JOURNAL OF BACTERIOLOGY, Mar. 1991, p. 2131–2133
0021-9193/91/062131-03$02.00/0

Vol. 173, No. 6

Nucleotide Sequence of the Gene Encoding the Vanadate-Sensitive Membrane-Associated ATPase of *Methanococcus voltae*

RITA DHARMAVARAM, PATRICK GILLEVET, AND JORDAN KONISKY*

Department of Microbiology, University of Illinois, Urbana, Illinois 61801

Received 9 October 1990/Accepted 7 January 1991

Methanococcus voltae **contains a membrane-associated ATPase whose structural gene has been sequenced. The gene encodes 565 amino acids and includes a 12-amino-acid N-terminal sequence which is not present in the purified enzyme. On the basis of its amino acid sequence, the *M. voltae* enzyme is unrelated to previously characterized ATPases.**

Methanococcus voltae is a marine methanogen and a member of the domain *Archaea* (13). We have previously reported the isolation of a membrane-associated, vanadate-sensitive ATPase from this organism whose catalytic cycle involves formation of a phosphorylated intermediate (3, 4). This is a property which it shares with a class of eubacterial and eukaryotic ATPases, termed P-type enzymes, which function in ion translocation (9, 10). However, in contrast to known P-type enzymes which are integral membrane proteins, the methanogen enzyme can be removed from the membrane by salt extraction (4), a property of some peripheral membrane proteins. That the methanogen enzyme is, indeed, a novel type of ATPase is confirmed by the present study, which demonstrates that the *M. voltae* enzyme is not structurally related to any previously studied ATPase.

Isolation of the ATPase gene and its nucleotide sequence. Plaques formed after infection of *Escherichia coli* XL-1 Blue with a lambda ZAP II genomic library of *M. voltae* containing 2- to 6-kb inserts in the *Eco*RI site of the bacteriophage vector were probed with a polyclonal antibody raised against purified ATPase (3). The custom library was constructed by Stratagene (La Jolla, Calif.), and the pico-Blue immunoscreening kit from Stratagene was used for screening. Of 1.5×10^5 plaques screened, 4 gave positive signals.

A single-stranded template was isolated from one of these clones after superinfection of XL-1 Blue with helper phage M13K07 and subjected to DNA sequence analysis by the dideoxy-chain termination method with the Sequenase 2.0 kit from United States Biochemicals (Cleveland, Ohio) and [α-^{32}P]dATP (Amersham, Arlington Heights, Ill.) as the labeled nucleotide. To facilitate sequencing, the Promega Erase a Base system was used to generate a series of unidirectional deletion clones. In all, 87 sequences were used to generate the complete sequence of the ATPase gene. DNA manipulations were carried out according to established procedures (8).

Features of the sequence. An open reading frame extends from nucleotides 35 to 1732 and is capable of encoding a protein of 565 amino acids with a calculated molecular weight of 59,702. While we have previously assigned an apparent molecular mass of 74 kDa on the basis of sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis analysis of the purified enzyme, the discrepancy is not surprising in view of the acidic character of the enzyme, which would lead to decreased binding to SDS. To verify that this open reading frame encodes the ATPase, we compared the derived sequence to the N-terminal amino acid sequence of the purified enzyme. DEAE-Sephadex-purified (3) *M. voltae* ATPase was excised from an SDS gel, and its N-terminal sequence was determined by the University of Illinois Biotechnology Center. The 14-residue N-terminal sequence obtained, V-E-K-I-G-D-V-E-G-F-K-V-I-D, matched perfectly with amino acid residues 13 through 26 of the open reading frame, indicating that the enzyme is processed. The possibility that the amino terminus of the protein sample used for sequencing was degraded was judged extremely unlikely on the basis of the excellent quality of the N-terminal sequence analysis.

The role of the 12-amino-acid leader sequence remains to be elucidated. It is possible that this region plays some role in localization of the enzyme to the membrane in a manner similar to that of signal sequences found in many eubacterial and eukaryotic membrane proteins (11). Indeed, the 12-amino-acid sequence does show limited similarity to known signal peptides, including a hydrophobic core region and a conserved cleavage site; 88% of prokaryotic and 70% of eukaryotic signal sequences have alanine or glycine at the cleavage site. However, the *M. voltae* sequence lacks a positively charged amino acid in its N-terminal region, a situation found in most reported signal sequences. That membrane association is based on direct interaction of the enzyme with lipids or interaction with an ancillary integral membrane protein are possibilities to be explored.

The initiation codon is immediately preceded by a presumptive ribosome-binding sequence which is similar to that identified upstream of other sequenced methanogen genes (1). We have no information on possible transcription initiation sequences, since the region where such a sequence is usually found in other archaebacteria, about 25 bp upstream of the initiation codon (1), was not present in our clones. A putative transcription termination structure comprising a region of dyad symmetry which is capable of forming a hairpin loop structure is illustrated in Fig. 1. The hairpin has a free energy of -10.5 kcal (1 cal = 4.184 J) calculated by the fold program of Zucker and Stiegler (14) using the energy values of Freier et al. (5) and is followed by a run of T residues, indicative of a rho-independent termination structure. Similar putative transcription termination structures have been identified in eubacteria and other archaebacteria (1).

While residues 432 to 436, E-D-K-E-D-T-K, show some similarity to the conserved region consisting of S/L-D-K-T-G-T-I/L-T (10), which is found in the region containing the

* Corresponding author.

```
      10                  30                  50                  70                  90
AAAAAAAATCGGTGCAATTGCTGCAGGTAGCGCTATGGTAGCTTCAGCATTAGCTACAGGCGTTTTCGCTGTAGAAAAAATCGGTGACGT
                ******* **M  V  A  S  A  L  A  T  G  V  F  A  V  E  K  I  G  D  V

            110                 130                 150                 170
AGAAGGATTTAAAGTAATAGACAACGGTGAACCAACCGCAGATATCGTAGTAGGTTCAACAGCTGCTGCTGCTGACGTTGTTTCAGCTGC
  E  G  F  K  V  I  D  N  G  E  P  T  A  D  I  V  V  G  S  T  A  A  A  A  D  V  V  S  A  A

      190                 210                 230                 250                 270
AAACGTAGCTGCAAAAGTTGGCTCAATGATGTTCAAAGAAGGAGAAGCTGCTTCAGGTTCAGCAAAATTAACAGTTAAAGCATCAGCAGA
  N  V  A  A  K  V  G  S  M  M  F  K  E  G  E  A  A  S  G  S  A  K  L  T  V  K  A  S  A  E

            290                 310                 330                 350
ATCAGACGATGCTAACTTAAAAAGCTTATTGACAAACGGTACAAACGACTTTACGGAATTAGATGCAGGTAAAGAAGCATTTGTAGTTGC
  S  D  D  A  N  L  K  S  L  L  T  N  G  T  N  D  F  T  E  L  D  A  G  K  E  A  F  V  V  A

      370                 390                 410                 430                 450
TGCTGCTGATAGTGATTATTCAGATGCTTTAATAAACGCAACAACAGGATTTGCAAACATTGCTGATAACGTATTATACGACCAGGCAAA
  A  A  D  S  D  Y  S  D  A  L  I  N  A  T  T  G  F  A  N  I  A  D  N  V  L  Y  D  Q  A  K

            470                 490                 510                 530
ATTAGCTGCTGCTGTTTCATTAGGAGACTTAAGCACATTATCAGTTGTTAAAGACATTGACCCATCAGACTGGTACGCTGATAAAAATAA
  L  A  A  A  V  S  L  G  D  L  S  T  L  S  V  V  K  D  I  D  P  S  D  W  Y  A  D  K  N  K

      550                 570                 590                 610                 630
AGCTGCAGACGTAGCAACCAAAGACTACTACGACCAAGATGGTGACGCAGTAGAGATGTTAATGGCTACAGTTGCTTCAAACGATGACGG
  A  A  D  V  A  T  K  D  Y  Y  D  Q  D  G  D  A  V  E  M  L  M  A  T  V  A  S  N  D  D  G

            650                 670                 690                 710
TAAATCATTAACCGTTGATGAAGATGGTGTATTATATGCTTCAATTGCATACGATGACGATAACGAAGACTTCCAAAGAGCTACACAAGT
  K  S  L  T  V  D  E  D  G  V  L  Y  A  S  I  A  Y  D  D  D  N  E  D  F  Q  R  A  T  Q  V

      730                 750                 770                 790                 810
TTTAAAAGAAGGTAACAGATTACCATTCTTAGGTGAAGAATACGCTTTAGTTAAATTAGATACAGATGACGACATTGTATACTTAGGTAA
  L  K  E  G  N  R  L  P  F  L  G  E  E  Y  A  L  V  K  L  D  T  D  D  D  I  V  Y  L  G  K

            830                 850                 870                 890
AGAAGTATTCGATGGAGTTTTAAAAGAAGGCGACACCTACAACATCGGTGACGGTTACGAATTAAAAGTTGTAGCTATCTTAAAATCTGG
  E  V  F  D  G  V  L  K  E  G  D  T  Y  N  I  G  D  G  Y  E  L  K  V  V  A  I  L  K  S  G

      910                 930                 950                 970                 990
TGACGAATACAAAATCAGCTTACAATTAATGAAAGATGGAAAAGTAGTTGCTGAAAAATTCGATAAAGTTAGTGCAACAAGCGCATTGAA
  D  E  Y  K  I  S  L  Q  L  M  K  D  G  K  V  V  A  E  K  F  D  K  V  S  A  T  S  A  L  K

           1010                1030                1050                1070
AATGATATACACACCAGGAAACATTGGTATTGTAGTTAACGAAGCATGGGAAAACGTTGGTCAAGACTATGGTTACGGAAGCACCTTAAT
  M  I  Y  T  P  G  N  I  G  I  V  V  N  E  A  W  E  N  V  G  Q  D  Y  G  Y  G  S  T  L  I

     1090                1110                1130                1150                1170
CACAAAAGATGTTATTGCTTTAGAACTTGGTGAAGAATACATACCAGACTGGGAAGTTGTTACAATTGAAAAAGACACAACTACAGATAA
  T  K  D  V  I  A  L  E  L  G  E  E  Y  I  P  D  W  E  V  V  T  I  E  K  D  T  T  T  D  N

           1190                1210                1230                1250
CACCAAAGATTCAAAAATGACACTCTCTGACGATAAAATCACAAAAGATAATACATACGGTATTGGTTTACAATATGTTGGTGACGAAGA
  T  K  D  S  K  M  T  L  S  D  D  K  I  T  K  D  N  T  Y  G  I  G  L  Q  Y  V  G  D  E  E

     1270                1290                1310                1330                1350
AGATAACTTCAAATCAGGAAAAGCAATCAAAATTGCAAAATACGCTGAACTTGAATTAGATGATGAAGACAAAGAAGATACAAAATTAAA
  D  N  F  K  S  G  K  A  I  K  I  A  K  Y  A  E  L  E  L  D  D  E  D  K  E  D  T  K  L  N

           1370                1390                1410                1430
CTTGTTCTTCTCAATGGATGAAACCAAAGAAGCAACATTAGCTGCTGGTCAAAAAGTTACAGTATTAAACTCAGACATTACCTTATCAGA
  L  F  F  S  M  D  E  T  K  E  A  T  L  A  A  G  Q  K  V  T  V  L  N  S  D  I  T  L  S  E

     1450                1470                1490                1510                1530
AGTAATGGCTGACGCTAAAGCTCCTGTAGCATTCAAAGCACCATTAGCTGTATTAGATACAGAAGTATCATTAGATGCTGCTAACAAAAA
  V  M  A  D  A  K  A  P  V  A  F  K  A  P  L  A  V  L  D  T  E  V  S  L  D  A  A  N  K  K

           1550                1570                1590                1610
ATTAATCTTAGTTGGTGGTCCTGTAGCTAACGCTTTAACAAAAGAATTAGCTGACGCTGGAAAAATAGAAATGACAGTTGAATCACCTGC
  L  I  L  V  G  G  P  V  A  N  A  L  T  K  E  L  A  D  A  G  K  I  E  M  T  V  E  S  P  A

     1630                1650                1670                1690                1710
AACATTAGCTGTAGTAGCTGGTGCTGCAAACGGTAACGATGTATTAGTAGTAGCTGGTGGCGACAGAGCTGCAACAGCAGAAGCTGCTAA
  T  L  A  V  V  A  G  A  A  N  G  N  D  V  L  V  V  A  G  G  D  R  A  A  T  A  E  A  A  N

           1730                1750                1770                1790
CGCATTAATCGAAATGTTATAATTTGAATTATTCAAACTATAATTTTTCAACCAATATTAAAAGTGGGATTTTTCCCACTTTTTTTATTT
  A  L  I  E  M  L  *
```

FIG. 1. Nucleotide and deduced amino acid sequences of *M. voltae* ATPase. The sequence of the nonsense strand is shown, with the single-letter designation for each amino acid indicated below each codon. The bases of the putative ribosomal binding site are indicated by asterisks. A region of dyad symmetry stretches from nucleotide 1770 to 1793.

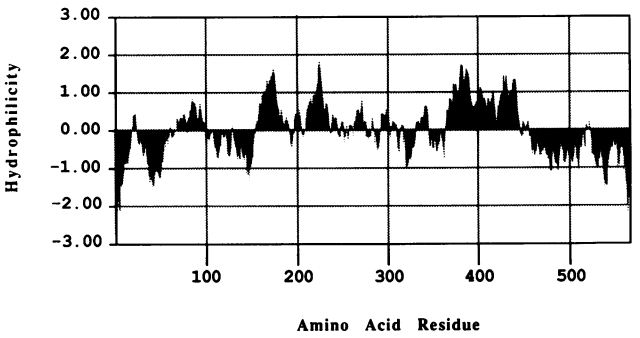

FIG. 2. Hydrophilicity plot of *M. voltae* ATPase generated by the method of Kyte and Doolittle (6). The window parameter used was 19 residues.

phosphorylated aspartate residue of ATPases known to form an acyl phosphate intermediate, it will be necessary to determine the phosphorylated residue of the *M. voltae* enzyme more directly.

While *M. voltae* ATPase is inhibited by vanadate and forms an acyl phosphate intermediate, unlike characterized P-type ATPases, it shows properties which suggest that it is not an integral membrane protein (4). This conclusion is supported by its hydropathy profile (6) (Fig. 2), which suggests that the enzyme does not have regions of transmembrane orientation, a common feature of characterized P-type enzymes (12).

Lack of homology with other ion-transporting ATPases. A search of the GenBank and EMBL data bases using either the GAP or the FASTP algorithm (7) provided in the University of Wisconsin Genetics Computer Group sequence analysis software package (2) failed to identify resident sequences that were significantly related to the *M. voltae* ATPase protein sequence. We next compared the *M. voltae* sequence to those of characterized ion-translocating ATPase structural genes by using the rdf2 program of Pearson and Lipman on the Vax/VMS 5.0 (7). *M. voltae* ATPase did not display significant identity to the following sequences resident in GenBank: P type (human Na^+-K^+, *Escherichia coli* Kdp, *Staphylococcus aureus* Ca^{2+}, *Streptococcus faecalis* K^+, *Arabidopsis thaliana* β subunit, *Neurospora crassa* β, and *Saccharomyces cerevisiae* β), F_1F_0 type (*Anabaena* sp. β and *E. coli* β), V type (*N. crassa*), and archaebacterial ATPases (*Methanobacterium barkeri* α and *Sulfolobus acidocaldarius* α).

The physiological function of *M. voltae* ATPase remains to be elucidated. Possibilities include a role in pH homeostasis regulation of internal sodium ion levels, generation and maintenance of transmembrane ionic gradients, and ATP production.

Nucleotide sequence accession number. The GenBank accession number for the sequence described here is M59200.

We thank Bruce Roe, University of Oklahoma, for making computer programs of the University of Oklahoma VAX available to us.

These studies were supported by National Institutes of Health grant GM41587.

REFERENCES

1. **Brown, J. W., C. J. Daniels, and J. N. Reeve.** 1989. Gene structure, organization, and expression in archaebacteria. Crit. Rev. Microbiol. **16**:287–338.
2. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12**:387–395.
3. **Dharmavaram, R., and J. Konisky.** 1989. Characterization of a P-type ATPase of the archaebacterium Methanococcus voltae. J. Biol. Chem. **264**:14085–14089.
4. **Dharmavaram, R. M., and J. Konisky.** 1987. Identification of a vanadate-sensitive, membrane-bound ATPase in the archaebacterium *Methanococcus voltae*. J. Bacteriol. **169**:3921–3925.
5. **Freier, S. M., R. Kierzek, J. H. Jaeger, N. Sugimoto, M. H. Caruthers, T. Neilson, and D. H. Turner.** 1986. Improved free-energy parameters for predictions of RNA duplex stability. Proc. Natl. Acad. Sci. USA **83**:9373–9377.
6. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157**:105–132.
7. **Lipman, D. J., and W. R. Pearson.** 1985. Rapid and sensitive protein similarity searches. Science **227**:1435–1441.
8. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
9. **Pedersen, P. L., and E. Carafoli.** 1987. Ion motive ATPases. I. Ubiquity, properties, and significance to cell function. Trends

Biol. Chem. **12**:146–150.
10. **Pedersen, P. L., and E. Carafoli.** 1987. Ion motive ATPases. II. Energy coupling and work output. Trends Biol. Chem. **12**:181–186.
11. **Pollitt, S., and M. Inouye.** 1987. Structure and function of the signal peptide, p. 117–139. *In* M. Inouye (ed.), Bacterial outer membranes as model systems. John Wiley & Sons, Inc., New York.
12. **Serrano, R.** 1988. Structure and function of proton translocating ATPase in plasma membranes of plants and fungi. Biochim. Biophys. Acta **947**:1–28.
13. **Woese, C. R., O. Kandler, and M. L. Wheelis.** 1990. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, Eucarya. Proc. Natl. Acad. Sci. USA **87**:4576–4579.
14. **Zucker, M., and P. Stiegler.** 1980. Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. Nucleic Acids Res. **9**:133–148.