JOURNAL OF BACTERIOLOGY, Aug. 1992, p. 4997–5004
0021-9193/92/154997-08$02.00/0

Vol. 174, No. 15

Identification of Formate Dehydrogenase-Specific mRNA Species and Nucleotide Sequence of the *fdhC* Gene of *Methanobacterium formicicum*

W. BRUCE WHITE† AND JAMES G. FERRY*

Department of Anaerobic Microbiology, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061-0315

Received 14 January 1992/Accepted 19 May 1992

The overlapping *fdhA* and *fdhB* genes of *Methanobacterium formicicum*, which encode the α and β subunits, respectively, of formate dehydrogenase, were cotranscribed as part of an approximately 4.5-kb transcript. An additional gene (*fdhC*) upstream of *fdhA* was cotranscribed with *fdhA* and *fdhB*. The deduced amino acid sequence suggested that *fdhC* has the potential to encode a hydrophobic polypeptide with a calculated molecular weight of 29,417. A hydropathy plot of the hypothetical polypeptide indicated several potential membrane-spanning regions. The putative *fdhC* gene product had 28% identity with the deduced amino acid sequence of the *nirC* gene from *Salmonella typhimurium*. Northern (RNA) blot analyses and primer extension assays located a transcription start site 268 bp upstream of the initiation codon of *fdhC*. A sequence identical to the consensus promoter sequence for methanogenic organisms was situated between −35 and −25 bp from the proposed transcription start site. In addition to the 4.5-kb transcript, Northern blot analyses detected a 1.1-kb transcript with an *fdhC*-specific probe and a 3.4-kb transcript with either an *fdhA*- or *fdhB*-specific probe. The levels of all three transcripts were significantly greater in cells grown in media supplemented with molybdate.

Methanobacterium formicicum is an archaeon (32) which utilizes either formate or carbon dioxide and hydrogen as the sole sources of carbon and energy. Formate dehydrogenase catalyzes the oxidation of formate, which supplies electrons for the reduction of carbon dioxide to methane. The formate dehydrogenase of *M. formicicum* is an iron-sulfur enzyme containing molybdopterin guanine dinucleotide and flavin adenine dinucleotide (13, 18, 25, 26). The enzyme contains two subunits with molecular weights of approximately 85,000 and 35,000 (25). The level of formate dehydrogenase protein in the cell is regulated in response to the amount of molybdate in the growth medium (18).

The *fdhA* and *fdhB* genes, which encode the two subunits of the formate dehydrogenase from *M. formicicum*, have been cloned and sequenced (27). The deduced amino acid sequence of *fdhA*, encoding the large subunit, has 40% identity with the deduced amino acid sequence of *fdhF*, which encodes the 80,000-Da subunit of formate dehydrogenase-H from *Escherichia coli* (36). The *fdhAB* genes of *M. formicicum* overlap by 1 bp, which suggests that they may be cotranscribed as part of a polycistronic operon. Approximately 3.6 kbp upstream of the initiation codon of *fdhA* is a 60-bp sequence with 61% identity to a sequence upstream of the *fdhF* gene of *E. coli* (4, 20). It was previously suggested that the *fdhA* and *fdhB* genes are cotranscribed as part of an approximately 12-kb transcript initiating within 50 bp downstream of this 60-bp sequence (20). Here, we reexamine the transcription of *fdhAB* and present evidence that these genes are cotranscribed as part of a 4.5-kb transcript. In addition, we report the nucleotide sequence of *fdhC*, which was cotranscribed with *fdhA* and *fdhB*.

* Corresponding author.
† Present address: National Animal Disease Center, USDA/ARS, Ames, IA 50010.

MATERIALS AND METHODS

Bacterial strains, bacteriophages, plasmids, and culture conditions. *M. formicicum* JF-1 was grown in 10-liter fermentors in medium with or without molybdate (0.1 mM) as described previously (18). *E. coli* P2392 (Stratagene, La Jolla, Calif.), JM109 (35), and DH5α (Bethesda Research Laboratories, Gaithersburg, Md.) were used as host strains for bacteriophage lambda DASH (Stratagene), bacteriophage M13mp18 (35), and plasmid pUC18-derived recombinant plasmids, respectively. Plasmid pUCfd18 (a gift from P. Schendel, Genetics Institute, Boston, Mass.) is a pUC18-derived recombinant plasmid which contains a 10.2-kbp fragment of *M. formicicum* DNA previously shown to contain the *fdhA* and *fdhB* genes (27). *E. coli* P2392 was grown on NZYM broth (22). *E. coli* DH5α was grown on LB broth (22) or LB broth supplemented with 100 μg of ampicillin per ml. *E. coli* JM109 was grown on YT broth (35).

DNA isolation, recombinant DNA methods, and labeling of oligonucleotides and restriction fragments. High-molecular-weight DNA was isolated from *M. formicicum* by the method of Marmur (17). For construction of a lambda DASH library, *Sau*3A partial digests of *M. formicicum* DNA were size-fractionated on sucrose gradients (22). Pooled fragments of approximately 10 to 22 kbp in size were ligated into lambda DASH arms which had been cut with *Bam*HI and dephosphorylated. Packaged phage were plated and screened for regions of DNA flanking the *fdhA* and *fdhB* genes of *M. formicicum* with a previously described oligonucleotide probe (20) specific for the upstream region of the formate dehydrogenase operon (5′-CGTCCTTGAAATT CAAAGAGG-3′). Plaque lifts, hybridizations, and Southern blotting of both recombinant phage DNA and *M. formicicum* DNA were performed by established procedures (3, 30, 33).

Synthetic oligonucleotides were labeled with [γ-^{32}P]ATP (3,000 Ci/mmol or greater) with T4 polynucleotide kinase as described elsewhere (22). Gel-purified restriction fragments were nick-translated with *E. coli* DNA polymerase I to a

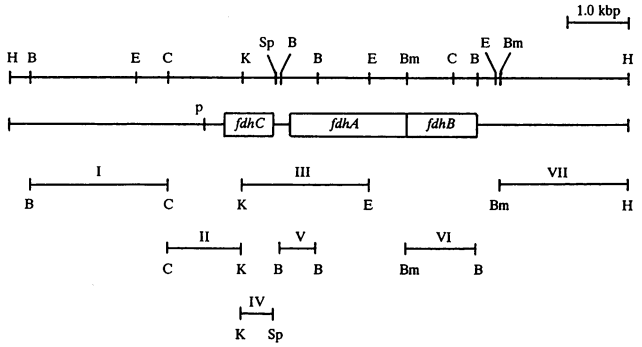

FIG. 1. Physical map of *M. formicicum* DNA containing *fdhC*, *fdhA*, and *fdhB*. At the top is a partial map of restriction enzyme cleavage sites. Abbreviations: B, *Bgl*II; Bm, *Bam*HI; C, *Cla*I; E, *Eco*RI; H, *Hin*dIII; K, *Kpn*I; Sp, *Sph*I. The 0.3-kbp *Hin*dIII-*Bgl*II fragment at the far left is from a lambda DASH clone. The remaining 10.2 kbp is from the insert of pUCfd18 (20). The second line indicates the locations of the *fdhA*, *fdhB*, and *fdhC* genes and a putative *fdh* promoter region (p). The locations of fragments used as probes I to VII in Northern blot analyses are indicated.

specific activity of at least 10^8 cpm/μg with a nick translation reagent kit (Bethesda Research Laboratories). Unincorporated label was removed with Nensorb 20 cartridges (NEN Research Products, Boston, Mass.) according to the manufacturer's specifications.

Plasmid DNA was isolated by a modification (22) of the method of Birnboim and Doly and further purified by CsCl gradient centrifugation. Lambda DNA was isolated on a large scale (29) and a miniprep scale (9) by established procedures. Single-stranded M13 DNA was prepared for sequencing by the procedure described by Bethesda Research Laboratories. Restriction fragments from pUCfd18 and lambda DASH clones containing the *fdhA* and *fdhB* genes and flanking regions were purified from agarose gels with GeneClean (Bio 101, La Jolla, Calif.) for subsequent subcloning or nick translations.

In preparation for sequencing, the 4.2-kbp *Bgl*II fragment directly upstream of the *fdhA* gene (Fig. 1) was subcloned from pUCfd18 into M13mp18 in each orientation by established procedures (22). The 4.2-kbp *Bgl*II fragment, the downstream-flanking 0.7-kbp *Bgl*II fragment which contained a portion of the *fdhA* gene, and the upstream-flanking 1.0-kbp *Bgl*II fragment were subcloned from several overlapping lambda DASH clones. Insert DNA from the lambda DASH clones was mapped and partially sequenced in order to verify the integrity of the pUCfd18 insert.

DNA sequencing. DNA sequences were obtained by the dideoxy chain termination method (23) with the Sequenase procedure of the United States Biochemical Corporation (Cleveland, Ohio), with single-stranded M13 DNA as templates. Both strands of the 4.2-kbp *Bgl*II fragment were completely sequenced. Appropriate subclones were generated for sequencing by the unidirectional progressive deletion method with exonucleases III and VII (35). Recessed 3′ ends proximal to the vector DNA generated by restriction digests were protected from subsequent exonuclease III digestion by being filled in with α-phosphorothioate deoxynucleotides. Commercially available universal sequencing primers (17-mers) and other synthetic oligonucleotides were used in the sequencing procedures. [^{35}S]dATP (1,000 to 1,500 Ci/mmol; NEN Research Products) was used for all sequencing reactions.

RNA isolation. Total nucleic acids for use in Northern (RNA) blotting experiments were isolated as follows. Samples of frozen cell paste (0.1 g) were resuspended in 0.5 ml of lysis buffer (29 mM EDTA and 0.1% sodium dodecyl sulfate in 50 mM Tris [pH 8.0]) in 1.5-ml Eppendorf tubes (17). An equal volume of Tris-saturated phenol (pH 7.0) containing 8-hydroxyquinoline (2.8 mM final concentration) was added to each cell suspension and mixed at 23°C with a 1,000-μl pipettor. The suspension was then centrifuged at 12,000 × *g* for 5 min. The aqueous phase was extracted again with Tris-saturated phenol. Sodium acetate was then added to a final concentration of 0.25 M, and the nucleic acids were precipitated overnight at −20°C with 2.5 volumes of ethanol. The total nucleic acid concentration was determined by the A_{260} after resuspension in distilled water.

Where indicated, Northern blotting was performed with total nucleic acids that were treated with DNase I as follows. The total nucleic acid pellets were washed with 70% ethanol and gently resuspended in 100 μl of 50 mM Tris (pH 8.0) which contained 2.5 mM $MgCl_2$. The samples were then incubated at 37°C for 30 min after the addition of 5 U of DNase I and 40 U of RNasin (Promega, Madison, Wis.). The samples were extracted once with phenol and once with chloroform and then ethanol precipitated at −20°C overnight. The concentration of DNase-treated total nucleic acids was determined by monitoring the A_{260} after resuspension in distilled water.

Total RNA for primer extension assays was prepared as follows. Total nucleic acids were first extracted from frozen cell paste as described above except that the lysis buffer contained 1% Sarkosyl (Fisher Scientific, Fair Lawn, N.H.) instead of 1% sodium dodecyl sulfate. Purified RNA was then isolated from the aqueous phase by pelleting through a 5.7 M CsCl cushion with an SW41 rotor (22). Purified RNA was washed several times with 70% ethanol and stored at −20°C in 70% ethanol until used. The RNA concentration was determined by A_{260} measurements following gentle resuspension in distilled water.

Northern analyses. Northern blotting was performed by established procedures (22). In brief, 10 to 20 μg of glyoxylated total cellular RNA was electrophoresed in a 1% agarose gel. The RNA was electroblotted from the gel onto a GeneScreen hybridization membrane (NEN Research Products) for 6 h at 1.0 A and then for 6 h at 0.25 A with a model TE50 Transphor apparatus (Hoefer Scientific Instruments, San Francisco, Calif.). The membranes were dried and vacuum baked at 80°C. Prehybridization and hybridization were performed as recommended by the manufacturers except that the hybridizations were performed at 48°C for 18 to 24 h. After hybridization, the membranes were washed three times at 23°C for 20 min each in 2× SSC (2× SSC is 0.3 M sodium chloride plus 0.03 M sodium citrate) containing 1% sodium dodecyl sulfate and then twice at 65°C for 30 min each in the same solution and finally rinsed in 0.2× SSC. The SB5 X-ray film (Kodak, Rochester, N.Y.) was exposed to dried membranes for 24 to 48 h at −80°C.

Primer extension assays. Primer extension assays were performed with oligonucleotides (20-mers) complementary to the sequences from positions 466 to 485 (5′-CCTTCTAA TCGACTATCTGT-3′) and 628 to 647 (5′-ACTTCAGCTAG AAGTCCTCC-3′) specific to *fdhC* and the adjacent upstream region, respectively (Fig. 2). Primer extension assays were also performed with an oligonucleotide (23-mer) complementary to the sequences from positions 1601 to 1624 (5′-CCACATCCAACACCGCAGTAGGG-3′) specific to *fdhA* (Fig. 2). Assays were conducted as described before

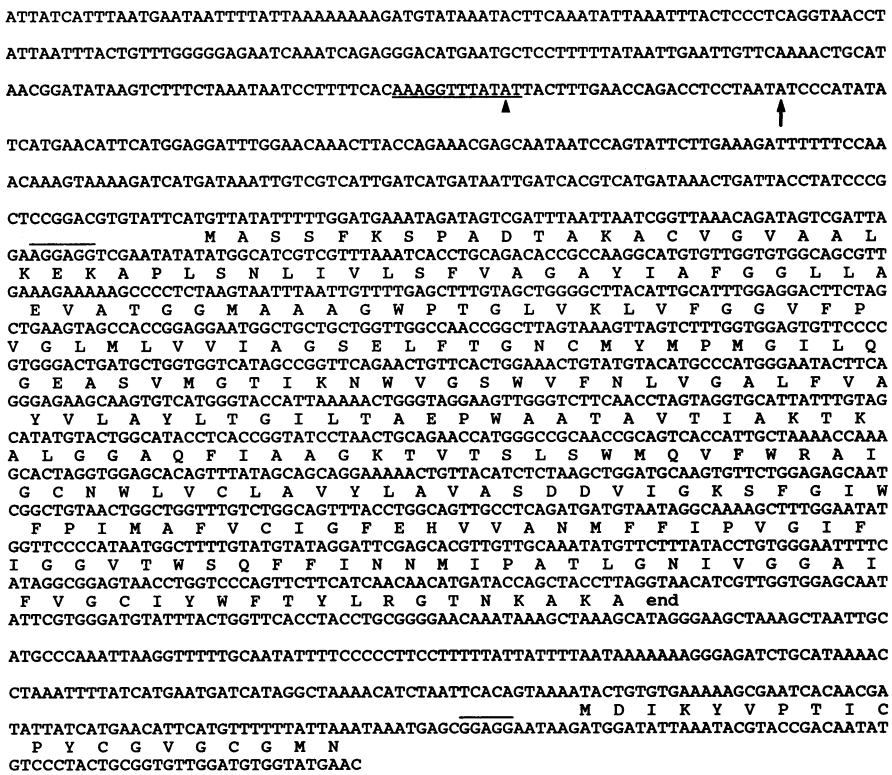

FIG. 2. Nucleotide sequence and flanking sequence of the *fdhC* gene from *M. formicicum*. The nucleotide sequence of one strand is presented in the 5′-to-3′ direction. The deduced amino acid sequence (in single-letter symbols) is shown above the first nucleotide of each codon. In addition, the first 20 codons of *fdhA* (27) are shown below the N-terminal amino acid sequence of the α subunit of formate dehydrogenase. Putative ribosome-binding sites are overlined. A proposed transcriptional start site for the formate dehydrogenase operon is indicated by the arrow pointing to nucleotide A at position 231. The putative −25 promoter region is underlined. The base at −25 relative to the proposed transcriptional start site is marked with an arrowhead.

(1), except that [α-^{35}S]dATP (1,000 to 1,500 Ci/mmol) was used instead of [α-^{32}P]dATP. Also, the assay buffer was modified to contain 50 mM NaCl, 6 mM $MgCl_2$, and 5 mM dithiothreitol in 34 mM Tris (pH 8.3) (8); the RNA appeared to be more soluble in this buffer. Avian myeloblastosis virus reverse transcriptase (NEN Research Products) was used. Sequence ladders, generated by using the same primer on the appropriate M13 DNA clone, were coelectrophoresed to determine the position of the 5′ end of each mRNA species.

Analysis of sequence data. Analysis of nucleic acid and protein sequence data was performed with the IBI/Pustell DNA analysis program (International Biotechnologies, New Haven, Conn.) and the Genetics Computer Group program (University of Wisconsin Biotechnology Center, Madison, Wis.). Computer searches of the GenPept protein data base were done with the FASTA amino acid alignment program (21).

Nucleotide sequence accession number. The GenBank accession number for the 4.2-kbp *Bgl*II fragment which contains the *fdhC* gene is M64798.

RESULTS

Analysis of DNA upstream of the *fdhAB* genes. Figure 1 gives a physical map of *M. formicicum* DNA which contained the *fdhAB* genes and flanking regions. The entire 4.2-kbp *Bgl*II fragment upstream of the *fdhA* gene, which included a previously identified open reading frame (ORF) (28), was sequenced. The ORF was designated the *fdhC* gene (Fig. 1) for reasons described below. Figure 2 shows the sequence of *fdhC* and additional flanking DNA. The putative translation initiation and stop codons were bounded by nucleotides 499 and 1341, respectively. A putative methanogen ribosome-binding site (AGGAGG) was located upstream of the initiation codon at positions 483 to 488. The translational stop for *fdhC* was located 235 nucleotides upstream of the initiation codon of the *fdhA* gene. The intergenic region between *fdhC* and *fdhA* contained no potential hairpin loops preceding oligo(dT) stretches, which are previously determined transcription termination sites for polypeptide-encoding genes from several methanogenic organisms (5).

The *fdhC* gene could encode a hypothetical polypeptide containing 280 amino acids with a calculated molecular weight of 29,417. An analysis of the deduced amino acids indicated that 59.3% were hydrophobic, 32.1% were semipolar, 3.2% were acidic, and 5.4% were basic, which suggested a very hydrophobic putative polypeptide. A hydropathy analysis identified seven hydrophobic regions of 20 or more amino acids, with an average hydropathy index of greater than 1.3 (Fig. 3), that could span the membrane. A computer search of the GenBank data base for matches to the hypothetical polypeptide encoded by *fdhC* revealed 28% identity (Fig. 4) with the amino acid sequence deduced from the *nirC* gene from *Salmonella typhimurium* (34). The *nirC* gene is required for nitrite reduction, but the function of the hypothetical polypeptide is unknown. The *nirC* gene has the potential to encode a polypeptide with a calculated molecular weight of 28,545. A hydropathy analysis of the putative

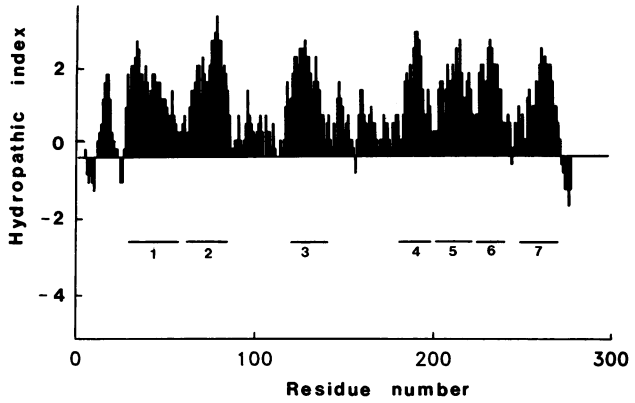

FIG. 3. Hydropathy profile of the deduced amino acid sequence for *fdhC* from *M. formicicum*. The algorithm of Kyte and Doolittle (15) was used, with a span of 9 amino acids. The bars indicate seven hydrophobic regions of 20 or more residues that could span the cytoplasmic membrane.

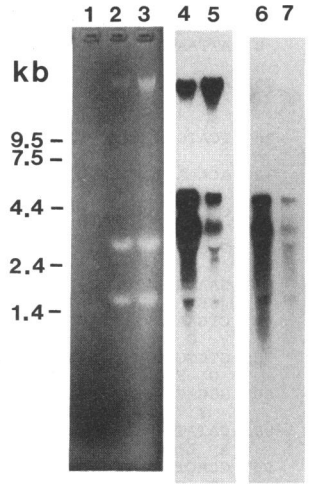

FIG. 5. Estimation of the sizes of formate dehydrogenase-specific mRNA species and effect of molybdate supplementation on the level of mRNA from *M. formicicum*. RNA size markers (Bethesda Research Laboratories) are shown in lane 1. Glyoxylated nucleic acids were fractionated by 1% agarose electrophoresis and stained with ethidium bromide (lanes 1 to 3). The probe used in the Northern blot analyses (lanes 4 to 7) was the ^{32}P-labeled 624-bp *Bgl*II fragment (probe V, Fig. 1). Total nucleic acids (20 μg) were from cells grown with (lanes 2 and 4) or without (lanes 3 and 5) molybdate. DNase-treated total nucleic acids (10 μg) were from cells grown with (lane 6) or without (lane 7) molybdate.

nirC gene product indicated a very hydrophobic protein with six potential membrane-spanning regions (data not shown). The FASTA alignment program identified a 260-nucleotide overlap between *fdhC* (positions 730 to 990, Fig. 2) and *nirC* with 54% identity (data not shown). The computer search also identified 99 nucleotides at the 3′ end of *fdhC* (positions 1223 to 1321, Fig. 2) that had 66% identity with the 3′ end of the ORF immediately upstream of the *E. coli* structural gene for pyruvate formate-lyase (24); the deduced amino acid sequences of the overlapping N termini were 52% identical (data not shown). Three other stretches of amino acids deduced from the *fdhC* gene (positions 61 to 91, 111 to 127, and 203 to 235 in Fig. 4) had 48, 61, and 56% identity, respectively, with stretches deduced from the *E. coli* ORF (data not shown). The ORF, which is cotranscribed with the pyruvate formate-lyase gene, has a predicted amino acid sequence which is very hydrophobic and contains seven potential membrane-spanning helices; however, the function of the hypothetical protein is unknown (24).

Two additional ORFs were identified within the 4.2-kbp *Bgl*II fragment that were in the same orientation as *fdhC* (data not shown). The initiation codons were located 2,706 and 2,216 bp upstream of the initiation codon of *fdhC*. These overlapping ORFs could encode polypeptides with calculated molecular weights of 20,954 and 16,584, respectively; however, as described below, no mRNA species were detected that were complementary to the region of DNA which contained these ORFs. Also, no potential methanogen ribosome-binding site was identified within 50 bp of the initiation codon of either ORF.

Northern analyses. Figure 5 shows the results of Northern blot analyses of total nucleic acids from cells grown with and without molybdate. Probe V (Fig. 1), used in these experiments, was specific for *fdhA*. Two hybridizing bands were detected which corresponded to *fdhA*-specific mRNA spe-

```
                10        20        30                40        50        60        70
fdhC   MASSFKSPADTAKACVGVAALKEKAPLSNLIVLSF-----VAGAYIAFGGLLAEVATGGMAAAGWPTGLVKLVFG
                .:. . :  .: :...::.:.   :.:      .:::...: .:  . .. .... .:      ::.:
nirC         MFTDSINKC--AAKLRASAPVSANNPLGFWVSSAMAGAYVGLGIILIFTLGNLLDPSVRP-----LVMG
                       10        20        30        40        50                60

                80        90        100       110       120       130       140
fdhC   GVFPVGLMLVVIAGSELFTGNCMYMPMGILQGEASVMGTIKNWVGSWVFNLVGALFVAYVLAYLTGILTAEPWAA
       ..: ..: ::.:::::::::. :....:. .:. :           .:. :::..::  : :         .:..
nirC   ATFGIALTLVIIAGSELFTGHTMFLTLGVKAGTISHGQMWAILPQTWLGNLVGSVFVA--LLY--------SWGG
            70        80        90       100       110       120

         150       160       170       180       190       200       210       220
fdhC   TAVTIAKTKALGGAQFIAAGKTVTSLSWMQVFWRAIGCNWLVCLAVYLAVASDDVIGKSFGIWFPIMAFVCIGFE
       ...  ..:. . ....  .. ..: :    :...  ::::::::...:. ....        : . ...   .
nirC   GSLLPVDTSIVHSVALAKTTAPATVL-----FFKGALCNWLVCLAIWMAIRTEGTAKFLAIWWCLLAFIASGYEH
       130       140       150             160       170       180       190

              230       240       250       260       270       280
fdhC   HVVANMFFIP--VGIFIGGVTWSQFFINNMIPATLGNIVGGAIFVGCIYWFTYLRGTNKAKA
        :..  .:     :   .. : .  . .:.. .:::...:..:.:  ::..  ...  : :
nirC   SVANMTLFALSWFGHHSDAYTLAG-IGHNLLWVTLGNTLSGVVFMGLGYWYATPKSERPAPAKINQPEAAANN
       200       210       220       230       240       250       260
```

FIG. 4. Amino acid alignment of *fdhC* from *M. formicicum* and *nirC* from *S. typhimurium*. Alignment was performed with the FASTA search program (21). Dashes represent alignment gaps. Exact matches are indicated with a colon, and conservative matches are marked with a period.

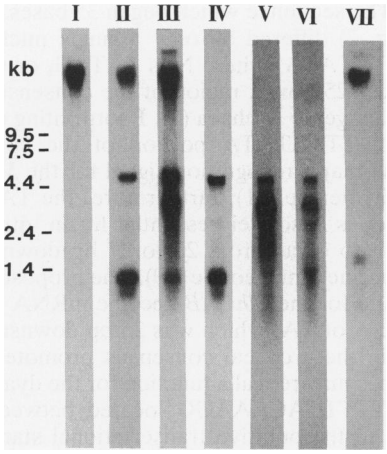

FIG. 6. Cotranscription of *fdhC* with *fdhAB* from *M. formicicum*. The radiolabeled probes (Fig. 1) used in the Northern blot analyses were (I) the 2,274-bp *Bgl*II-*Cla*I fragment upstream of *fdhC*; (II) the 1,278-bp *Cla*I-*Kpn*I fragment, which contains a portion of *fdhC* plus 952 bp of upstream sequence flanking *fdhC*; (III) the 2,029-bp *Kpn*I-*Eco*RI fragment, which contains portions of both *fdhC* and *fdhA* plus the intergenic region; (IV) the 540-bp *Kpn*I-*Sph*I fragment of *fdhC*; (V) the 624-bp *Bgl*II fragment containing a portion of *fdhA* plus 147 bp of upstream sequence flanking *fdhA*; (VI) the 1,196-bp *Bam*HI-*Bgl*II fragment of *fdhB;* and (VII) the 2.3-kbp *Bam*HI-*Hin*dIII fragment downstream from *fdhB*. Total nucleic acids (20 μg) (lanes 1 to 3 and 7) and DNase-treated total nucleic acids (10 μg) (lanes 4 to 6) were from cells grown with molybdate.

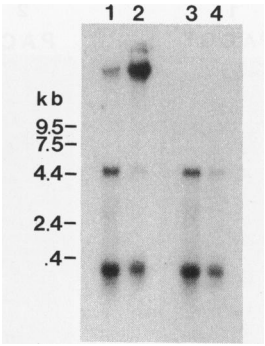

FIG. 7. Effect of molybdate supplementation on the level of *fdhC*-specific mRNA from *M. formicicum*. The probe used in the Northern analyses was the ^{32}P-labeled 540-bp *Kpn*I-*Sph*I fragment (probe IV, Fig. 1) of *fdhC*. Total nucleic acids (20 μg) were from cells grown with (lane 1) or without (lane 2) molybdate. DNase-treated total nucleic acids (10 μg) were from cells grown with (lane 3) or without (lane 4) molybdate.

cies of approximately 4.5 and 3.4 kb. Degradation of the mRNA species was suggested by the hybridization, which tailed below the 3.4-kb band (lanes 4 to 7, Fig. 5); this tailing was interrupted by the 23s and 16s rRNA bands (lanes 2 and 3, Fig. 5). An additional strongly hybridizing band was detected near the top of the gel except when total nucleic acids were treated with DNase I (lanes 6 and 7, Fig. 5), a result which indicated that this band was DNA. When total nucleic acids were isolated from molybdate-supplemented cells, the amount of DNA was always less than when total nucleic acids were isolated from nonsupplemented cells (lanes 2 and 3, Fig. 5). The amount of DNA paralleled the amount of hybridization with *fdh*-specific probes (lanes 4 and 5, Fig. 5). The presence of DNA in the total nucleic acids interfered with the quantitation of RNA; thus, total nucleic acids were treated with DNase I and then quantitated for Northern analyses (lanes 6 and 7, Fig. 5). The results indicated that a significantly greater amount of *fdhA*-specific mRNA was isolated from cells grown with molybdate than from cells grown without molybdate.

Northern blots probed with the *fdhB*-specific probe VI (Fig. 1) gave results similar to those from experiments with the *fdhA*-specific probe V (Fig. 1); both probes hybridized strongly with mRNA species of approximately 4.5 and 3.4 kb (lanes V and VI, Fig. 6). Probe VI also hybridized with an apparently degraded mRNA which migrated ahead of 16s rRNA. The *fdhC*-specific probes II and IV (Fig. 1) strongly hybridized to two mRNA species of approximately 4.5 and 1.1 kb (lanes II and IV, Fig. 6). As was the case for *fdhA*-specific mRNA, the level of *fdhC*-specific mRNA was significantly greater in DNase-treated total nucleic acids isolated from cells grown with molybdate (lanes 3 and 4, Fig. 7). The differences in the levels of the 4.5-kb mRNA appeared to be greater than the differences in the levels of the 1.1-kb mRNA, an unexplained result. Probe III (Fig. 1), which spanned *fdhC* and *fdhA*, hybridized with the 4.5-, 3.5-, and 1.1-kb mRNA species (lane III, Fig. 6). There was no detectable hybridization when the 2.3-kbp *Bgl*II-*Cla*I fragment (probe I, Fig. 1) from upstream of *fdhC* was used as a probe (lane I, Fig. 6). When the 2.3-kbp *Bam*HI-*Hin*dIII fragment (probe VII, Fig. 1) from downstream of *fdhB* was used as a probe, only a faint signal at approximately 1.8 kb was detected (lane VII, Fig. 6). The results suggest cotranscription of the *fdhC, fdhA,* and *fdhB* genes on a 4.5-kb mRNA. The detection of an *fdhC*-specific 1.1-kb mRNA and an *fdhAB*-specific 3.4-kb mRNA was consistent with an additional transcription start site between *fdhC* and *fdhA*, processing of the 4.5-kb mRNA, or both; however, we were unable to distinguish conclusively among these possibilities.

Determination of the 5′ ends of mRNA species by primer extension. The results of Northern blot analyses (Fig. 5, 6, and 7) were consistent with the 5′ ends of the 4.5- and 1.1-kb mRNA species being complementary to any of several hundred nucleotides upstream of *fdhC*. Primer extension assays (Fig. 8) yielded only one major signal which was complementary to nucleotide A at position 231 (Fig. 2). A faint signal (Fig. 8) was also identified which was complementary to nucleotide A at position 260 (Fig. 2). The other primer used in these assays (Materials and Methods) yielded identical results (data not shown). No other mRNAs were detected with 5′ ends complementary to any of 400 bases directly upstream of *fdhC*.

The results from Northern blots (Fig. 5 and 6) were consistent with the 5′ end of the 3.4-kb mRNA species being complementary to bases within the intergenic region between *fdhC* and *fdhA*. Primer extension assays indicated a major signal (Fig. 8), complementary to nucleotide A at position 1443 (Fig. 2), which identified the potential 5′ end of the major mRNA species responsible for the *fdhAB*-specific 3.4-kb band in Fig. 5 and 6. Multiple minor signals were also present both upstream and downstream of this site (Fig. 8); other minor signals further downstream of the major signal were also detected (data not shown). It could not be determined from the results whether these minor signals arose from processing of a larger message or represented additional transcription start sites. However, it is unlikely that any of the signals upstream of the major signal (Fig. 8)

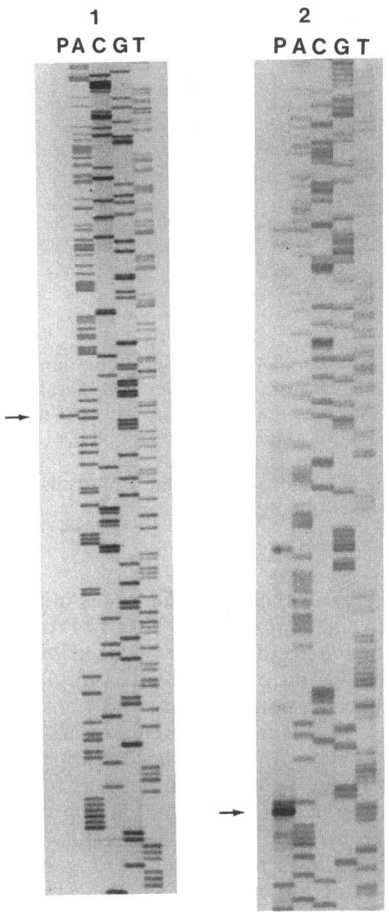

FIG. 8. Determination of the 5′ ends of transcripts by primer extension analysis of RNA from *M. formicicum*. Total RNA was isolated from cells grown with molybdate. Lane P, products of primer extension. Lanes A, C, G, and T, products of deoxynucleotide sequencing reactions with single-stranded M13 DNA as the template and the same primer as for lane P. (Panel 1) Analysis of the 400-bp region directly upstream of *fdhC*. An oligonucleotide complementary to the sequence from positions 466 to 485 in Fig. 2 (5′-CCTTCTAATCGACTATCTGT-3′) was used as the primer. The band at the bottom of lane C corresponds to the base complementary to nucleotide G at position 326 in Fig. 2. The arrow points to the major primer extension signal which was complementary to nucleotide A at position 231 in Fig. 2. An additional faint signal was visible which was complementary to nucleotide A at position 260 in Fig. 2. (Panel 2) Analysis of the 235-bp region between *fdhC* and *fdhA*. The primer is described in Materials and Methods. The band at the bottom of lane C corresponds to the base complementary to the G at nucleotide position 1459 in Fig. 2. The arrow points to the primer extension signal which was complementary to nucleotide A at position 1443 in Fig. 2.

represented the 5′ end of the 4.5-kb mRNA species, since probe II (Fig. 1) hybridized with the 4.5-kb band (lane II, Fig. 6).

DISCUSSION

The results of the present study suggest that the *fdhC* gene of *M. formicicum* was cotranscribed with the *fdhAB* genes on an *fdhCAB*-specific mRNA species of approximately 4.5 kb. We propose a start site for the *fdhCAB*-specific mRNA located 268 bases upstream of the initiation codon of the *fdhC* gene. The sequence which began 35 bases upstream of this site (Fig. 2) differed by only a single nucleotide from AAANNTTTATATA (where N is A, T, G, or C), which is the proposed −25 box A region of the consensus promoter for the methanogenic Archaea (5). Footprinting experiments implicate the TTTATATA portion of the consensus sequence as the major recognition signal for the *Methanococcus* RNA polymerase (31); furthermore, the TATA (TATA box) sequence is absolutely essential for in vitro transcription initiation to occur from 22 to 27 bp downstream at a pyrimidine-purine dinucleotide (10). The proposed transcriptional start site for the *fdhCAB*-specific mRNA (Fig. 2) was at nucleotide A of TA, which was 25 bp downstream of the TATA box in the archaeal consensus promoter sequence. Further studies may reveal a function for the dyad symmetry element CCTTTTCACAAAGG located between −44 and −31 relative to the putative transcriptional start site (positions 187 to 200 in Fig. 2). The *fdhC*-specific 1.1-kb mRNA species and the *fdhAB*-specific 3.4-kb mRNA species appear to account for the entire length of the *fdhCAB*-specific 4.5-kb mRNA. Additional studies are necessary to determine whether the 1.1- and 3.4-kb mRNAs arise from processing of the 4.5-kb mRNA. Although the intergenic region between *fdhC* and *fdhA* did not contain sequences which strictly conformed to the archaeal consensus promoter, the possibility of a transcriptional start site between *fdhC* and *fdhA* cannot be ruled out.

This study indicated that the level of formate dehydrogenase-specific mRNA was significantly higher in cells grown in medium with molybdate. The relative amount of formate dehydrogenase protein increases in molybdate-supplemented cells (18), which is consistent with regulation by molybdenum of either transcription initiation or stability of formate dehydrogenase-specific mRNA. In bacteria, a regulatory role for molybdenum has been suggested at the level of transcription for genes associated with nitrogen fixation (12, 16) and other processes (6, 11, 14, 19).

Previously, it was concluded that the *fdhAB* genes were cotranscribed as part of an approximately 12-kb transcript (20). The conclusion was based partly on Northern blot analyses with *fdhA*- and *fdhB*-specific probes that hybridized only to a high-molecular-weight band (20). It was also previously reported that, when total nucleic acids were isolated from molybdate-supplemented cells, hybridization of the probes to the high-molecular-weight band was less than when total nucleic acids were isolated from cells that were not supplemented (20). These previous results were repeated here; however, the results indicated that the high-molecular-weight band was DNA. Thus, the previous conclusion (20) that the *fdhAB* genes are cotranscribed as part of a large transcript of approximately 12 kb is apparently based on invalid results, which we retract. Likewise, the previous suggestion that molybdenum inversely affects transcription of the *fdhAB* genes is also invalid (18, 20). It is not clear why the 4.5-, 3.4-, and 1.1-kb mRNA species were not detected previously when *fdhA*- or *fdhAB*-specific fragments were used as hybridization probes in Northern blot analyses (20); it is possible that mRNA was degraded during isolation and manipulation of nucleic acids or that the electroblotting transfer times were suboptimal. Previously (20), primer extension assays located the 5′ end of an mRNA species complementary to *M. formicicum* DNA approximately 50 bases downstream from a 60-bp region which has greater than 61% identity with a 60-bp region between −88 and −29 bp upstream of the *fdhF* gene of *E. coli* (4). This site is located approximately 3.7 kbp upstream of the *fdhA* gene

and within the 2.3-kbp *Bgl*II-*Cla*I fragment (Fig. 1). In the present study, we were unable to detect any mRNAs that hybridized with the 2.3-kbp *Bgl*II-*Cla*I fragment. From the results presented here, it is clear that additional studies are necessary to determine whether the 60-bp region has any functional significance relative to transcription of the *fdhCAB* genes.

Transcriptional linkage of *fdhC* to *fdhAB* suggests that the putative *fdhC* gene product is in some manner associated with formate metabolism. The hydropathy plot of the deduced amino acid sequence (Fig. 3) is consistent with a transmembrane carrier protein. Although formate dehydrogenase is found in the soluble fraction after French pressure cell disruption, immunogold labeling experiments indicate that the enzyme is located primarily at the cell membrane (2). It is possible that formate dehydrogenase is associated with the cytoplasmic membrane through interaction with a hydrophobic protein such as the putative *fdhC* gene product. Interestingly, the deduced amino acid sequence of *fdhC* had substantial identity with the hydrophobic membrane-spanning proteins predicted from the *nirC* gene of *S. typhimurium* (34) and the ORF upstream of the *E. coli* pyruvate formate-lyase structural gene (24). Further experiments are necessary to identify and characterize the putative *fdhC* gene product in *M. formicicum* before any conclusions can be drawn regarding its function.

ACKNOWLEDGMENTS

We thank Deborah J. Pheasant for assistance in culturing *M. formicicum*. We are grateful to Shu-Zhen Wang and John L. Johnson for providing us with the DNA sequencing procedures.

This study was supported in part by grant 5086-260-1255 from the Gas Research Institute and grant BIO-89-005 from the Virginia Center for Innovative Technology, Institute of Biotechnology.

REFERENCES

1. **Alam, J., R. A. Whitaker, D. W. Krogmann, and S. E. Curtis.** 1986. Isolation and sequence of the gene for ferredoxin I from the cyanobacterium *Anabaena* sp. strain PCC7120. J. Bacteriol. **168:**1265–1271.
2. **Baron, S. F., D. S. Williams, H. D. May, P. S. Patel, H. C. Aldrich, and J. G. Ferry.** 1989. Immunogold localization of coenzyme F_{420}-reducing formate dehydrogenase in *Methanobacterium formicicum*. Arch. Microbiol. **151:**307–313.
3. **Benton, W. D., and R. W. Davis.** 1977. Screening λgt11 recombinant clones by hybridization to single plaques *in situ*. Science **196:**180–182.
4. **Birkmann, A., F. Zinoni, G. Sawers, and A. Böck.** 1987. Factors affecting transcriptional regulation of the formate-hydrogenlyase pathway of *Escherichia coli*. Arch. Microbiol. **148:**44–51.
5. **Brown, J. W., C. J. Daniels, and J. N. Reeve.** 1989. Gene structure, organization, and expression in archaebacteria. Crit. Rev. Microbiol. **16:**287–338.
6. **Cotter, P. A., and R. P. Gunsalus.** 1989. Oxygen, nitrate, and molybdenum regulation of *dmsABC* gene expression in *Escherichia coli*. J. Bacteriol. **171:**3817–3823.
7. **Frey, G., M. Thomm, B. Brudigam, H. P. Gohl, and W. Hausner.** 1990. An archaebacterial cell-free transcription system: the expression of tRNA genes is mediated by a transcription factor. Nucleic Acids Res. **18:**1361–1368.
8. **Graham, A., J. Steven, D. McKechnie, and J. W. Harris.** 1985. Direct DNA sequencing using avian myeloblastosis virus and Moloney murine leukemia virus reverse transcriptase. Focus **8:**2.
9. **Grossberger, D.** 1987. Minipreps of DNA from bacteriophage lambda. Nucleic Acids Res. **15:**6737.
10. **Hausner, W., G. Frey, and M. Thomm.** 1991. Control regions of an archaeal gene: a TATA box and an initiator element promote cell-free transcription of the $tRNA^{Val}$ gene of *Methanococcus vannielii*. J. Mol. Biol. **222:**495–508.
11. **Iuchi, S., and E. C. C. Lin.** 1987. Molybdenum effector of fumarate reductase repression and nitrate reductase induction in *Escherichia coli*. J. Bacteriol. **169:**3720–3725.
12. **Jacobson, M. R., R. Premakumar, and P. E. Bishop.** 1986. Transcriptional regulation of nitrogen fixation by molybdenum in *Azotobacter vinelandii*. J. Bacteriol. **167:**480–486.
13. **Johnson, J. L., N. R. Bastian, N. L. Schauer, J. G. Ferry, and K. V. Rajagopalan.** 1991. Identification of molybdopterin guanine dinucleotide in formate dehydrogenase from *Methanobacterium formicicum*. FEMS Microbiol. Lett. **77:**213–216.
14. **Kalman, L. V., and R. P. Gunsalus.** 1990. Nitrate- and molybdenum-independent signal transduction mutations in *narX* that alter regulation of anaerobic respiratory genes in *Escherichia coli*. J. Bacteriol. **172:**7049–7056.
15. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157:**105–132.
16. **Luque, F., and R. N. Pau.** 1991. Transcriptional regulation by metals of structural genes for *Azotobacter vinelandii* nitrogenases. Mol. Gen. Genet. **227:**481–487.
17. **Marmur, J.** 1961. A procedure for the isolation of deoxyribonucleic acid from microorganisms. J. Mol. Biol. **3:**208–218.
18. **May, H. D., P. S. Patel, and J. G. Ferry.** 1988. Effect of molybdenum and tungsten on synthesis and composition of formate dehydrogenase in *Methanobacterium formicicum*. J. Bacteriol. **170:**3384–3389.
19. **Miller, J. B., D. J. Scott, and N. K. Amy.** 1987. Molybdenum-sensitive transcriptional regulation of the *chlD* locus of *Escherichia coli*. J. Bacteriol. **169:**1853–1860.
20. **Patel, P. S., and J. G. Ferry.** 1988. Characterization of the upstream region of the formate dehydrogenase operon of *Methanobacterium formicicum*. J. Bacteriol. **170:**3390–3395.
21. **Pearson, W. R., and D. J. Lipman.** 1988. Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA **85:**2444–2448.
22. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
23. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
24. **Sawer, G., and A. Böck.** 1989. Novel transcriptional control of the pyruvate formate-lyase gene: upstream regulatory sequences and multiple promoters regulate anaerobic expression. J. Bacteriol. **171:**2485–2498.
25. **Schauer, N. L., and J. G. Ferry.** 1986. Composition of coenzyme F_{420}-dependent formate dehydrogenase from *Methanobacterium formicicum*. J. Bacteriol. **165:**405–411.
26. **Schauer, N. L., J. G. Ferry, J. F. Honeck, W. H. Orme-Johnson, and C. Walsh.** 1986. Mechanistic studies of the coenzyme F_{420}-dependent formate dehydrogenase from *Methanobacterium formicicum*. Biochemistry **25:**7163–7168.
27. **Shuber, A. P., E. C. Orr, M. A. Recny, P. F. Schendel, H. D. May, N. L. Schauer, and J. G. Ferry.** 1986. Cloning, expression, and nucleotide sequence of the formate dehydrogenase genes from *Methanobacterium formicicum*. J. Biol. Chem. **261:**12942–12947.
28. **Shuber, A. P., E. C. Orr, M. A. Recny, P. F. Schendel, H. D. May, N. L. Schauer, and J. G. Ferry.** Unpublished data.
29. **Silhavy, T. J., M. L. Berman, and L. W. Enquist.** 1984. Experiments with gene fusions. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
30. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98:**503–517.
31. **Thomm, M., G. Wich, J. W. Brown, G. Frey, B. Sherf, and G. S. Beckler.** 1989. An archaebacterial promoter sequence assigned by RNA polymerase binding experiments. Can. J. Microbiol. **35:**30–35.
32. **Woese, C. R., O. Kandler, and M. L. Wheelis.** 1990. Towards a natural system of organisms: proposal for the domains Archaea,

Bacteria, and Eucarya. Proc. Natl. Acad. Sci. USA **87:**4576–4579.

33. **Woods, D.** 1984. Oligonucleotide screening of cDNA libraries. Focus **6:**1–2.
34. **Wu, J.-Y., L. M. Siegel, and N. M. Kredich.** 1991. High-level expression of *Escherichia coli* NADPH-sulfite reductase: requirement for a cloned *cysG* plasmid to overcome limiting siroheme cofactor. J. Bacteriol. **173:**325–333.
35. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33:**103–119.
36. **Zinoni, F., A. Birkmann, T. C. Stadtman, and A. Böck.** 1986. Nucleotide sequence and expression of the selenocysteine-containing polypeptide of formate dehydrogenase (formate hydrogen-lyase linked) from *Escherichia coli*. Proc. Natl. Acad. Sci. USA **83:**4650–4654.