
Termination of a transcription unit comprising highly expressed genes in the archaeobacterium *Methanococcus voltae*

B.Müller, R.Allmansberger and A.Klein*

Molecular Genetics, Department of Biology, Philipps University, P.O. Box 1929, D-3550 Marburg, FRG

Received 23 July 1985; Revised and Accepted 3 September 1985

ABSTRACT

The 3' termini of transcripts originating from genes organized in a highly expressed transcription unit were analyzed in the archaeobacterium *Methanococcus voltae*. The putative termination signals were found in an AT-RICH intergenic region following the 3'-terminal gene. The two detected signals both contain oligo(T) sequences. A possible stem/loop structure immediately precedes one of the oligo(T) tracts. This secondary structure is considered to have an additional function in stabilizing the transcripts.

INTRODUCTION

Methanogenic bacteria are the largest known subgroup of archaeobacteria, a procaryotic division of organisms, which, on the basis of 16S rRNA cataloguing has been considered a separate kingdom beside the eubacteria and the eucaryotes. This view has been substantiated by numerous investigations showing many different traits which are typical for archaeobacteria (see (1) for review).

We have been interested in mechanisms of gene expression of methanogenic bacteria on both levels, transcription and translation. The genetic code (2, 3) as well as the translation signals (i.e. ribosome binding sites) appear to be identical to those used in eubacteria (2, 3, 4). The same situation cannot be expected for transcription, since it has been shown that the DNA-dependent RNA polymerases of archaeobacteria differ in their compositions from their eucaryotic and eubacterial counterparts. Still, immunological analysis indicates that they might be of monophyletic origin (5).

We have previously shown that transcription starts occur in AT-rich intergenic regions on the genome of the methanogenic bacterium *Methanococcus voltae* (4). Such AT-rich sequences are commonly found flanking structural genes in methanococci (2, 4). Promoter structures have not yet been identified in these regions even though attempts were made to define consensus sequences (6, 7).

Recently a polycistronic message was detected in Mc. voltae. It is the transcript of the strongly expressed structural genes of methyl CoM reductase component C (8). Strongly expressed transcription units usually end in strong terminators. We therefore expected to be able to detect termination signals in the region following the reductase genes. The 3'-terminal gene encodes the subunit α of the reductase complex. We have concentrated on the region following this gene to look for the area of transcription termination. We describe here the sites and the apparent signals used in termination of the transcript and discuss their relationships to known eubacterial and eucaryotic terminators.

MATERIALS AND METHODS

Bacteria.

Mc. voltae (DSM 1537) was obtained from the German Collection of Microorganisms, Göttingen, and grown as described previously (9). Escherichia coli HB101 (10) was a gift from F. Zinoni (Munich).

Plasmids.

Plasmids pNPT20 (8) or pACYC184 (11) were used for cloning the HindIII or HindIII/BamHI fragments, respectively, both carrying the 3' end of the methyl CoM reductase gene α (8, compare also Figure 1).

Enzymes.

Restriction endonucleases were purchased from Boehringer (Mannheim, FRG) or PL Biochemicals (Freiburg, FRG), T4 polynucleotide kinase and DNA polymerase, large fragment were purchased from PL Biochemicals.

Radiochemicals.

^{32}P -nucleotides were purchased from Amersham Buchler GmbH (Braunschweig, FRG), γ - ^{32}P -ATP was a gift from J. Schallenberg (Marburg).

Isolation of RNA from Mc. voltae.

The cells were harvested anaerobically at a cell density of 2×10^8 cells/ml by centrifugation at $12,000 \times g$ for 5 min at 4°C . The supernatant was discarded and 1×10^{11} cells were resuspended in 10 ml of homogenization buffer (4.5 M guanidinium thiocyanate, 50 mM Tris HCl pH 7.6, 0.2% sodium laurylsarcosinate, 0.14 M β -mercaptoethanol). The suspension was mixed with an equal volume of phenol, equilibrated with 0.1 M Tris HCl pH 8, and gently agitated at 65°C . After 10 min 5 ml of acetate buffer (0.1 M sodium acetate pH 5, 10 mM Tris HCl pH 7.4, 1 mM EDTA pH 8) was added and the mixture was agitated for 10 min at 65°C . Subsequently 10 ml of chloroform/isoamylalcohol, 24 : 1 were added and the suspension was agitated for 10 min at 65°C . The

aqueous phase was separated by centrifugation and the organic phase reextracted with 5 ml of homogenization buffer. The combined aqueous phases were extracted with chloroform/isoamylalcohol. The nucleic acids were precipitated by adding 0.1 volume 3 M sodium acetate pH 5.2 and 2.5 volumes ethanol and cooling the sample for at least 4 h at -20°C . The precipitate was collected by centrifugation at $12,000 \times g$ for 15 min at 4°C and washed twice with 80% ethanol.

Other methods.

Nucleotide sequence determination of DNA was performed using the chemical cleavage method (12). DNA recombinant work was done according to methods described by Maniatis et al. (13). Analysis of DNA/RNA hybrids using nuclease S1 has been described previously (8).

RESULTS

The 3' end of the continuous open reading frame comprising previously defined gene sequences (8) is located in between two EcoRI sites as shown on the restriction map in Figure 1. Figure 2 gives the sequence of the 3' end of the gene and the region immediately downstream. This intergenic region is very AT-rich (85% compared to approximately 60% of coding regions). Within the noncoding sequence an inverted repeat exists, which can be drawn to form a stem/loop structure (see Figure 4, below). This structure is followed by two oligo(T) sequences spaced 5 nucleotide pairs apart.

In order to check whether the 3' end of the reductase transcript might be located close to this area, the pACYC184 plasmid containing the Methanococcus HindIII/BamHI-fragment (compare Methods and Figure 1) was cut with BstEII, the 3' termini were ^{32}P -labelled, and the BstEII/BamHI-fragment was isolated. It was hybridized against Mc. voltae cellular RNA and the hybrid treated with nuclease S1. The resulting material was electrophoretically analyzed on a denaturing polyacrylamide gel. The result is shown in Figure 3. The 3' ends of the transcripts appear mainly in the oligo(T) tract following the hairpin

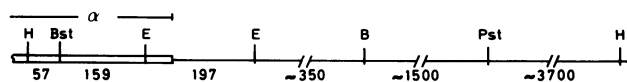


Figure 1. Restriction map of the 3' end of the Methanococcus voltae methyl CoM reductase α gene and the adjacent region. The cleavage sites indicated are B, BamHI; Bst, BstEII; H, HindIII; Pst, PstI. The 3' end of the α gene is indicated on the left. The numbers indicate fragment lengths given in base pairs.

```

--- GCT AAA ATC TAA gtttaattactaattttattattataattttattatttagattgggcaaaaata
   CGA TTT TAG ATT caattaaatgattaaataataataataataataataataataatactaacccgttttat

```

```

gtaaaagaaaactaaaggaaccctaataatggtttcctttttttatataatttttaaaaattgatt---
cattttccttttgatttcccttttgattataccaagggaataataataaaaaatttttaactaa---

```

Figure 2. Partial sequence of the *EcoRI* fragment comprising the 3' end of the α gene and part of the intergenic region downstream. Sequencing of both strands was performed using the *HindIII/PstI* fragment (Figure 1) labelled at the 5' or 3' ends of the opposite strands at the *HindIII* restriction site. The coding sequence is given in triplets with the stop codon indicated in bold type. The inverted repeats in the noncoding region are underlined.

structure and, to a lesser extent, in the oligo(T) region immediately downstream. An S1 signal is also seen in the loop region of the putative hairpin. It is remarkable that S1 signals are absent from AT-rich sequences following further downstream. This was confirmed by overexposure of the autoradiogram, whereby faint bands became visible only upstream of the strong S1 signals, whereas no such material was detectable downstream, up to the *BamHI* end of the restriction fragment used (data not shown).

DISCUSSION

It could be argued that the 3' ends of the RNA molecules seen in our experiment might be caused by processing of a primary transcript. As mentioned in the results, however, protection of the labelled DNA fragment could not be detected beyond the 3' ends seen, even at minor levels, although such bands were detected in small amounts upstream of the terminus, indicating that degradation products and/or nascent chains of RNA could be picked up by the assay. This does not absolutely rule out that the signal(s) detected are processing signals but makes homology to eubacterial transcription terminators much more likely. The structural resemblance discussed below adds to this argument.

Our results suggest that oligo(T) sequences serve as termination sites in the methyl reductase gene transcription. The termination is likely to be influenced by secondary structure of the template and/or transcript close to the termination site. The putative stem/loop structure described above is probably sufficiently stable under *in vivo* conditions as are those found in eubacterial cells, as discussed below. In fact, the high internal salt concentration in *Methanococcus* cells (14) would favor hydrogen bonding. Also, the S1 signal found in the possible hairpin loop is most likely due to its

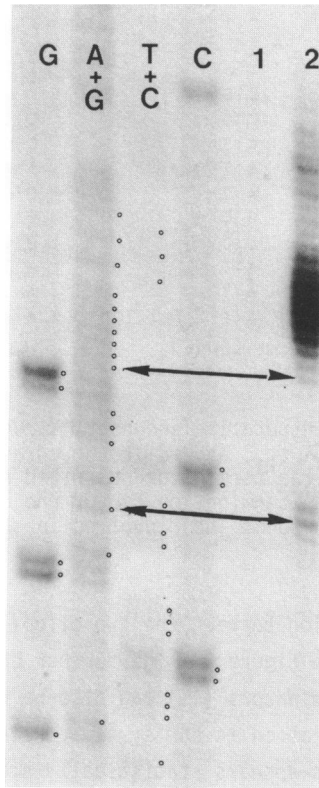


Figure 3. S1 mapping of the 3' termini of *Methanococcus voltae* cellular RNA in the noncoding region following the α gene. The HindIII/BamHI fragment shown in Figure 1 was subcloned in pACYC184. The BstEII/BamHI subfragment labelled with ^{32}P at the BstEII 3' end was used for the hybridization against the RNA. S1 digestion and electrophoretic analysis are described in the methods. Lanes: G, A + G, T + C, and C, sequence of the labelled noncoding strand obtained by the chemical method; 1, control: S1 digestion without added RNA; 2, S1 digestion after DNA/RNA hybrid formation. Note that the sequencing fragments move ahead approximately 1 nucleotide as compared to the S1 fragments due to the elimination of the modified bases. The alignment is indicated by the double headed arrows.

formation under the hybridization conditions applied. The putative hairpin structure strikingly resembles eubacterial terminators (15) as shown in Figure 4. Hyphenated inverted repeats are regular constituents of such factor independent termination signals. Thus, in *E. coli* poly(U) tracts following hairpins in the transcripts lead to termination, independent of the known termination factor rho (16). This is interpreted in terms of the low stability of dA : rU compared to dT : rA base pairs. This view is consistent with

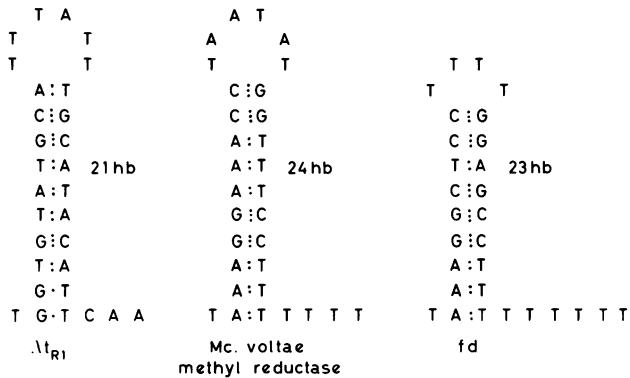


Figure 4. Comparison between putative secondary structures of the terminators of the *E. coli* phages λ (λ_{R1}) and fd (both redrawn from Rosenberg and Court (15)) and the *Mc. voltae* terminator described here. Note the lack of an oligo(T) at the 3' end of the termination factor rho dependent terminator. The maximal numbers of hydrogen bonds (hb) involved in stem formation are indicated.

the observed weak termination between the two oligo(T) tracts in the α gene termination region (compare Figure 3). One further characteristic feature known from eubacterial terminators (15) can also be recognized in the sequence implicated in termination in our system: preceding the hairpin we see 4 contiguous GC pairs. This appears statistically significant in a region containing an average of only 15% GC pairs.

How does the termination region described here compare to other termination regions in archaeobacteria and to eucaryotic termination signals? A putative termination signal downstream of the *Halobacterium halobium* bacterio-opsin gene exhibits an inverted repeat but lacks an oligo(T) tract (19). Sequences capable of forming hairpin structures have been found at the end of the rRNA transcription units in *Mc. vanniellii* (7). The lower number of hydrogen bonds involved would suggest the requirement for termination factors at these signals, should they be implicated in transcription termination. In eucaryotes RNA polymerase III (17) and possibly RNA polymerase I (18) recognize oligo(T) tracts as termination signals. A possible stem/loop structure has been described in the region downstream of the mouse β -globin (major) gene. Its large loop, however, would tend to destabilize the possible secondary structure.

The relative importance of the elements found in the termination region described here could best be evaluated in an *in vitro* transcription system.

Recent initial work with such a system from Methanococcus thermolithotrophicus (20) would appear promising in this respect.

The regulation mechanisms of gene expression include differential messenger stabilities (see (16) for review). Therefore, we interpret the hairpin structure at the end of the methyl CoM reductase transcript as a stabilizing element in addition to its function as part of a termination signal. Analysis of the relative stability of the reductase messenger compared to other RNA species in Mc. voltae is in progress.

ACKNOWLEDGMENTS

We thank H. Bestgen for expert technical assistance and H. Steinebach for help in preparation of the manuscript. This work was supported by the Deutsche Forschungsgemeinschaft.

*To whom correspondence should be addressed

REFERENCES

1. Kandler, O. ed. (1982) Archaeobacteria. Gustav Fischer Verlag, Stuttgart.
2. Cue, R., Beckler, G.S., Reeve, J.N., Konisky, J. (1985) Proc. Natl. Acad. Sci. USA, in press.
3. Allmansberger, R., Bollschweiler, C., Konheiser, U., Müller, B., Muth, E., Pasti, G., Klein, A. (1985) Syst. Appl. Microbiol., in press.
4. Bollschweiler, C., Kühn, R., Klein, A. (1985) EMBO J. 4, 805-811.
5. Zillig, W., Schnabel, R., Stetter, K., Thomm, M., Gropp, F., Reiter, W.D. (1985) in: FEMS Symposium on the Evolution of Prokaryotes, Academic Press, New York, in press.
6. Reeve, J.N., Hamilton, P.T., Beckler, G.S., Morris, C.J. (1985) Syst. Appl. Microbiol., in press.
7. Wich, G., Böck, A. (1985) Syst. Appl. Microbiol., in press.
8. Konheiser, U., Pasti, G., Bollschweiler, C., Klein, A. (1984) Mol. Gen. Genet. 198, 146-152.
9. Whitman, W.B., Ankwarda, E., Wolfe, R.S. (1982) J. Bact. 149, 852-863.
10. Boyer, H.W., Roulland-Dussoix, D. (1969) J. Mol. Biol. 41, 459-472.
11. Chang, A.C.Y., Cohen, S.N. (1978) J. Bact. 134, 1141-1156.
12. Maxam, A.M., Gilbert, W. (1980) Methods Enzymol. 65, 499-560.
13. Maniatis, T., Fritsch, E.F., Sambrook, J. (1982) Molecular cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y.
14. Jarrel, K.F., Sprott, G.D., Matheson, A.T. (1984) Can. J. Microbiol. 30, 663-668.
15. Holmes, W.M., Platt, T., Rosenberg, M. (1983) Cell 32, 1029-1032.
16. Rosenberg, M., Court, D. (1979) Ann. Rev. Genet. 13, 319-353.
17. Citron, B., Falck-Pedersen, E., Salditt-Georgieff, M., Darnell, J.E. (1984) Nucleic Acids Res. 12, 8723-8731.
18. Bakken, A., Morgan, G., Morgan, G., Sollner-Webb, B., Roan, J., Busby, S., Reeder, R.H. (1982) Proc. Natl. Acad. Sci. USA 79, 56-50.
19. Dassarma, S., Rajbhandary, U.L., Khorana, H.G. (1984) Proc. Natl. Acad. Sci. USA 81, 125-129.
20. Thomm, M., Stetter, K.O. (1985) Eur. J. Biochem. 149, 345-351.