), which sug-

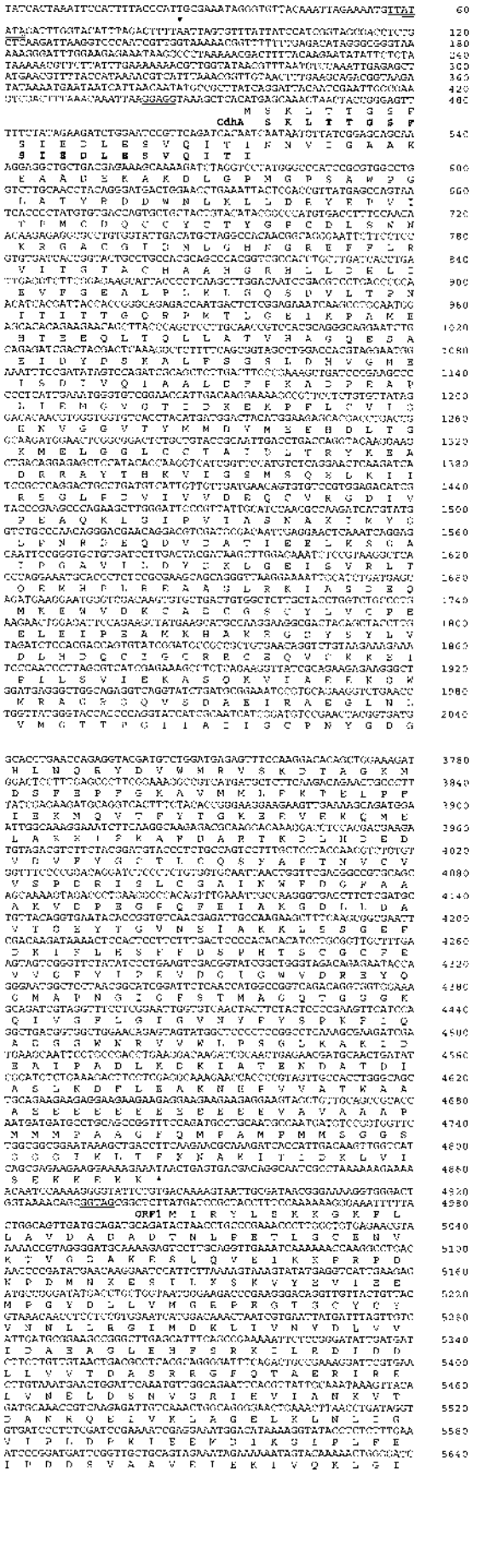

FIG. 2. Nucleic acid sequences and predicted amino acid sequences of *cdhA*, *cdhB*, *cdhC*, and ORF1 of the *cdh* operon from *M. thermophila*. The DNA is presented in the 5′-to-3′ direction. The predicted amino acid sequences are shown in single-letter code directly below the first base of each codon. The N-terminal sequences experimentally determined for the α, ε, and β subunits of the CODH/ACS complex are represented below the sequences predicted for CdhA, CdhB, and CdhC, where X is an uncertain assignment. Symbols: ▼, bases complementary to the 5′ ends of the primer extension products (Fig. 4, 5, and 6) (24, 35); =, putative archaeal box A promoter sequences; −, presumed ribosome-binding sites; ∗, initial base of translation stop codon; . . ., gaps introduced within *cdhD* and truncating *cdhE*.

gested that *cdhABCDE* and ORF1 are cotranscribed as an operon. The previously reported 9.5-kb *cdhA*-specific transcript supports this conclusion, although DNA standards which may have overestimated the size of the transcript were used (24). All probes, except probe II, hybridized to smaller mRNA species in addition to the 8.7-kb transcript (Fig. 3). The *cdhA* probe (I) hybridized with approximately 2.6- and 1.5-kb mRNA. The *cdhC*- and ORF1-specific (III and IV) probes hybridized with 5.5- to 5.0-kb mRNA. Probe IV also hybridized with RNA of approximately 2.8 kb. Previous results showed

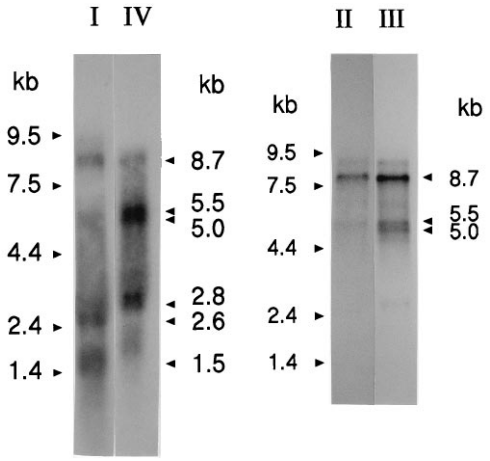

FIG. 3. Northern analysis of the *cdh* operon of *M. thermophila*. Total RNA isolated from acetate-grown *M. thermophila* was probed with DNA fragments containing portions of the coding regions for *cdhA* (I), *cdhB* (II), *cdhC* (III), and ORF1 (IV) (Fig. 1) (see Materials and Methods). RNA molecular weight markers (left) and transcript sizes (right) are indicated.

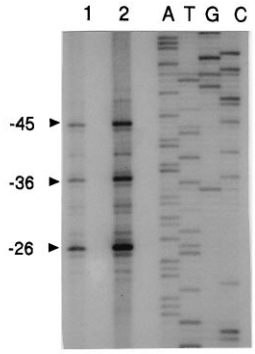

FIG. 5. Mapping of the 5′ end of *cdhC*-specific mRNA. Primer extension products with total RNA from acetate-grown *M. thermophila* cells (lanes 1 and 2) and DNA sequencing reactions (lanes A, T, G, and C) in which the same oligonucleotide (see Materials and Methods) was used. The numbers at the left indicate the bases relative to the putative translational start site (+1) of *cdhC* (Fig. 2) that were complementary to the 5′ ends of the primer extension products.

that a *cdhD*-specific probe hybridizes with approximately 5.0- and 3.0-kb mRNA species and a *cdhE*-specific probe hybridizes with approximately 5.0-, 3.0-, and 1.6-kb species (24).

The <8.7-kb transcripts identified by Northern analysis utilizing probes specific for each gene (Fig. 3) suggested that mRNA processing and/or multiple promoters are involved in transcript synthesis; thus, transcription of the *cdh* operon was investigated further by primer extension analysis. Results of previous primer extension analysis utilizing a primer within *cdhA* identified a putative transcriptional start located −370 bp relative to the *cdhA* translational start (35) (Fig. 2). When an oligonucleotide primer within *cdhB* was used, the 5′ ends of mRNA mapped to G nucleotides −189 and −186 and A nucleotides −185 and −184 relative to the predicted translational start site of the *cdhB* gene (Fig. 2 and 4). Transcription of genes from the Archaea typically initiates 22 to 27 bp downstream of a box A sequence [consensus, 5′-TTA(T/A)ATA-3′] at a purine-pyrimidine dinucleotide, usually in a box B sequence (consensus, 5′-ATGC-3′) (33). One of the 5′ ends (nucleotide −189) mapped to a purine-pyrimidine dinucleotide which was located 20 bp downstream of a 6-bp sequence (5′-TATATT-3′) with identity to the consensus box A (Fig. 2), suggesting a potential transcription start site; however, mRNA processing of the 8.7-kb transcript at this site cannot be ruled out. The primer extension results suggested that an approximately 5.5-kb mRNA species should be detected in Northern hybridizations utilizing probe II; however, only a faint band was detected at approximately 5.5 kb (Fig. 3), suggesting that this transcript, if produced, is less stable. Primer extension of RNA using an oligonucleotide within *cdhC* mapped the 5′ ends of mRNA to G nucleotides −45 and −36 and an A nucleotide, −26, relative to the presumed translational start codon of the *cdhC* gene (Fig. 2 and 5). An archaeal consensus promoter was not identified 22 to 27 bp upstream of the 5′ ends, which suggests the possibility of mRNA processing. Previous primer extension analysis yielded multiple signals complementary to nucleotides upstream of *cdhD* (24) (Fig. 2), none of which were at the expected position 20 to 27 bp downstream of a recognizable box A sequence, which also suggests the possibility of mRNA processing of a larger transcript. Primer extension with an oligonucleotide within *cdhE* mapped the 5′ ends of mRNA to A, T, C, and G nucleotides-163, −122, −121, and −85, respectively, relative to the putative *cdhE* translational start site (Fig. 2 and 6). Only the G (−85) nucleotide is the expected distance (27 bp) downstream of a sequence (5′-TTTATGAT-3′) with identity to the consensus box A sequence; thus, the 5′ ends of mRNA which mapped to nucleotides upstream of *cdhE* could have resulted from both mRNA processing of a larger mRNA species and initiation of transcription. These results are consistent with those of previously reported Northern hybridizations in which *cdhE*-specific, but not *cdhD*-specific, probes hybridized to mRNA species less than 3.0 kb (24). Primer extension products were not identified when an oligonucleotide within the ORF1 coding region was used (data not shown). The Northern blotting and primer extension analyses presented here and elsewhere (Fig. 2) (24) support processing of the 8.7-kb primary transcript and possibly additional transcriptional start sites. The exact stoichiometry and mechanism of assembly for subunits of the CODH/ACS complex from *M. thermophila* are unknown. The two potential *cdh* promoters and three sites of putative *cdh*-mRNA processing (Fig. 1) may provide smaller transcripts leading to the requisite translation for proper assembly of the complex.

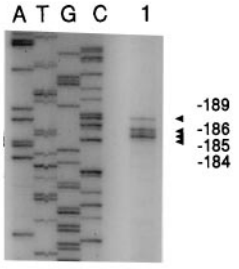

FIG. 4. Mapping of the 5′ end of *cdhB*-specific mRNA. Primer extension products with total RNA from acetate-grown *M. thermophila* cells (lane 1) and DNA sequencing reactions (lanes A, T, G, and C) in which the same oligonucleotide (see Materials and Methods) was used. The numbers at the right indicate the bases relative to the putative translational start site (+1) of *cdhB* (Fig. 2) that were complementary to the 5′ ends of the primer extension products.

Sequence analysis of CdhA, CdhB, CdhC, and the ORF1 gene product. (i) CdhA and CdhB. The *cdhA* gene encodes a putative 806-amino-acid protein (CdhA) with a calculated pI of 5.12 (Fig. 2). The anhydrous molecular mass of 88,536 Da calculated for CdhA is in agreement with 89 kDa estimated by

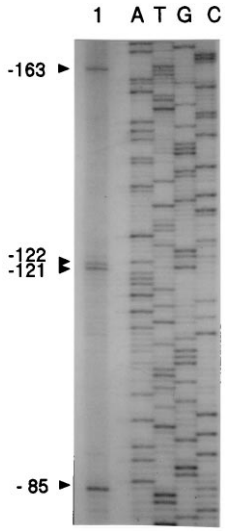

FIG. 6. Mapping of the 5′ end of *cdhE*-specific mRNA. Primer extension reaction products using total RNA from acetate-grown *M. thermophila* cells (lane 1) and DNA sequencing reactions (lanes A, T, G, and C) in which the same oligonucleotide (see Materials and Methods) was used. The numbers at the left indicate the bases relative to the putative translational start site (+1) of *cdhE* (Fig. 2) that were complementary to the 5′ ends of the primer extension products.

FIG. 7. Alignment of the predicted sequences of the CODH/ACS α subunits from *M. thermophila* (*Mt* CdhA) and *M. soehngenii* (*Ms* CdhA). Identical (:) and functionally similar (.) amino acid residues are noted. Dashes indicate gaps introduced in protein sequence alignment. Identical or functionally similar amino acid residues common to the archaeal CODH/ACS α subunits, Ni-containing CODH/ACS from *R. rubrum* (19), and the β subunit of the CODH/ACS from *C. thermoaceticum* (25) are highlighted. Conserved cysteine and histidine residues are underlined. The GenBank accession number for *Ms* CdhA is M55280 (6).

SDS-PAGE for the α subunit in the αε component of the CODH/ACS complex isolated from *M. thermophila* (37). The *cdhB* gene encodes a 170-amino-acid protein (CdhB) with a pI estimated at 10.11 (Fig. 2). The calculated anhydrous molecular mass of 18,739 Da is in agreement with the 19 kDa determined by SDS-PAGE for the ε subunit of the αε component (37). Only one histidine and no cysteine residues are present in the ε subunit; thus, metal clusters A, B, and C identified by EPR spectroscopy are apparently localized to the α subunit which contains several cysteine and histidine residues likely to coordinate metal clusters A, B, and C (Fig. 7). CdhA (α subunit) has a high level of identity and similarity to the analogous 89-kDa α subunits of the $\alpha_2\beta_2$ components from *M. soehngenii* (40 and 60%) (Fig. 7) and acetate-grown *M. frisia* (91 and 94%) (data not shown). Likewise, CdhB (ε subunit) from *M. thermophila* has a high level of identity and similarity to the analogous 19-kDa β subunits of the $\alpha_2\beta_2$ components from *M. soehngenii* (40 and 65%) (Fig. 8) and acetate-grown *M. frisia* (91 and 96%) (data not shown). The identity among the α subunits extends to cysteine and histidine residues, suggesting structural conservation of metal centers. Although EPR evidence for cluster A has been elusive for the *M. soehngenii* and *M. frisia* components, the presence of this cluster is suggested by sequence conservation with the *M. thermophila* component. The acetyl-CoA–CO exchange (C-C and C-S cleavage-synthesis) activity reported for the *M. soehngenii* $\alpha_2\beta_2$ component is also consistent with the presence of a cluster A. It is curious that a high-spin prismane-like EPR signal is observed for the *M. soehngenii* and *M. frisia* $\alpha_2\beta_2$ components but not for the analogous αε component from *M. thermophila*. A relationship between the absence of a cluster A EPR signal and the presence of a high-spin signal has not been investigated. Cluster A is located in the α subunit (38, 39) of the clostridial $\alpha_2\beta_2$ CODH/ACS; however, sequence alignment of *M. thermophila* CdhA with the *C. thermoaceticum* α subunit revealed only 13% identity (data not shown). The EPR evidence for cluster A in *M. thermophila* CdhA is compelling (21). Although more structural information is needed to draw conclusions, the simplest explanation for the lack of sequence identity between CdhA and the clostridial α subunit is convergent evolution dictated by the chemical requirements for catalysis of the same reaction. Conservation of an active-site zinc center in the recently described new class of carbonic anhydrase from *M. thermophila* (20) is a clear example of convergent evolution, which is not an unexpected event when one considers that the aceotrophic methanoarchaea and acetogenic clostridia are at phylogenetic extremes.

The two CXXCXXCXXXC ferredoxin-like cysteine motifs in CdhA (α subunit) from *M. thermophila* (residues 417 to 427

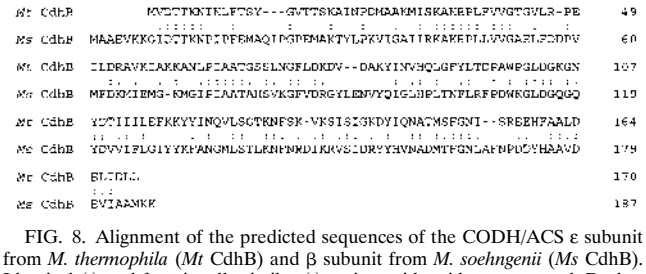

FIG. 8. Alignment of the predicted sequences of the CODH/ACS ε subunit from *M. thermophila* (*Mt* CdhB) and β subunit from *M. soehngenii* (*Ms* CdhB). Identical (:) and functionally similar (.) amino acid residues are noted. Dashes indicate gaps introduced in protein sequence alignment. The GenBank accession number for *Ms* CdhB is M55280 (6).

and 455 to 465, Fig. 7) show strict conservation with the CdhA sequences from *M. soehngenii* (Fig. 7) and *M. frisia* (data not shown) and are the most probable motifs for coordinating two B clusters; indeed, Fe_4S_4 EPR signals are observed at midpoint potentials of −444 and below −500 mV for the αε component, which is consistent with two B clusters (21). EPR spectroscopy of the $\alpha_2\beta_2$ CODH from *M. frisia* identifies two Fe_4S_4 clusters with weak dipolar interaction (7).

It was previously reported that the CdhA sequence from *M. soehngenii* has significant similarity to sequences deduced from the genes encoding the monomeric Ni-containing CODH from *R. rubrum* (19) (47%) and the β subunit of the CODH/ACS from *C. thermoaceticum* (25) (67%). It was recently reported that CdhA from *M. frisia* also has a high level of similarity with these proteins (7). Thus, it is not surprising that CdhA from *M. thermophila* also has a high level of similarity with the *R. rubrum* CODH (46%) and the β subunit of the CODH/ACS from *C. thermoaceticum* (49%) (data not shown). Several cysteine and histidine residues present in CdhA from *M. thermophila* (residues 76 to 94, 117 to 123, 183, 250, 277 to 278, 323, 523, 552, and 587, Fig. 7) are conserved among all of the above proteins, which have in common the ability to oxidize CO (3, 7, 28, 37). The proposed site for CO oxidation is cluster C (2, 16, 21, 36). Although similar in structure to cluster A, cluster C does not catalyze acetyl-CoA synthesis or cleavage. On the basis of the structural and functional similarities among the above proteins, it is likely that one or more of the conserved motifs in CdhA (Fig. 7) are involved in the structure of cluster C. All of the above proteins have a conventional Fe_4S_4 center (cluster B) proposed to shuttle electrons to or from cluster C; however, only CdhA from the methanoarchaea have standard ferredoxin-like cysteine motifs likely to coordinate Fe_4S_4 centers. The assignment of structure and function to amino acid motifs conserved among these proteins is further complicated by EPR evidence for cluster A in the CdhA (α subunit) from *M. thermophila* and prismane-like Fe_6S_6 clusters in CdhA from *M. soehngenii* and *M. frisia*. Neither cluster A nor a prismane-like cluster is present in the CODH from *R. rubrum* or the β subunit of the CODH/ACS from *C. thermoaceticum*.

(ii) CdhC. The *cdhC* gene encodes a 469-amino-acid protein with a calculated pI of 4.48 (Fig. 2). The CdhC (β subunit) calculated anhydrous molecular mass of 51,881 Da is significantly different from the 71 kDa estimated by SDS-PAGE of the purified CODH/ACS β subunit (37). The highly charged nature of the C terminus of CdhC suggests that the β subunit may bind less SDS than the reference proteins producing anomalous migration during SDS-PAGE (26). The function of the glutamic acid patch (residues 404 to 415, Fig. 9) in the C terminus is unknown, but the patch may be a signal for proteolysis (32) by the recently described proteasome from *M. thermophila* (23). The C-terminal residues EKKEKK (residues 464 to 469, Fig. 9) resemble the KEKE motif proposed to be involved in protein-protein interactions (31).

The N-terminal 397 amino acids of CdhC are highly identical (42%) and similar (68%) to residues 317 to 729 of the α subunit from the CODH/ACS of *C. thermoaceticum* (Fig. 9). Both proteins contain highly conserved four- and two-cysteine motifs (CdhC residues 189 to 211 and 278 to 280, Fig. 9). Apart from these motifs, there are only four other cysteine residues in the clostridial α subunit which are located in the N-terminal half (residues 1 to 316) and are spaced randomly (data not shown); thus, the conserved cysteine motifs (Fig. 9) are the most likely to coordinate cluster A in the clostridial α subunit (38). The high level of identity between *M. thermophila* CdhC (β subunit) and the clostridial α subunit (Fig. 9) indicates that

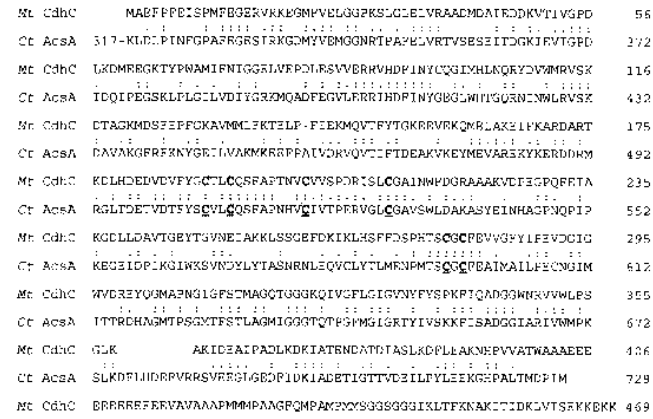

FIG. 9. Alignment of the predicted sequences of the CODH/ACS β subunit from *M. thermophila* (*Mt* CdhC) and α subunit from *C. thermoaceticum* (*Ct* AcsA). Identical (:) and functionally similar (.) amino acid residues are noted. Dashes indicate gaps introduced in protein sequence alignment. Conserved cysteine residues are highlighted and underlined. The GenBank accession number for AcsA is M62727 (25).

CdhC could potentially contain a cluster A. The Ni content reported for the five-subunit CODH/ACS complex from *M. thermophila* ranges from 2.3 to 3.6 atoms per mole of the complex (1, 37), a result that is consistent with a Ni-containing cluster in addition to Ni-containing clusters A and C previously identified in the αε component. Furthermore, a proteolytic fragment of the β subunit of the CODH/ACS complex from *M. barkeri* displays acetyl-CoA–CoA (C-S bond cleavage-synthesis) activity (13), which is also an activity ascribed to cluster A in the clostridial CODH/ACS. Alignment of CdhC with the β subunit from the clostridial CODH/ACS revealed no significant sequence identity (16%), suggesting no functional similarity with the clostridial β subunit which is thought to contain cluster C but not cluster A (39). Although biochemical evidence is needed for a firm conclusion, the results reported here strongly implicate the presence of a cluster A in the CdhC (β) subunit in addition to a cluster A in the CdhA (α) subunit of the CODH/ACS complex from *M. thermophila.*

Although there are similarities between the *M. thermophila* and *C. thermoaceticum* CODH/ACS that suggest a fundamentally common mechanism of acetyl-CoA synthesis or cleavage, there are also notable differences. Only one cluster A is known for the clostridial CODH/ACS, and the gene organization and subunit composition are significantly different. Furthermore, the acetyl-CoA–CoA exchange (C-S bond cleavage-synthesis) activity of the *M. thermophila* CODH/ACS complex is fivefold greater than that for the clostridial enzyme; conversely, the CO–acetyl-CoA exchange (C-C and C-S cleavage-synthesis) activity is 24-fold lower (30). These differences between the *M. thermophila* and clostridial CODH/ACS could reflect the different physiological roles for these enzymes. The *M. thermophila* CODH/ACS complex functions to cleave acetyl-CoA, whereas in the acetogenic clostridia CODH/ACS synthesizes acetyl-CoA. It is conceivable that in the *M. thermophila* CODH/ACS complex the acetyl group is cleaved from acetyl-CoA at cluster A in the β subunit and transferred to the αε component, where C-C bond cleavage takes place at cluster A and oxidation of the carbonyl group at cluster C. Consistent with this hypothesis is the acetyl-CoA–CoA exchange (C-S bond cleavage) activity reported for a proteolytic fragment of the β subunit from the *M. barkeri* CODH/ACS complex (13). Separation of C-C and C-S bond cleavage activities at the two cluster A sites could provide a mechanism for favoring cleav-

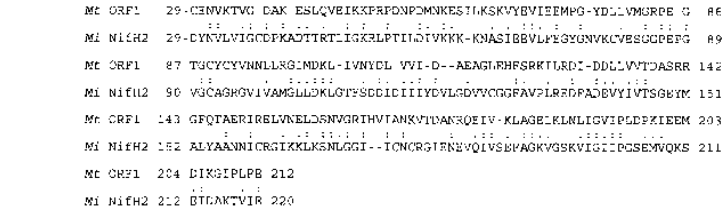

FIG. 10. Alignment of predicted ORF1 protein sequence from *M. thermophila* (*Mt*) and NifH2 protein from *M. ivanovii* (*Mi*). Dashes represent gaps introduced for alignment. Identical (:) and functionally similar (.) amino acid residues are indicated. The GenBank accession number for NifH2 is X07501 (34).

age of acetyl-CoA. This hypothesis also offers an explanation for the absence of acetyl-CoA synthesis activity of the *M. thermophila* αε component when resolved from the β subunit of the CODH/ACS complex (1). Regardless of the function of CdhC, the collective evidence suggests that the β subunit is required for acetyl-CoA cleavage. In a previous report (7), a 144-kbp DNA sequence was included downstream of the *cdhAB* genes from *M. frisia*, which contains an initiator codon followed by DNA with a deduced sequence 100% identical to the first 40 N-terminal residues of *M. thermophila* CdhC (data not shown). Thus, it appears that CdhC is indispensable for acetotrophic growth of methanosarcina, which further supports a function for CdhC in catalysis of acetyl-CoA cleavage or maturation of the CODH/ACS complex.

The high level of sequence identity (42%) between *M. thermophila* CdhC and the clostridial α subunit is surprising when it is considered that *M. thermophila* and *C. thermoaceticum* are classified separately at the highest taxonomic levels (Archaea and Bacteria domains). The high level of sequence identity is similar to the previously reported (24) sequence identities (37%) between the γδ subunits of the corrinoid–iron-sulfur component from the CODH/ACS complex of *M. thermophila* and the analogous subunits of the corrinoid–iron-sulfur protein from *C. thermoaceticum*. The high-level sequence identities can be explained by either horizontal gene transfer or common ancestral genes. There are other examples of significant levels of identity between proteins from methanosarcina and clostridia, including heat shock protein 70 (1) and several *nif* gene products (4, 5). These sequence identities have led to the recent proposal of close evolutionary linkage between the Archaea and gram-positive Bacteria (15).

(iii) ORF1 protein. An ORF with unknown function (ORF1) was located 119 bp downstream of *cdhC* and terminated 248 bp upstream of *cdhD* (Fig. 2). The predicted ORF1 protein sequence of 232 amino acids has significant identity (29%) and similarity (49%) to NifH2 from *Methanobacterium ivanovii* (Fig. 10) (34). The dimeric *Azotobacter vinelandii* NifH1 Fe protein, which transfers electrons to the nitrogenase MoFe protein, coordinates an Fe_4S_4 center utilizing two cysteines from each monomer (9). Of the three cysteine residues in the putative ORF1 protein, only residue 89 is conserved with *A. vinelandii* NifH1 and *M. ivanovii* NifH2 (Fig. 10); thus, it is not possible to predict that the putative ORF1 gene product contains an Fe-S center. Phylogenetic analysis of *nifH* gene products (4) indicates three families which function in nitrogen fixation. A fourth family, which includes *M. ivanovii* NifH2, consists of genes from the methanoarchaea which are located remotely from *nif* genes and presumed to serve functions other than those for nitrogen fixation (4); however, thus far there has been no indication of a potential function. The cotranscription of ORF1 with other *cdh* genes suggests that the putative ORF1 gene product may be a component, or be required for maturation, of an active CODH/ACS complex.

ACKNOWLEDGMENTS

We thank K. T. Shanmugam for the use of his laboratory at the University of Florida Department of Microbiology and Cell Science. We are also grateful to David Grahame, Stephen Ragsdale, and Kevin Sowers for supportive discussions.

This research was supported by grant 1RO1GM44661 from the National Institutes of Health and grant DE-FG05-87ER13730 from the Department of Energy–Basic Energy Sciences to J.G.F. J.M.-F. was supported by National Institutes of Health Individual National Research Service Award 1F32GM15877.

REFERENCES

1. **Abbanat, D. R., and J. G. Ferry.** 1991. Resolution of component proteins in an enzyme complex from *Methanosarcina thermophila* catalyzing the synthesis or cleavage of acetyl-CoA. Proc. Natl. Acad. Sci. USA **88:**3272–3276.
2. **Anderson, M. E., and P. A. Lindahl.** 1994. Organization of clusters and internal electron pathways in CO dehydrogenase from *Clostridium thermoaceticum*: relevance to the mechanism of catalysis and cyanide inhibition. Biochemistry **33:**8702–8711.
3. **Bonam, D., and P. W. Ludden.** 1987. Purification and characterization of carbon monoxide dehydrogenase, a nickel, zinc, iron-sulfur protein, from *Rhodospirillum rubrum*. J. Biol. Chem. **262:**2980–2987.
4. **Chien, Y. T., and S. H. Zinder.** 1994. Cloning, DNA sequencing, and characterization of a *nifD*-homologous gene from the archaeon *Methanosarcina barkeri* 227 which resembles *nifD1* from the eubacterium *Clostridium pasteurianum*. J. Bacteriol. **176:**6590–6598.
5. **Chien, Y. T., and S. H. Zinder.** 1996. Cloning, functional organization, transcript studies, and phylogenetic analysis of the complete nitrogenase structural genes (*nifHDK2*) and associated genes in the archaeon *Methanosarcina barkeri* 227. J. Bacteriol. **178:**143–148.
6. **Eggen, R. I. L., A. C. M. Gerrling, M. S. M. Jetten, and W. M. de Vos.** 1991. Cloning, expression, and sequence analysis of the genes for carbon monoxide dehydrogenase of *Methanothrix soehngenii*. J. Biol. Chem. **266:**6883–6887.
7. **Eggen, R. I. L., R. van Kranenburg, A. J. M. Vriesema, A. C. M. Gerrling, M. F. J. M. Verhagen, W. R. Hagen, and W. M. de Vos.** 1996. Carbon monoxide dehydrogenase from *Methanosarcina frisia* Göl. J. Biol. Chem. **271:**14256–14263.
8. **Ferry, J. G.** 1995. CO dehydrogenase. Ann. Rev. Microbiol. **49:**305–333.
9. **Georgiadis, M. M., H. Komiya, P. Chakrabarti, D. Woo, J. J. Kornuc, and D. C. Rees.** 1992. Crystallographic structure of the nitrogenase iron protein from *Azotobacter vinelandii*. Science **257:**1653–1659.
10. **Grahame, D. A.** 1991. Catalysis of acetyl-CoA cleavage and tetrahydrosarcinapterin methylation by a carbon monoxide dehydrogenase-corrinoid enzyme complex. J. Biol. Chem. **266:**22227–22233.
11. **Grahame, D. A.** 1993. Substrate and cofactor reactivity of a carbon monoxide dehydrogenase corrinoid enzyme complex. Stepwise reduction of iron-sulfur and corrinoid centers, the corrinoid $Co^{2+/1+}$ redox midpoint potential, and overall synthesis of acetyl-CoA. Biochemistry **32:**10786–10793.
12. **Grahame, D. A., and E. DeMoll.** 1995. Substrate and accessory protein requirements and thermodynamics of acetyl-CoA synthesis and cleavage in *Methanosarcina barkeri*. Biochemistry **34:**4617–4624.
13. **Grahame, D. A., and E. DeMoll.** 1996. Partial reactions catalyzed by protein components of the acetyl-CoA decarbonylase synthase enzyme complex from *Methanosarcina barkeri*. J. Biol. Chem. **271:**8352–8358.
14. **Grahame, D. A., S. Khangulov, and E. DeMoll.** 1996. Reactivity of a paramagnetic enzyme-CO adduct in acetyl-CoA synthesis and cleavage. Biochemistry **35:**593–600.
15. **Gupta, R. S., and B. Singh.** 1994. Phylogenetic analysis of 70 kD heat shock protein sequences suggests a chimeric origin for the eukaryotic cell nucleus. Curr. Biol. **4:**1104–1114.
16. **Hu, Z. G., N. J. Spangler, M. E. Anderson, J. Q. Xia, P. W. Ludden, P. A. Lindahl, and E. Munch.** 1996. Nature of the C-cluster in Ni-containing carbon monoxide dehydrogenases. J. Am. Chem. Soc. **118:**830–845.
17. **Jablonski, P. E., W.-P. Lu, S. W. Ragsdale, and J. G. Ferry.** 1993. Characterization of the metal centers of the corrinoid/iron-sulfur component of the CO dehydrogenase enzyme complex from *Methanosarcina thermophila* by EPR spectroscopy and spectroelectrochemistry. J. Biol. Chem. **268:**325–329.
18. **Jetten, M. S. M., W. R. Hagen, A. J. Pierik, A. J. M. Stams, and A. J. B. Zehnder.** 1991. Paramagnetic centers and acetyl-coenzyme A/CO exchange activity of carbon monoxide dehydrogenase from *Methanothrix soehngenii*. Eur. J. Biochem. **195:**385–391.
19. **Kerby, R. L., S. S. Hong, S. A. Ensign, L. J. Coppoc, P. W. Ludden, and G. P. Roberts.** 1992. Genetic and physiological characterization of the *Rhodospirillum rubrum* carbon monoxide dehydrogenase system. J. Bacteriol. **174:** 5284–5294.
20. **Kisker, C., H. Schindelin, B. E. Alber, J. G. Ferry, and D. C. Rees.** 1996. A

left-handed β-helix revealed by the crystal structure of a carbonic anhydrase from an archaeon. EMBO J. **15:**2323–2330.
21. **Lu, W. P., P. E. Jablonski, M. Rasche, J. G. Ferry, and S. W. Ragsdale.** 1994. Characterization of the metal centers of the Ni/Fe-S component of the carbon-monoxide dehydrogenase enzyme complex from *Methanosarcina thermophila*. J. Biol. Chem. **269:**9736–9742.
22. **Lu, W. P., I. Schiau, J. R. Cunningham, and S. W. Ragsdale.** 1993. Sequence and expression of the gene encoding the corrinoid/iron-sulfur protein from *Clostridium thermoaceticum* and reconstitution of the recombinant protein to full activity. J. Biol. Chem. **268:**5605–5614.
23. **Maupin-Furlow, J., and J. G. Ferry.** 1995. A proteasome from the methanogenic archaeon *Methanosarcina thermophila*. J. Biol. Chem. **270:**28617–28622.
24. **Maupin-Furlow, J., and J. G. Ferry.** 1996. Characterization of the *cdhD* and *cdhE* genes encoding subunits of the corrinoid/iron-sulfur enzyme of the CO dehydrogenase complex from *Methanosarcina thermophila*. J. Bacteriol. **178:** 340–346.
25. **Morton, T. A., J. A. Runquist, S. W. Ragsdale, T. Shanmugasundaram, H. G. Wood, and L. G. Ljungdahl.** 1991. The primary structure of the subunits of carbon monoxide dehydrogenase/acetyl-CoA synthase from *Clostridium thermoaceticum*. J. Biol. Chem. **266:**23824–23828.
26. **Pitt-Rivers, R., and F. S. A. Impiombato.** 1968. The binding of sodium dodecyl sulphate to various proteins. Biochem. J. **109:**825–830.
27. **Qiu, D., M. Kumar, S. W. Ragsdale, and T. G. Spiro.** 1994. Nature's carbonylation catalyst: raman spectroscopic evidence that carbon monoxide binds to iron, not nickel, in CO dehydrogenase. Science **264:**817–819.
28. **Ragsdale, S. W., J. E. Clark, L. G. Ljungdhal, L. L. Lundie, and H. L. Drake.** 1983. Properties of purified carbon monoxide dehydrogenase from *Clostridium thermoaceticum*, a nickel, iron-sulfur protein. J. Biol. Chem. **258:**2364–2369.
29. **Ragsdale, S. W., M. Kumar, J. Seravalli, D. Qui, and T. G. Spiro.** 1996. Anaerobic carbon monoxide dehydrogenase, p. 191–196. *In* M. Lidstrom and R. Tabita (ed.), Microbial growth on C_1 compounds. Kluwer Academic, Dordrecht, Germany.
30. **Raybuck, S. A., S. E. Ramer, D. R. Abbanat, J. W. Peters, W. H. Orme-Johnson, J. G. Ferry, and C. T. Walsh.** 1991. Demonstration of carbon-carbon bond cleavage of acetyl coenzyme A by using isotopic exchange catalyzed by the CO dehydrogenase complex from acetate-grown *Methanosarcina thermophila*. J. Bacteriol. **173:**929–932.
31. **Realini, C., S. W. Rogers, and M. Rechsteiner.** 1994. KEKE motifs. Proposed roles in protein-protein association and presentation of peptides by MHC class I receptors. FEBS Lett. **348:**109–113.
32. **Rechsteiner, M.** 1991. Natural substrates of the ubiquitin proteolytic pathway. Cell **66:**615–618.
33. **Reeve, J. N.** 1993. Structure and organization of genes, p. 493–527. *In* J. G. Ferry (ed.), Methanogenesis. Ecology, physiology, biochemistry & genetics. Chapman and Hall, New York.
34. **Souillard, N., M. Magot, O. Possot, and L. Sibold.** 1988. Nucleotide sequence of regions homologous to *nifH* (nitrogenase Fe protein) from the nitrogen-fixing archaebacteria *Methanococcus thermolithotrophicus* and *Methanobacterium ivanovii*: evolutionary implications. J. Mol. Evol. **27:**65–76.
35. **Sowers, K. R., T. T. Thai, and R. P. Gunsalus.** 1993. Transcriptional regulation of the carbon monoxide dehydrogenase gene (*cdhA*) in *Methanosarcina thermophila*. J. Biol. Chem. **268:**23172–23178.
36. **Stephens, P. J., M.-C. McKenna, S. A. Ensign, D. Bonam, and P. W. Ludden.** 1989. Identification of a Ni- and Fe-containing cluster in *Rhodospirillum rubrum* carbon monoxide dehydrogenase. J. Biol. Chem. **264:**16347–16350.
37. **Terlesky, K. C., M. J. K. Nelson, and J. G. Ferry.** 1986. Isolation of an enzyme complex with carbon monoxide dehydrogenase activity containing corrinoid and nickel from acetate-grown *Methanosarcina thermophila*. J. Bacteriol. **168:**1053–1058.
38. **Xia, J., and P. A. Lindahl.** 1996. Assembly of an exchange-coupled [Ni:Fe_4S_4] cluster in the α metallosubunit of carbon monoxide dehydrogenase from *Clostridium thermoaceticum* with spectroscopic properties and CO-binding ability mimicking those of the acetyl-CoA synthase active site. J. Am. Chem. Soc. **118:**483–484.
39. **Xia, J., J. F. Sinclair, T. O. Baldwin, and P. A. Lindahl.** 1996. Carbon monoxide dehydrogenase from *Clostridium thermoaceticum*: quaternary structure, stoichiometry of its SDS-induced dissociation, and characterization of the faster-migrating form. Biochemistry **35:**1965–1971.