JOURNAL OF BACTERIOLOGY, Oct. 1994, p. 6127–6130
0021-9193/94/$04.00+0

Vol. 176, No. 19

NOTES

The Reductive Acetyl Coenzyme A Pathway: Sequence and Heterologous Expression of Active Methyltetrahydrofolate:Corrinoid/ Iron-Sulfur Protein Methyltransferase from *Clostridium thermoaceticum*†

DAVID LEE ROBERTS,[1]‡ SHAYING ZHAO,[1] TZANKO DOUKOV,[2] AND STEPHEN W. RAGSDALE[1]*

Department of Biochemistry, University of Nebraska, Lincoln, Nebraska 68583-0718,[1] and Department of Chemistry, University of Nebraska, Lincoln, Nebraska 68583-0304[2]

Received 10 May 1994/Accepted 22 July 1994

The methyltransferase (MeTr) from *Clostridium thermoaceticum* transfers the N^5-methyl group of (*6S*)-methyltetrahydrofolate to the cobalt center of a corrinoid/iron-sulfur protein in the acetyl coenzyme A pathway. MeTr was purified to homogeneity and shown to lack metals. The *acsE* gene encoding MeTr was sequenced and actively expressed in *Escherichia coli* at a level of 9% of cell protein. Regions in the sequences of MeTr and the *E. coli* cobalamin-dependent methionine synthase were found to share significant homology, suggesting that they may represent tetrahydrofolate-binding domains.

The reductive acetyl coenzyme A (acetyl-CoA) pathway (22) is a major mechanism of CO_2 fixation in anaerobic environments. Serving an important role in the global carbon cycle, this pathway has been best studied in *Clostridium thermoaceticum*, an anaerobic, thermophilic, acetogenic bacterium. Sulfate reducers, methanogens, and acetogens use the acetyl-CoA pathway to convert CO or CO_2 to cell carbon (12, 17, 22). In some methanogens, key enzymes operate in reverse to convert acetic acid to methane (10, 11, 27).

In a key step in the pathway, an enzyme called methyltrans-

```
130   GCAGGCGGCT ACAGCGGTCC GGACCGAGCC GTAGTGCTGC GGGCGTGCAA CTTCTGAGCA AAAGGATATA TCCCGCCCGT CCAGAAAATA CCAAATTCTA

-30   AAAGAACGTT ATTTCAAAGG AGGCAATCCC

  1   fM L I  I G E R  I N G  M F G  D I K R  A I Q  E R D  P A P V  Q E W  A R R
  1   ATGCTCATTA TCGGTGAACG GATTAACGGT ATGTTCGGCG ACATCAAACG GGCCATCCAG GAACGGGACC CGGCCCCGGT ACAGGAGTGG GCCCGGCGCC

 34   Q E E G  G A R  A L D  L N V G  P A V  Q D K  V S A M  E W L  V E V  T Q E V
101   AGGAGGAAGG CGGCGCCCGC GCCCTGGACC TGAACGTGGG CCCGGCCGTC CAGGACAAGG TCAGCGCCAT GGAGTGGCTG GTCGAGGTCA CCCAGGAGGT

 68     S N L  T L C  L D S T  N I K  A I E  A G L K  K C K  N R A  M I N S  T N A
201   CAGCAACCTG ACCCTGTGTC TGGATTCCAC CAATATCAAA GCCATCGAAG CCGGGCTGAA AAAATGTAAA AACCGGGCCA TGATCAACTC CACCAATGCC

101    E R E  K V E K  L F P  L A V  E H G A  A L I  G L T  M N K T  G I P  K D T
301   GAGCGAGAAA AGGTAGAAAA GCTCTTCCCC CTGGCCGTCG AACATGGTGC CGCCCTTATA GGCCTGACCA TGAATAAAAC TGGTATTCCC AAGGATAGCG

134   D T R L  A F A  M E L  V A A A  D E F  G L P  M E D L  Y I D  P L I  L P A N
401   ATACCCGCCT GGCCTTTGCC ATGGAGTTGG TAGCCGCTGC CGATGAGTTC GGGCTGCCCA TGGAGGATTT GTACATTGAT CCCCTGATTC TTCCTGCCAA

168    V A Q  D H A  P E V L  K T L  Q Q I  K M L A  D P A  P K T  V L G L  S N V
501   TGTAGCCCAG GACCATGCTC CGGAAGTCCT GAAAACTCTC CAGCAGATCA AGATGCTGGC CGATCCGGCT CCCAAGACAG TCCTGGGTTT AAGTAATGTC

201    S Q N  C Q N R  P L I  N R T  F L A M  A M A  C G L  M S A  I A D A  C D E
601   TCCCAGAACT GCCAGAATCG GCCCTTGATT AACCGGACCT TCCTTGCCAT GGCCATGGCC TGCGGCTTGA TGTCCGCCA TCGCCGACGC TTGCGACGAA

234   A L I  E T A A  T A E  I L L  N Q T V  Y C D  S F V  K M F K  T R *
701   GCCCTCATCG AAACGGCAGC CACGGCAGAG ATCTTGCTCA ACCAAACCGT TTACTGCGAT TCTTTTGTCA AGATGTTTAA AACCAGGTAG GAATTTTTCA

801   TAGTAACGTC GAAATGTAAT TAAAGCATTG TTAAAAGATT ACTACTTCAA AATCTCCAGC GCCAGCAGCA ATACGTAAAT GCCCCCAGGG AATACCGTGA

901   CGGCAGGTGC CAGCGGATTC GT
```

FIG. 1. DNA and amino acid sequences of MeTr. The asterisk indicates the termination codon of the peptide.

* Corresponding author. Phone: (402) 472-2943. Fax: (402) 472-7842. Electronic mail address: sragsdale@crcvms.unl.edu.

† This article has been assigned journal series no. 10720, Agricultural Research Division, University of Nebraska.

‡ Present address: Department of Biochemistry, Medical College of Wisconsin, Milwaukee, WI 53226.

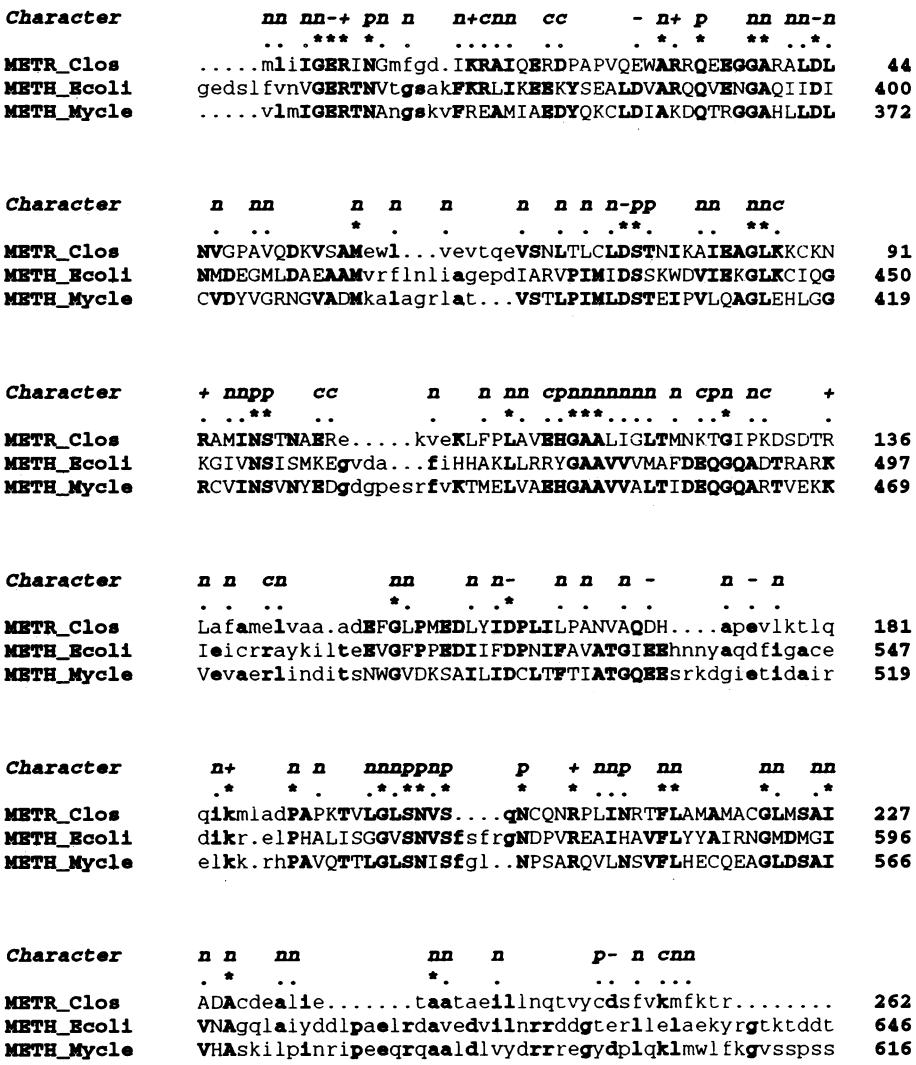

FIG. 2. Sequence comparisons between the MeTr of *C. thermoaceticum*, MetH of *E. coli* (amino acids 350 to 650), and methionine synthase from *M. leprae*. The accession numbers are L34780, J04975, and U00017, respectively. Alignment was performed with the PILEUP subprogram of the GCG program. The lowercase letters indicate amino acids that are not in well-conserved regions; capital letters indicate amino acids in well-conserved regions. Boldface type indicates identity in two of three sequences, asterisks indicate identity in all three, and dots indicate conserved residues. The top letters indicate the character of the conserved residue: *n*, hydrophobic; *p*, polar; *c*, charged; − or +, the conserved charge.

ferase (MeTr) catalyzes transfer of the N^5-methyl group of (6*S*)-methyltetrahydrofolate (CH_3-H_4folate) to the cobalt center of a corrinoid/iron-sulfur protein (C/Fe-SP). Both MeTr and the C/Fe-SP have been purified to homogeneity (7, 23). The MeTr reaction is similar to the first step of the reaction catalyzed by methionine synthase (5), a 139-kDa protein that, after activation by *S*-adenosyl-L-methionine, transfers the methyl group of CH_3-H_4folate to bound cobalamin, forming methyl-Co^{3+}. Subsequently, the methyl group is transferred to homocysteine to form methionine.

The *acsE* gene encoding MeTr has previously been cloned from *C. thermoaceticum* DSM521 into plasmid pCtJ9A and expressed at low levels in *Escherichia coli* (24). The *acsE* gene is part of a cluster containing the genes encoding CO dehydrogenase (*acsA* and *acsB*) (21) and the C/Fe-SP (*acsC* and *acsD*) (18). For DNA sequencing and increasing the level of heterologous expression, the *acsE* gene was subcloned by digesting pCtJ9A (24) with *Pst*I and *Eco*RI and ligating the fragment into Bluescript vectors SK^- and KS^- digested with the same enzymes. Plasmid DNA was transformed into *E. coli* (13), and colonies were selected and screened by growth on plates containing ampicillin, 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) and isopropyl-β-D-thiogalactopyranoside (IPTG) and by colony hybridization with ^{32}P-labeled oligonucleotide probes specific to MeTr (24). Plasmid DNA was prepared on a large scale by alkali lysis (6) followed by precipitation with polyethylene glycol (20). To sequence the *acsE* gene in the Bluescript construct, six deletion clones were generated from the 5′ end after digestion with *Kpn*I and *Sal*I and six were generated from the 3′ end after digestion with *Sac*I and *Xba*I. The DNA was then digested with exonuclease III (14) and treated with S1 nuclease by using the Erase-A-Base kit (Promega). The DNA was electrophoresed, excised from a gel, ligated, and transformed (13) into *E. coli* JM109.

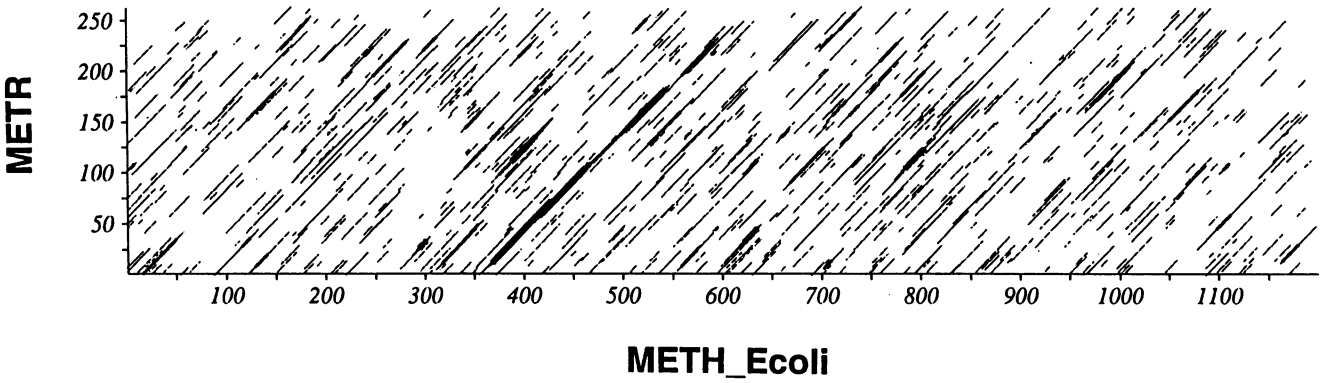

FIG. 3. Search matrix of the *E. coli* methionine synthase (abscissa) with MeTr (ordinate). Alignment was performed by using the Argos program (3) with a constant search window length of 25 residues and a cutoff value of 35. Since the length of the hashed marking is a measure of homology, the hashed diagonal line indicates the region of significant homology.

Plasmids were selected for size by using the "quick-screen" method (Strategene manual). Nucleotide sequences were determined by the dideoxynucleotide method (25) for double-stranded DNA with Sequenase version 2.0 (U.S. Biochemical Corp., Cleveland, Ohio) and [α-^{32}P]dATP (ICN Biochemicals).

Both DNA strands were sequenced from 129 bases upstream to 100 bases downstream of the *acsE* open reading frame (Fig. 1). The codon probability of the three possible reading frames using the established codon usage of *C. thermoaceticum* (21) and the program of Staden and McLachlan (26) indicated that the determined open reading frame for *acsE* was greater than 99% probable (results not shown). These results indicate that no frameshift errors occurred during sequence analysis. The deduced subunit molecular weight, 28,640, was identical to that determined by mass-spectroscopic analysis (below), and the sequence of the first 27 amino acids of MeTr originating with *N*-formylmethionine and ending with PAPV had been determined earlier and was identical to that area in the deduced sequence (24). The deduced and experimentally determined amino acid compositions were in agreement. These combined results indicate that the sequence shown in Fig. 1 is correct.

Homology searches of the Brookhaven, GenBank, and EMBL data banks by using BLAST (1) revealed significant homology with the methionine synthases from *E. coli* and *Mycobacterium leprae*. Results of analysis by the Genetics Computer Group (GCG) program BESTFIT, shown in Fig. 2, revealed 50% similarity and 25% identity with the *E. coli* methionine synthase region between amino acids 350 and 600. With the *M. leprae* protein, the identity reaches 31%. The results of the Argos program (3), shown in Fig. 3, reveal the striking similarity between the MeTr protein sequence from residues 1 to 200 and the *E. coli* methionine synthase from residues 350 to 650. The hatched line depicts the region of highest homology. It is known that the H_4folate-binding domain of methionine synthase is in a 98.4-kDa amino-terminal tryptic fragment that includes residues 1 through 896 (8). Since the region from residues 650 to 850 includes the cobalamin-binding domain (4, 19), our results suggest that the regions from 350 to 650 of methionine synthase and 1 to 200 of MeTr may include their H_4folate-binding domains.

In the Bluescript construct, MeTr can be expressed at high levels in *E. coli*. The activity of the enzyme was measured at 525 nm by monitoring the demethylation of methyl-B_{12} by H_4folate to form B_{12} and CH_3-H_4folate ($\Delta\varepsilon_{525}$ = 8.6 mM^{-1} cm^{-1}). This assay is similar to the radioactive measurement described previously (15) and contained 66 μM methyl-B_{12}, 300 μM H_4folate, and 5 mM dithiothreitol in a total volume of 0.8 ml. It is notable that MeTr is not stimulated by adenosylmethionine. The specific activity of MeTr in *E. coli* cell extracts was 13 μmol min^{-1} mg^{-1} (U mg^{-1}) at 55°C, approximately fivefold higher than in extracts from *C. thermoaceticum* (3 U mg^{-1}). Western immunoblot analysis of cell extracts with purified MeTr from *C. thermoaceticum* as a standard demonstrated that MeTr represents 8 to 10% of soluble cell protein in *E. coli* and ~2% in *C. thermoaceticum*. The high level of heterologous expression may result from the upstream DNA sequences being similar to consensus transcriptional and translational control sequences in *E. coli*. A putative ribosome-binding site (AGGAGGC) with strong homology to the *E. coli* consensus sequence was located with a standard spacing of 7 bases before the ATG start codon. Upstream of the coding region for *acsE* is a sequence, TCCAGA-X_{22}-TTATTT, that is similar to that of the *E. coli* consensus transcriptional control region, TTGAGA-X_{21}-TATAAT.

MeTr was isolated from *E. coli* and *C. thermoaceticum*, and the native and heterologously expressed proteins were compared. *C. thermoaceticum* was grown as described previously (2), and MeTr was purified 51-fold under anaerobic conditions in a Vacuum Atmospheres chamber at 16°C to apparent homogeneity by slightly modifying a previously reported procedure (7). The same procedure was used to purify the enzyme from *E. coli* except that, after the disrupted cells were centrifuged, the supernatant was incubated at 70°C for 10 min and centrifuged at 10,000 × *g* before being applied to the DEAE column. MeTr was routinely measured by Western hybridization (24), and the concentration of protein was determined by the Rose Bengal method (9). The specific activity of purified MeTr from both *E. coli* and *C. thermoaceticum* was between 150 and 200 U mg^{-1} at 55°C. Activity measurements thus agree with immunological analysis that active MeTr represents 8 to 10% of soluble cell protein in *E. coli* and ~2% of soluble cell protein in *C. thermoaceticum*. MeTr was found by plasma emission spectroscopy (16) to contain no metals at levels above 0.1 g-atom/mol of protein. It was also shown to lack chromophoric prosthetic groups since the absorption spectrum arises from the aromatic amino acids; the ε_{280} of 13,400 for the monomer is in agreement with the calculated value (i.e., 14,000). The subunit molecular weights determined by matrix-assisted laser desorption mass spectroscopy for the MeTr

expressed in *C. thermoaceticum* (28,640 ± 35; $n = 4$) and in *E. coli* (28,550 ± 110; $n = 4$) were identical within error limits to that deduced from the DNA sequence (i.e., 28,641) and determined by sodium dodecyl sulfate electrophoresis (i.e., 27,000). In addition, the native protein is a dimer whose molecular weight estimated by gel filtration is 58,900 (7). The combined results indicate that the native protein and heterologously expressed protein are identical.

Nucleotide sequence accession number. The nucleotide sequence presented here has been assigned accession no. L34780 by GenBank.

Metal content was determined by Rebecca Risley, Plasma Emission Laboratory, University of Georgia. We thank Fred Wagner, UN-L, and BioNebraska for performing mass-spectroscopic analysis of MeTr.

We thank John Stezowski, UN-L, for partial support of T.D. This work was supported by NIH grant GM39451 to S.W.R. and an EPSCoR Metallobiochemistry Grant.

REFERENCES

1. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215**:403–410.
2. **Andreesen, J. R., A. Schaupp, C. Neurater, A. Brown, and L. G. Ljungdahl.** 1973. Fermentation of glucose, fructose, and xylose by *Clostridium thermoaceticum*: effect of metals on growth yield, enzymes, and the synthesis of acetate from CO_2. J. Bacteriol. **114**:743–751.
3. **Argos, P.** 1987. A sensitive procedure to compare amino acid sequences. J. Mol. Biol. **193**:385–396.
4. **Banerjee, R. V., N. L. Johnston, J. K. Sobeski, P. Datta, and R. G. Matthews.** 1989. Cloning and sequence analysis of the *Escherichia coli metH* gene encoding cobalamin-dependent methionine synthase and isolation of a tryptic fragment containing the cobalamin-binding domain. J. Biol. Chem. **264**:13888–13895.
5. **Banerjee, R. V., and R. G. Matthews.** 1990. Cobalamin-dependent methionine synthase. FASEB J. **4**:1450–1459.
6. **Berger, S. L., and A. R. Kimmel.** 1987. Guide to molecular cloning techniques. Academic Press, Inc., San Diego, Calif.
7. **Drake, H. L., S.-I. Hu, and H. G. Wood.** 1981. Purification of five components from *Clostridium thermoaceticum* which catalyze synthesis of acetate from pyruvate and methyltetrahydrofolate. Properties of phosphotransacetylase. J. Biol. Chem. **256**:11137–11144.
8. **Drummond, J. T., S. Huang, R. M. Blumenthal, and R. G. Matthews.** 1993. Assignment of enzymatic function to specific protein regions of cobalamin-dependent methionine synthase from *Escherichia coli*. Biochemistry **32**:9290–9295.
9. **Elliott, J. I., and L. G. Ljungdahl.** 1982. Isolation and characterization of an Fe_8-S_8 ferredoxin (ferredoxin II) from *Clostridium thermoaceticum*. J. Bacteriol. **151**:328–333.
10. **Ferry, J. G.** 1992. Biochemistry of methanogenesis. Crit. Rev. Biochem. Mol. Biol. **27**:473–502.
11. **Ferry, J. G.** 1992. Methane from acetate. J. Bacteriol. **174**:5489–5495.
12. **Fuchs, G.** 1986. CO_2 fixation in acetogenic bacteria: variations on a theme. FEMS Microbiol. Rev. **39**:181–213.
13. **Hanahan, D.** 1983. Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. **166**:557–580.
14. **Henikoff, S.** 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene **28**:351–359.
15. **Hu, S.-I., E. Pezacka, and H. G. Wood.** 1984. Acetate synthesis from carbon monoxide by *Clostridium thermoaceticum*. Purification of the corrinoid protein. J. Biol. Chem. **259**:8892–8897.
16. **Jones, J. B., Jr.** 1977. Elemental analysis of soil extracts and plant tissue by plasma emission spectroscopy. Commun. Soil Sci. Plant Anal. **8**:349–365.
17. **Ljungdahl, L. G.** 1986. The autotrophic pathway of acetate synthesis in acetogenic bacteria. Annu. Rev. Microbiol. **40**:415–450.
18. **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.
19. **Luchinsky, C. L., J. T. Drummond, R. G. Matthews, and M. L. Ludwig.** 1992. Crystallization and preliminary X-ray diffraction studies of the cobalamin binding domain of methionine synthase from Escherichia coli. J. Mol. Biol. **222**:557–560.
20. **Miller, H.** 1987. Practical aspects of preparing phage and plasmid DNA: growth, maintenance, and storage of bacteria and bacteriophage. Methods Enzymol. **152**:145–170.
21. **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 CO dehydrogenase/acetyl-CoA synthase from *Clostridium thermoaceticum*. J. Biol. Chem. **266**:23824–23838.
22. **Ragsdale, S. W.** 1991. Enzymology of the acetyl-CoA pathway of autotrophic CO_2 fixation. Crit. Rev. Biochem. Mol. Biol. **26**:261–300.
23. **Ragsdale, S. W., P. A. Lindahl, and E. Münck.** 1987. Mössbauer, EPR, and optical studies of the corrinoid/Fe-S protein involved in the synthesis of acetyl-CoA by *Clostridium thermoaceticum*. J. Biol. Chem. **262**:14289–14297.
24. **Roberts, D. L., J. E. James-Hagstrom, D. K. Smith, C. M. Gorst, J. A. Runquist, J. R. Baur, F. C. Haase, and S. W. Ragsdale.** 1989. Cloning and expression of the gene cluster encoding key proteins involved in acetyl-CoA synthesis in *Clostridium thermoaceticum*: CO dehydrogenase, the corrinoid/Fe-S protein, and methyltransferase. Proc. Natl. Acad. Sci. USA **86**:32–36.
25. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
26. **Staden, R., and A. D. McLachlan.** 1982. Codon preference and its use in identifying protein coding regions in long DNA sequences. Nucleic Acids Res. **10**:141–156.
27. **Thauer, R. K., D. Möller-Zinkhan, and A. M. Spormann.** 1989. Biochemistry of acetate catabolism in anaerobic chemotrophic bacteria. Annu. Rev. Microbiol. **43**:43–67.