The Formylmethanofuran: Tetrahydromethanopterin Formyltransferase from *Methanobacterium thermoautotrophicum* ΔH

**NUCLEOTIDE SEQUENCE AND FUNCTIONAL EXPRESSION OF THE CLONED GENE**

(Received for publication, June 21, 1989)

Anthony A. DiMarco, Karen A. Smentz, Jordan Konisky, and Ralph S. Wolfe§

*From the Department of Microbiology, University of Illinois, Urbana, Illinois 61801*

The formylmethanofuran: tetrahydromethanopterin formyltransferase (FTR) from *Methanobacterium thermoautotrophicum* ΔH was cloned and its sequence was determined. The clone was contained on a 4.8-kb BamHI fragment of *M. thermoautotrophicum* DNA ligated into pBR329. When this fragment was subcloned into the phagemid pTZ18R, a functional enzyme was synthesized under control of the lac promoter. Sequence analysis revealed the presence of a ribosome binding site and a possible terminator structure. The absence of an identifiable promoter lends credibility to the open reading frame which is present 5’ to *ftr*. The *ftr* gene encodes an acidic protein with a calculated molecular weight of 31,401. The sequence of FTR does not appear to be homologous to any other sequenced proteins, including proteins which use pterin substrates.

*Methanobacterium thermoautotrophicum* ΔH obligately reduces CO2 with hydrogen to yield methane. Despite the existence of gaps in the pathway, several intermediate conversions of methanogenesis have been well studied in this organism, and some of the enzymes have been purified (1-5). Many of these enzymes are unique to methanogens in that they catalyze interconversions among novel coenzymes (6). One of these enzymes, the formylmethanofuran:tetrahydromethanopterin (FTR),1 has been demonstrated as an essential enzyme in the synthesis of methane (1). FTR transfers the formyl group of formylmethanofuran to tetrahydromethanopterin to yield 5-formyltetrahydromethanopterin (1). This enzyme differs remarkably from nonmethanogenic formyl-

* Experiments—Restriction enzymes and DNA modification enzymes were obtained from Bethesda Research Laboratories or Boehringer Mannheim. T7 DNA polymerase (Sequenase™) was from United States Biochemical Corp. (Cleveland, OH). All enzymes were used according to the protocols of the suppliers. Immunological reagents were from Bethesda Research Laboratories. 35S- and 32P-labeled ATP were from Amersham Corp.

**EXPERIMENTAL PROCEDURES**

Reagents—Restriction enzymes and DNA modification enzymes were obtained from Bethesda Research Laboratories or Boehringer Mannheim. T7 DNA polymerase (Sequenase™) was from United States Biochemical Corp. (Cleveland, OH). All enzymes were used according to the protocols of the suppliers. Immunological reagents were from Bethesda Research Laboratories. 35S- and 32P-labeled ATP were from Amersham Corp.

**Protein Sequencing and Preparation of Oligonucleotide for Use as a Probe**—Homogeneous FTR was obtained as a gift from Mark Donnell or purified as described (1). The amino-terminal sequence was determined on a gas-liquid phase protein sequenator from Applied Biosystems at the University of Illinois Genetic Engineering Facility. Oligonucleotides were synthesized on an Applied Biosystems Model 380A DNA synthesizer at the same facility. A synthetic oligonucleotide (5’-CGTGTATCAT-3’) corresponding to the reverse complement of amino acids 1-4 of the protein sequence was used as a
Nucleotide Sequence and Functional Expression of FTR

**RESULTS AND DISCUSSION**

The characterization of genes involved in the biogenesis of methane has been hindered by their indispensability in organisms like *M. thermoautotrophicum* which have no alternative metabolic pathway. In addition, the thermophilic and anaerobic nature of this organism necessitates the use of special culture conditions, thereby aggravating manipulative techniques. To circumvent the standard genetic strategies, we have decided to use an approach which takes advantage of the relatively large amount of biochemical information available, and clone the gene which encodes the formylmethanofuran:terehydromethanopterin formyltransferase. Genomic digests of *M. thermoautotrophicum* DNA were probed with the partially pooled mix of oligonucleotides described under "Experimental Procedures." Each digest yielded two hybridizing bands of unequal intensity. The sizes of the hybridizing bands were as follows: *BamHI*, 4.44 and 4.81 kb; *EcoRI*, 5.07 and 3.45 kb; *PstI*, 4.81 and 1.94 kb, with the stronger of the two bands listed first. The *BamHI* fragments were pursued by isolating *BamHI*-digested DNA in the size range of 4-5 kb, ligating into *pBR322*, and transforming *E. coli* HB101 with the recombinant plasmid. Clones containing the two hybridizing fragments were identified and designated *ADM5* (*pADM5* contained the 4.8-kb fragment), and *ADM7* (*pADM7* contained the 4.4-kb fragment). These fragments were then subcloned into the phagemid *pTZ18R*, and the region hybridizing to the oligonucleotides was sequenced. The weaker hybridizing fragment (*pADM5*) was found to contain a nucleotide sequence which matched identically with the first 13 nucleotides of the 14 nucleotide oligomer probe, however the flanking region did not agree with the determined amino acid sequence. The absence of additional hybridizing bands in the genomic digests is convincing evidence that the *ftr* gene is present as a single copy on the *M. thermoautotrophicum* chromosome.

**Structure and Sequence of ftr**

A physical map of the 4.8-kb fragment is shown in Fig. 1; the synthetic oligonucleotide pool hybridized to the 0.35-kb *BamHI/PstI* fragment on the 5' end. The nucleotide sequence of *ftr* and its flanking regions are shown in Fig. 2. The open reading frame, starting at the 5' end, is 888 nucleotides long. The open reading frame is 888 nucleotides long and contains a single ATG codon, and is preceded by a 5'-untranslated region of 41 nucleotides. The 5'-untranslated region did not agree with the determined amino acid sequence, and the 5'-untranslated region did not contain a potential ribosome binding site. The 5'-untranslated region did not contain a potential ribosome binding site, and the 5'-untranslated region did not contain a potential ribosome binding site. The 5'-untranslated region did not contain a potential ribosome binding site.

**Restriction map of the BamHI fragment of pADM97**

The restriction map of the *BamHI* fragment of pADM97 is shown in Fig. 1. The synthetic oligonucleotide pool hybridized to the 0.35-kb *BamHI/PstI* fragment on the 5' end. The nucleotide sequence of *ftr* and its flanking regions are shown in Fig. 2. The open reading frame, starting at the 5' end, is 888 nucleotides long. The open reading frame is 888 nucleotides long and contains a single ATG codon, and is preceded by a 5'-untranslated region of 41 nucleotides. The 5'-untranslated region did not agree with the determined amino acid sequence, and the 5'-untranslated region did not contain a potential ribosome binding site. The 5'-untranslated region did not contain a potential ribosome binding site, and the 5'-untranslated region did not contain a potential ribosome binding site. The 5'-untranslated region did not contain a potential ribosome binding site, and the 5'-untranslated region did not contain a potential ribosome binding site, and the 5'-untranslated region did not contain a potential ribosome binding site.
acidic residues which resulted in a net charge of -17.7 at pH 7, and an overall pI of 4.8; highly charged peptides have been found to bind far less SDS per g of protein than the calibration proteins (24). This is supported by examples of anomalous behavior of highly acidic proteins on SDS-PAGE systems. Their genes, with molecular weights of 75,725 (pI 4.84) and 43,927 (pI 4.49). There is fairly good agreement.

The ribosome binding site is indicated as **RBS**, and the putative termination signals are **underlined with arrows** and **bold lines**. **TER** indicates the termination codon of **ftr**. The coding-like sequence present 5′ to **ftr** is indicated and terminates at position -61.

**Expression of the ftr Gene Product in E. coli**—To express the **ftr** gene product in **E. coli**, a 2.5-kb BamHI/SalI fragment containing **ftr** was ligated into pTZ18R. This construction, as part of an operon. Support for this hypothesis is based on the existing pattern of gene arrangements studied in methanogenic bacteria which display a motif of operonic clustering (9). Investigation of the upstream open reading frame would possibly encode the COOH terminus of a gene which precedes **ftr** as part of an operon. Support for this hypothesis is based on the existing pattern of gene arrangements studied in methanogenic bacteria which display a motif of operonic clustering (9).
was an FTR immunocross-reacting band on the immunoblot staining at M, 41,000 in the SDS-PAGE (Fig. 4, lane C), there was an FTR immunocross-reacting band on the immunoblot with antibody raised against homogeneous denatured FTR (E–D). Lanes A and I contain 80 and 8 μg of cell extract from HB101/pTZ18R (+IPTG). Lanes B and H contain 80 and 8 μg of cell extract from HB101/pADM11 (+glucose). Lanes C and G contain 80 and 8 μg of cell extract from HB101/pADM11 (+IPTG). Lanes D and F contain 5 and 0.5 μg of homogeneous FTR from M. thermoautotrophicum ΔH. Lane E contains 80 μg of cell extract from HB101/pADM97 (+IPTG).

(Fig. 4, lane G) with a mobility that coincided with homogeneous FTR from M. thermoautotrophicum, indicating that initiation of translation occurs at the same site. Were translation to start at either of the nearest in-frame initiation codons, the resulting polypeptides would differ by 2.8 kDa (initiation at position –85) or 7.8 kDa (initiation at position 220). While the accuracy of SDS-PAGE systems for the determination of molecular weights can be variable, the fact that the cloned ftr product and purified FTR from M. thermoautotrophicum have identical mobilities suggests that they are identical. Induction of the lac promotor appears to be essential for the synthesis of FTR in the pADM11 construct since no immunocross-reacting band is visible in extracts from cells grown under conditions where lac is repressed (Fig. 4, lane H), or where lac does not precede ftr (Fig. 4, lane E).

To test if the expressed FTR was active, ADM11 was grown under lac-inducing conditions (+IPTG), and the extract was assayed as described by Donnelly and Wolfe (1) at 60 °C. ADM11 cell lysates exhibited a specific activity which was 3-fold higher than the activity from M. thermoautotrophicum. However, when ADM11 was grown under lac repressing conditions (–IPTG, +glucose), no FTR activity was detected. The same is true for ADM97 extracts grown under inducing conditions (pADM97 contains ftr in the opposite orientation with respect to lac). This clearly demonstrates that ftr transcription is entirely dependent on the lac promotor in the construct pADM11.

The functional expression of this thermostable enzyme in cells grown aerobically at 37 °C is an interesting phenomenon which has been demonstrated previously with other enzymes from both eubacterial (28) and archaeabacterial (29, 30) sources. The implication that FTR is folded correctly in a

### Table I

| Codon | ftr Total no. of codons | Percent | mcr (A + B + C) Total no. of codons | Percent |
|-------|------------------------|---------|------------------------------------|---------|
| Phe UUU | 2 | 15 | 10 | 20 |
| Phe UUC | 11 | 85 | 39 | 79 |
| Leu UUA | 0 | 0 | 4 | 4 |
| Leu UUG | 0 | 0 | 2 | 2 |
| Leu CUC | 5 | 28 | 20 | 22 |
| Leu CUC | 5 | 28 | 36 | 40 |
| Leu CUA | 2 | 11 | 4 | 4 |
| Leu CUG | 6 | 33 | 24 | 26 |
| Ile AUU | 3 | 11 | 7 | 11 |
| Ile AUC | 3 | 11 | 20 | 32 |
| Ile AUA | 22 | 78 | 34 | 55 |
| Met AUG | 7 | 42 |
| Val GGU | 12 | 54 | 25 | 25 |
| Val GUC | 5 | 23 | 35 | 35 |
| Val GUA | 3 | 14 | 11 | 11 |
| Val GUG | 2 | 9 | 29 | 29 |
| Ser UCU | 0 | 0 | 4 | 7 |
| Ser UCC | 1 | 7 | 6 | 11 |
| Ser UCA | 9 | 64 | 22 | 56 |
| Ser UCG | 0 | 0 | 0 |
| Ser AGU | 2 | 14 | 5 | 9 |
| Ser AGC | 2 | 14 | 7 | 13 |
| Pro CCU | 5 | 42 | 12 | 23 |
| Pro CCC | 0 | 0 | 7 | 13 |
| Pro CCA | 7 | 58 | 31 | 60 |
| Pro CGG | 0 | 0 | 3 | 5 |
| Thr ACU | 1 | 6 | 4 | 6 |
| Thr ACC | 8 | 50 | 24 | 36 |
| Thr ACA | 6 | 38 | 38 | 57 |
| Thr ACG | 1 | 6 | 0 | 0 |
| Ala GCC | 7 | 23 | 24 | 17 |
| Ala GCC | 9 | 30 | 34 | 24 |
| Ala GCA | 13 | 43 | 79 | 56 |
| Ala GCG | 1 | 3 | 2 | 1 |
| Tyr UAU | 2 | 25 | 8 | 14 |
| Tyr UAC | 6 | 75 | 46 | 85 |
| His CAU | 2 | 67 | 1 | 4 |
| His CAC | 1 | 33 | 22 | 95 |
| Gin CAA | 0 | 0 | 2 | 4 |
| Gin CAG | 4 | 100 | 45 | 95 |
| Asn AAU | 2 | 18 | 2 | 4 |
| Asn AAC | 9 | 82 | 47 | 95 |
| Lys AAA | 6 | 27 | 28 | 50 |
| Lys AAG | 16 | 73 | 27 | 49 |
| Asp GAU | 9 | 47 | 16 | 20 |
| Asp GAC | 10 | 53 | 64 | 80 |
| Glu GAA | 10 | 38 | 73 | 69 |
| Glu GAG | 16 | 62 | 32 | 30 |
| Cys UGU | 1 | 20 | 7 | 77 |
| Cys UGC | 4 | 80 | 2 | 22 |
| Trp UGG | 0 | 7 |
| Arg CGU | 0 | 0 | 8 | 13 |
| Arg GGC | 0 | 0 | 1 | 1 |
| Arg GCA | 0 | 0 | 3 | 5 |
| Arg GCG | 0 | 0 | 0 | 0 |
| Arg AGA | 2 | 40 | 13 | 22 |
| Arg AGC | 3 | 60 | 34 | 57 |
| Gly GGU | 11 | 33 | 56 | 53 |
| Gly GGC | 7 | 21 | 11 | 10 |
| Gly GGA | 10 | 30 | 32 | 30 |
| Gly GGG | 5 | 15 | 6 | 8 |

*From Brown et al. (9).*

Induced ADM11 extracts showed no dramatic increase in staining at M, 41,000 in the SDS-PAGE (Fig. 4, lane C), there was an FTR immunocross-reacting band on the immunoblot with antibody raised against homogeneous denatured FTR (E–D). Lanes A and I contain 80 and 8 μg of cell extract from HB101/pTZ18R (+IPTG). Lanes B and H contain 80 and 8 μg of cell extract from HB101/pADM11 (+glucose). Lanes C and G contain 80 and 8 μg of cell extract from HB101/pADM11 (+IPTG). Lanes D and F contain 5 and 0.5 μg of homogeneous FTR from M. thermoautotrophicum ΔH. Lane E contains 80 μg of cell extract from HB101/pADM97 (+IPTG).

FIG. 3. Expression construct pADM11. The ftr gene follows downstream of the truncated lacZ' of pTZ18R and is expressed under control of the lac promotor. The lighter line indicates vector DNA.

FIG. 4. SDS-PAGE (A–D) and the corresponding immunoblot with antibody raised against homogeneous denatured FTR (E–D). Lanes A and I contain 80 and 8 μg of cell extract from HB101/pTZ18R (+IPTG). Lanes B and H contain 80 and 8 μg of cell extract from HB101/pADM11 (+glucose). Lanes C and G contain 80 and 8 μg of cell extract from HB101/pADM11 (+IPTG). Lanes D and F contain 5 and 0.5 μg of homogeneous FTR from M. thermoautotrophicum ΔH. Lane E contains 80 μg of cell extract from HB101/pADM97 (+IPTG).
Acknowledgments—We gratefully acknowledge the technical assistance of Victor Gabriel. We also thank J. Cronan, Jr., and M. Henry for helpful discussions.

REFERENCES

1. Donnelly, M. I., and Wolfe, R. S. (1986) J. Biol. Chem. 261, 16559-16560
2. DiMarco, A. A., Donnelly, M. I., and Wolfe, R. S. (1986) J. Bacteriol. 168, 1372-1377
3. Hartzell, P. L., Zvilius, G., Escalante-Semerena, J. C., and Donnelly, M. I. (1985) Biochem. Biophys. Res. Commun. 133, 884-890
4. Edlefsen, W. L., and Wolfe, R. S. (1981) J. Bacteriol. 256, 4259-4262
5. Rouviere, P. E., Escalante-Semerena, J. C., and Wolfe, R. S. (1985) J. Bacteriol. 162, 61-66
6. Wolfe, R. S. (1986) Trends Biochem. Sci. 10, 396-399
7. Rader, J. L., and Huennekens, F. M. (1974) in The Enzymes (Boyer, P. D., ed) Vol. XI, pp. 197-223, Academic Press, New York
8. Kay, L. D., Osborn, M. J., Hatefi, Y., and Huennekens, F. M. (1980) J. Biol. Chem. 255, 199-201
9. Brown, J. W., Daniels, C. J., and Reeve, J. N. (1989) CRC Crit. Rev. Microbiol. 16, 297-308
10. Zeikus, J. G., and Wolfe, R. S. (1972) J. Bacteriol. 109, 707-713
11. Bokranz, M., Baumann, G., Allmannsberger, R., Ankil-Fuchs, D., and Klein, A. (1988) J. Bacteriol. 170, 565-577
12. Konheiser, U., Pasti, G., Bollshweiler, C., and Klein, A. (1984) Mol. Gen. Genet. 198, 146-152
13. Cram, D. S., Sherf, B. A., Libby, R. T., Mattaliano, R. J., Ramachandran, K. L., and Reeve, J. N. (1987) Proc. Natl. Acad. Sci. U. S. A. 84, 3002-3006
14. Reeve, J. N., Beckler, G. S., Cram, D. S., Hamilton, P. T., Brown, J. W., Krzycki, J. A., Kolodziej, A. F., Alex, L., Orme-Johnson, W. H., and Walsh, C. T. (1989) Proc. Natl. Acad. Sci. U. S. A. 86, 3031-3035
15. Hamilton, P. T., and Reeve, J. N. (1985) J. Mol. Evol. 22, 351-359
16. Wallace, R. B., Schold, M., Johnson, M. J., Dembek, P., and Itakura, K. (1981) Nucleic Acids Res. 9, 3674-3686
17. Gunsalus, R. P., Romesser, J. A., and Wolfe, R. S. (1978) Biochemistry 17, 2374-2377
18. Miele, L., Kiener, A., and Leininger, T. (1983) Mol. Gen. Genet. 181, 480-484
19. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
20. Covarrubias, L., and Bolivar, F. (1982) Gene (Amst.) 17, 79-89
21. Kraft, R., Tardif, J., Krauter, K. S., and Leinwand, L. A. (1988) Biotechniques 6, 544-549
22. Blacksheer, P. J. (1984) Methods Enzymol. 104, 237-255
23. Towbin, H., Staehelin, T., and Gordon, J. (1979) Proc. Natl. Acad. Sci. U. S. A. 76, 4350-4354
24. Pitt-Rivers, R., and Impiombato, F. S. A. (1968) Biochem. J. 109, 519-527
25. Rock, C. O., and Cronan, J. E., Jr. (1979) J. Biol. Chem. 254, 9778-9785
26. Shuber, A. P., Orr, E. C., Recny, M. A., Schendel, P. F., May, H. D., Schauer, N. L., and Ferry, J. G. (1986) J. Biol. Chem. 261, 12942-12947
27. Neidhardt, F. C. (1987) Escherichia coli and Salmonella typhimurium. Cellular and Molecular Biology, pp. 1231-1240, American Society of Microbiology, Washington, D. C.
28. Lovell, C. R., Przybyla, A., and Ljungdahl, L. G. (1988) Arch. Microbiol. 149, 280-285
29. Weil, C. P., Beckler, G. S., and Reeve, J. N. (1987) J. Bacteriol. 169, 4857-4860
30. Fabry, S., Lehmacher, A., Bode, W., and Hensel, R. (1988) F e d. Eur. Biochem. Soc. 237, 213-217
31. Shannon, K. W., and Rabinowitz, J. C. (1988) J. Biol. Chem. 263, 7717-7725
32. Staben, C., and Rabinowitz, J. C. (1986) J. Biol. Chem. 261, 4629-4637
33. Hum, D. W., Bell, A. W., Rozen, R., and MacKenzie, R. E. (1988) J. Biol. Chem. 263, 15946-15950
34. Needleman, S. F., and Wunsch, C. D. (1970) J. Mol. Biol. 48, 443-453
35. Chu, F. K., Maley, G. F., Maley, F., and Belford, M. (1984) Proc. Natl. Acad. Sci. U. S. A. 81, 3049-3053
36. Grumont, R., Washtien, W. L., Caputi, D., and Santi, D. V. (1986) Proc. Natl. Acad. Sci. U. S. A. 83, 5387-5397
37. Freisheim, J. H., Bitar, K. G., Reddy, A. V., and Blankenship, D. T. (1978) J. Biol. Chem. 253, 6437-6444
38. Masters, J. N., and Attardi, G. (1983) Gene (Amst.) 21, 59-63
39. Takeishi, K., Kaneda, D. A., Kimiko, S., Gotoh, O., and Seno, T. (1985) Nucleic Acids Res. 13, 2005-2004

foreign host when grown at a much lower temperature than in its natural host is especially encouraging for future work in cloning additional thermostable enzymes involved in methanogenesis.

Comparison of FTR with Folate Binding Proteins—FTR represents a novel enzyme with no true catalytic homolog. To see if FTR shared structural homology with other proteins, a search of the Genbank™ protein sequence data bank was conducted. In addition, the protein sequences of formyltetrahydrofolate synthetase from a number of sources (31-33) were aligned with FTR by the method of Needleman and Wunsch (34) using the Aalign program of DNASTAR™ (Madison, WI). No extensive homologies were found. Despite the absence of extensive homology, it is possible that these or other folate utilizing proteins might share a common pterin binding structure with FTR. To test this, protein sequences from thymidylate synthetase (5,10-methylentetrahydrofolate:UMP C-methyltransferase, EC 2.1.1.45) (35, 36), and dihydrofolate reductase (EC 1.5.1.3) (37, 38) were compared with FTR as described above. No common sequences were observed which could be designated as possible pterin binding sites. This is not surprising in light of the absence of common folate binding structures among different folate utilizing enzymes (32, 39). No extensive homologies were found. Despite the absence of common sequences, which could be designated as possible pterin binding sites. This is not surprising in light of the absence of common folate binding structures among different folate utilizing enzymes (32, 39). It appears that FTR has a novel pterin binding structure; whether or not this will be conserved among other tetrahydrofolate-utilizing enzymes is unknown.

Acknowledgments—We gratefully acknowledge the technical assistance of Victor Gabriel. We also thank J. Cronan, Jr., and M. Henry for helpful discussions.
The formylmethanofuran: tetrahydromethanopterin formyltransferase from Methanobacterium thermoautotrophicum delta H. Nucleotide sequence and functional expression of the cloned gene.
A A DiMarco, K A Sment, J Konisky and R S Wolfe

J. Biol. Chem. 1990, 265:472-476.