Biology of Anaerobic Bacteria, edited by H.C. Dubourguier et al.
Elsevier Science Publishers B.V., Amsterdam, 1986 — Printed in The Netherlands

ANALYSIS OF FUNCTIONALLY RELATED GENE GROUPS IN METHANOGENIC BACTERIA

J. Schallenberg[1], M. Thomm[2], K.O. Stetter[2], M. Moes[1], M. Truss[1], R. Allmansberger[1], M. Bokranz[1], E. Muth[1] and A. Klein[1]

[1]Molecular Genetics, Dept. of Biology, Philipps University, P.O. Box 1929, D-3550 Marburg (FRG)

[2]Institute of Biochemistry, Genetics and Microbiology, University of Regensburg, Universitätsstr. 31, D-8400 Regensburg (FRG)

SUMMARY

Structural genes of DNA-dependent RNA polymerase and methyl reductase subunits were cloned out of Methanobacterium thermoautotrophicum or Methanococcus voltae. Their arrangements were determined. Both gene groups most likely form transcription units. Striking peptide sequence homologies were detected between an RNA polymerase subunit from the methanogen with corresponding ones from E. coli and yeast. Very high conservation of the polypeptide sequences of two corresponding reductase subunits was found among the two methanogens. Two small open reading frames of so far unknown function were detected in between two reductase genes in Mc. voltae.

INTRODUCTION

Genetic studies on the molecular level can contribute in various ways to the understanding of the biology of a group of physiologically or phylogenetically related organisms.

The analysis of the arrangement of genes involved in a metabolic pathway allows predictions about a possible regulation of their expression. It might also lead to the detection of genes which define previously unknown components of the pathway. The analysis of the gene structures leads to the deduction of primary structures of the gene products and permits the prediction of secondary structures of the proteins, within certain limits. Finally, the analysis of regions flanking structural genes allows the detection of signals used for gene expression. Taken together these approaches help define the genetic basis for the physiological characteristics of the organism.

With bacteria like the methanogens which almost always exist as components of multiorganism consortia the understanding of gene structure and expression mechanisms also gives clues to the possibilities and limits of the exchange of genetic information among members of such bacterial ecosystems. Our interest in methanogens has two aspects. Methanogens are terminal links in anaerobic degradative food chains and therefore play an important ecological as well as biotechnological role. In addition, methanogens are a major group of archaebacteria, a group of procaryotes distinct from both eucaryotes and eubac-

teria. These organisms display characteristic features of the gene expression apparatus (refs. 1-4).

Two principal approaches towards the elucidation of the genetics of an organism are available. Classical genetics use mutants and exploit recombination and complementation studies. Due to the lack of efficient gene transfer systems this approach has not been applicable to methanogens even though mutants of various species are being accumulated (refs. 5-10). The alternative approach is known as reverse genetics. It starts from the gene products and employs gene technology in identifying the corresponding genes as will be detailed below (compare Fig. 1).

We have concentrated our efforts on the elucidation of the structure and function of genes encoding two constitutive multisubunit enzymes, namely DNA-dependent RNA polymerase and methyl reductase of methanogenic bacteria. We hope to gain insight into the structures of the gene clusters, the transcription units and the nature of expression signals. In addition, we compare the primary structures of the proteins with homologous proteins of other organisms. In this way, we expect to get information about structure-function-relationships and about phylogenetic pathways leading to the present state of the genetic constitution of these two groups of functionally related genes.

RESULTS

RNA polymerase genes of methanogenic bacteria

(i) Identification of structural genes of Methanobacterium thermoautotrophicum. DNA-dependent RNA polymerase of Mb. thermoautotrophicum consists of at least 8 subunits (ref. 11). Employing antisera raised against the isolated subunits of RNA polymerase of Mb. thermoautotrophicum (strain Winter) recombinant plasmids could be identified which carried at least fragments of the structural genes of the four largest RNA polymerase subunits A, B', B'' and C. These plasmids are listed in Table 1. The methods used are outlined in Fig. 1. Fig. 2

TABLE 1

pEx31 plasmids (ref. 12) carrying Mb. thermoautotrophicum (strain Winter) RNA polymerase gene fragments

Expressed subunit antigenic determinants	Insert size
A	1.6 kb
A*/C	2.6 kb
B'/B''	1.0 kb
B''*	0.2 kb
B''	0.2 kb

Asterisks indicate that the antigenic determinants are carried on fusion proteins. The cloning procedures have been described in ref. 13.

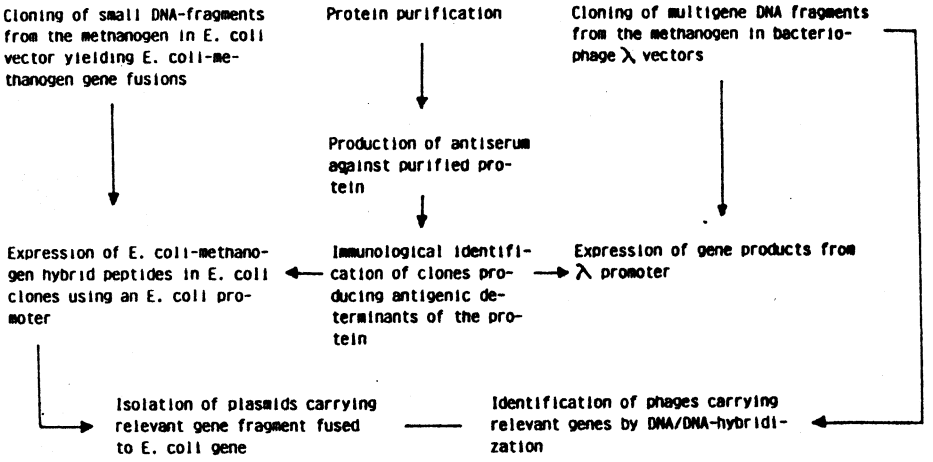

Fig. 1. Flow diagram of methods employed in this study to identify structural genes starting from the known gene products.

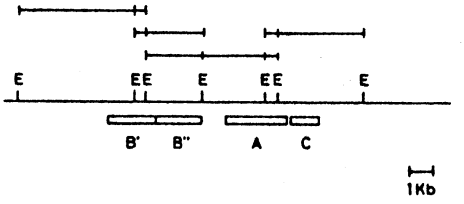

Fig. 2. EcoRI restriction map of the Mb. thermoautotrophicum (strain Winter) DNA segment carrying the RNA polymerase subunit genes A, B', B'' and C. The bars above show inserts of λgt10 vectors, the boxes below indicate the location of the genes as determined by DNA/DNA-hybridization (ref. 14).

shows the relative positions of the four genes, taken from DNA/DNA-hybridization data (ref. 14) of plasmid against DNA from recombinant λ phage taken from a genomic library of Mb. thermoautotrophicum generated in the λ insertion vector λgt10 (ref. 15). The 4 kb EcoRI fragment comprising the 3' end of the A gene was subcloned in the expression vector λgt11 (ref. 16) which allowed to express the complete C gene product (Fig. 3) encoded on the same fragment and thus to confirm both the neighborhood of the A and C genes and their common direction of transcription.

(ii) Arrangement of polymerase genes in other methanogenic bacteria. We wanted to determine the gene order of the identified structural genes for the A, B', B'' and C subunits of the RNA polymerase in other methanogenic bacteria. Experiments performed towards this goal showed that DNA sequence homologies among different members of the group only allowed hybridization between Methano-

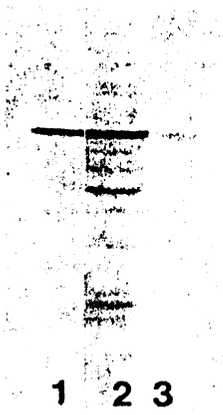

Fig. 3. Identification of vector encoded C subunit polypeptide by an immunoblot technique (compare ref. 13). Lanes, 1: Extract from Mb. thermoautotrophicum (strain Winter); 2: Extract from E. coli containing an expression plasmid carrying a 2.6 kb genomic DNA fragment of Mb. thermoautotrophicum; 3: Protein obtained from λgt11 carrying a 4.0 kb Mb. thermoautotrophicum DNA fragment which overlaps the one contained in the plasmid.

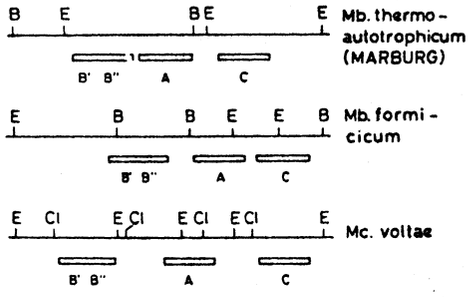

Fig. 4. Gene order of RNA polymerase subunit structural genes in methanogenic bacteria. Genome fragments are shown with relevant restriction sites and the areas indicated to which probes of identified Mb. thermoautotrophicum gene fragments (open bars) of the indicated genes hybridize.
E = EcoRI, B = BamHI, Cl = ClaI restriction sites

bacterium and Methanococcus species. Mb. thermoautotrophicum probes did not hybridize with DNA of Methanosarcina barkeri. The hybridization between Mb. thermoautotrophicum and Mc. voltae is especially remarkable because of the difference in the DNA GC contents of these two members of different orders of methanogenic bacteria (refs. 17, 18). As can be seen from Fig. 4 the gene order of the four polymerase genes mapped is identical in all four organisms in which it could be investigated.

(iii) Possible relationship between RNA polymerase genes among methanogenic and sulfur dependent archaebacteria. Methanogenic bacteria are a subgroup of archaebacteria, which differ from both eubacteria and eucaryotes with respect to various features of the gene expression apparatus including their RNA polymerases (ref. 19).

Since all organisms contain RNA polymerases this enzyme lends itself for studies on the evolution of gene and enzyme structure. It has been found that within the archaebacteria the RNA polymerase structures differ. One interesting difference occurs among the methanogenic bacteria and the sulfur dependent genus Sulfolobus as schematically shown in the insert of Fig. 5b. The largest subunit of the Sulfolobus enzyme has antigenic determinants which appear distributed on two subunits of the methanogen enzymes, namely B' and B''. As shown above, the genes for these two subunits are located next to each other on the chromosome. It appeared interesting to look for the structure of the intergenic region between the B' and B'' genes in Methanobacterium in order to get a clue concerning a possible phylogenetic pathway leading from a polymerase subunit B type found in sulfur dependent archaebacteria to the methanogen B'/B'' subunit arrangement. Sequence determination in between a sequence coding for a B' antigenic determinant and a B'' coding sequence in Mb. thermoautotrophicum (both characterized by antisera after expression on plasmids) showed two open reading frames which overlap. The sequence ATGA includes both the stop codon TGA terminating the B' open reading frame and an ATG codon most likely constituting the start of the B'' gene (Fig. 5a). It is clear from this sequence that the deletion of one base at the overlap would lead to the fusion of the two genes or else these two genes could have originated from one gene by a one base pair insertion (e.g. the A or T of the ATG codon).

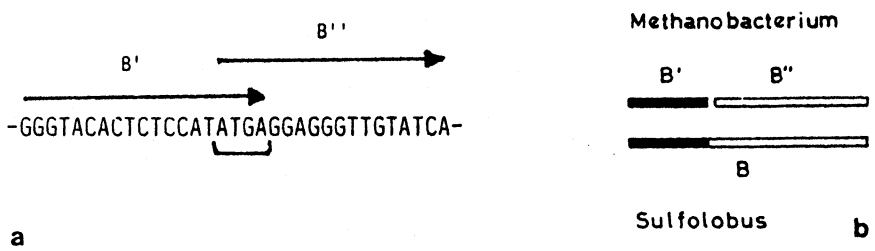

Fig. 5. a) Overlapping open reading frames at the presumptive border of the B' and B'' polymerase subunit genes of Mb. thermoautotrophicum (strain Winter). b) Schematic representation of the gene arrangements in Sulfolobus and Methanobacterium (redrawn from ref. 19). Sequence determination was performed using the chemical cleavage method (ref. 20).

(iv) Structural homologies of RNA polymerases of Mb. thermoautotrophicum, Escherichia coli and Saccharomyces cerevisiae. As mentioned above the archaebacterial RNA polymerases differ from both the eubacterial and eucaryotic enzymes. Yet, a common ancestor is very likely and immunological studies have indeed indicated structural relationships among polymerases of all three groups (ref. 19). We have therefore compared available sequences of E. coli, yeast and Mb. thermoautotrophicum polymerase genes and found an example of striking peptide sequence homology (as derived from the corresponding DNA sequences). This is shown in Fig. 6 and points to a highly conserved domain of the large RNA polymerase subunits of the three organisms, ß', RPO21 and A of E. coli, S. cerevisiae and Mb. thermoautotrophicum, respectively (ref. 21).

```
SC  -Met Gly Gly Arg Glu Gly Leu Ile Asp Thr Ala Val Lys Thr Ala Glu Thr Gly Tyr
      *   *   *   *   *   *   *  (*)  *   *   *  (*) (*)  *   *       (*)  *   *
MT  -Met Gly Gly Arg Glu Gly Leu Val Asp Thr Ala Ile Arg Thr Ala Gln Ser Gly Tyr
          * (*)  *       *   *       *   *   *  (*) (*)  *   *  (*)  *   *   *
EC  -His Gly Ala Arg Lys Gly Leu Ala Asp Thr Ala Leu Lys Thr Ala Asn Ser Gly Tyr

SC   Ile Gln Arg Arg Leu Val Lys Ala Leu Glu Asp Ile Met Val-
       *   *   *   *   *       *   *       *  (*)      *
MT   Met Gln Arg Arg Leu Val Asn Ala Leu Gln Asp Leu Thr Val-
               *   *   *   *                   *   *   *      *
EC   Leu Thr Arg Arg Leu Val Asp Val Ala Gln Asp Leu Val Val-
```

Fig. 6. Sequence homologies of RNA polymerase subunits. The data for E. coli and S. cerevisiae were taken from ref. 21.

The methyl reductase gene group

In contrast to RNA polymerase methyl CoM reductase component C (methyl reductase) is an enzyme found only in methanogenic bacteria. It is the central enzyme in the energy metabolism, catalyzing the terminal step of carbon reduction to methane (ref. 4). This reaction is generally performed at the expense of a methyl group initially bound to an unique coenzyme (coenzyme M). and electrons derived from hydrogen with the help of a hydrogenase as schematically shown in Fig. 7. The enzyme has been purified from different methanogens and shown to contain three subunits, named α, ß and γ (ref. 22). It is a very abundant protein constituting roughly 5% of the total cellular protein in Mc. voltae.

(i) Gene arrangement and Expression. We have previously described that the genes encoding the three known subunits of methyl reductase are arranged in the order ß, γ and α in Mc. voltae (ref. 23). They are located in close proximity to each other. This same gene order of the three genes has also been found in Mb. thermoautotrophicum (Marburg) (Bäumner and Bokranz, unpublished results). The genes probably constitute a transcription unit (ref. 13 and unpublished

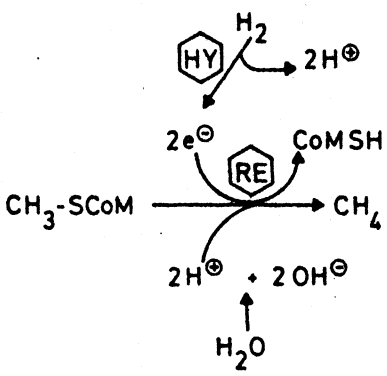

Fig. 7. Schematic representation of the terminal steps of C1-reduction in methanogenic bacteria. HY = hydrogenase, RE = methyl reductase, CoM = coenzyme M (thioethanesulfonate).

observations). The high amount of the enzyme requires very efficient expression of the genes which occur only once in the genome of the cell. This efficient expression will most likely entail coordination of that expression which at first approximation could be accomplished by cotranscription on a polycistronic messenger molecule.

During sequence analysis we have recently detected two small open reading frames intervening between the ß and γ reductase genes (Fig. 8).

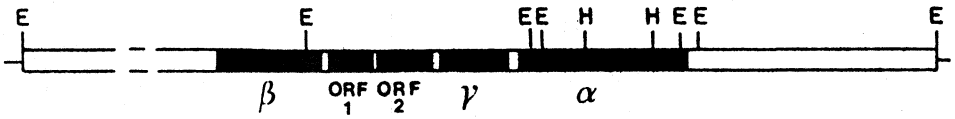

Fig. 8. Arrangement of methyl reductase genes and two interspersed open reading frames of unknown function in Mc. voltae.

Their function is unknown and it will be interesting to see whether similar genes are found in the same position in other methanogens. In a common transcription unit with the three structural genes of the reductase the small open reading frames would be transcribed at the same rate as the reductase genes.

It has been found previously that structural genes in methanogens are preceded by sequences which are complementary to the 16S rRNA 3' end (ref. 24). They are assumed to constitute ribosome binding sites (refs. 25, 26), i.e.

translation signals. Such signals in front of the two reductase genes γ and α and the two open reading frames in between the reductase ß and γ genes are shown together with the complementary 3' terminal sequence of the 16S rRNA in Fig. 9.

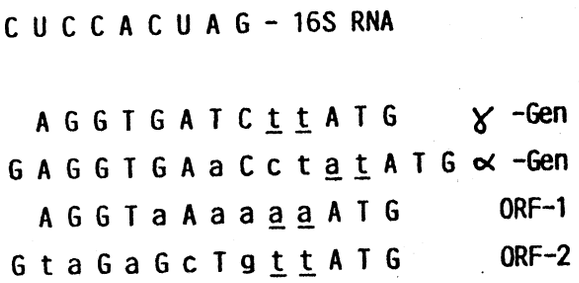

Fig. 9. Putative ribosome binding sites in front of methyl reductase genes and and ORF 1 and 2. The shown RNA sequence is the 3' terminus of methanococcal 16S rRNA (ref. 24).

It can be seen that the complementarity is much weaker with the signals in front of the small open reading frames than with those in front of the reductase genes which might indicate differential expression on the translational level.

(ii) Homology of the methyl reductase α genes of Mc. voltae and Mb. thermoautotrophicum (Marburg). It is clear that the methyl reductases in methanogens are homologous enzymes. Still, comparison of their structures in different methanogens may yield insight into the functional organization of the proteins. It is conceivable that distantly related methanogens have highly conserved sequences only in functionally important domains like cofactor and substrate binding sites and the reactive center.

Nothing is known so far about the structure-function-relationship of the reductase. We have started comparing the primary structures of the enzymes by DNA sequence determination of their structural genes in Mc. voltae (an AT rich mesophilic methanogen) and Mb. thermoautotrophicum which has a higher GC content and is thermophilic. Very high conservation has been found in the first pair of genes looked at, the α genes. The homology is roughly 71% on both the protein and the DNA levels. Base exchanges in the wobble position of the codons allow adaptation to the GC contents of the DNA without changing the resulting polypeptide sequence.

Fig. 10 demonstrates this effect in an extremely conserved region: even con-

servative amino acid exchanges such as serine to threonine at the corresponding position follow that rule.

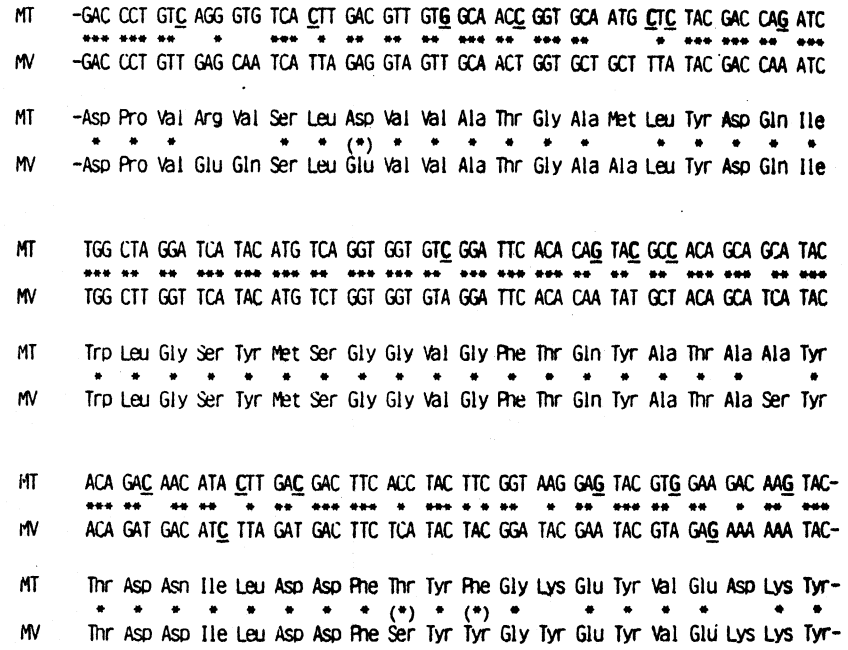

Fig. 10. Sequence homologies of a DNA and corresponding polypeptide section of the methyl reductase ∝ genes from Mc. voltae (MV) and Mb. thermoautotrophicum (Marburg) (MT).

It will be interesting to see whether or not the high conservation of the polypeptide sequence of the ∝ subunit is exceptional compared to the other subunits thus pointing to a central functional role of that subunit.

DISCUSSION

Our studies on the RNA polymerase of Mb. thermoautotrophicum strengthen the hypothesis that RNA polymerases are phylogenetically derived from one precursor in all organisms. This is inferred from the sequence homology between the large subunits of a eubacterial, eucaryotic and archaebacterial gene products (fig. 6).

The arrangements of homologous bacterial RNA polymerase genes differ. We have found a conserved gene order and close proximities of the structural genes

for the largest RNA polymerase subunits in the investigated methanogenic bacteria. These genes are likely to form a transcription unit in Mb. thermoautotrophicum since the direction of transcription is identical for all genes (our unpublished results). Due to the serological relationship between the homologous subunits this arrangement can be compared to the corresponding gene order in E. coli (ref. 27). In this organism only two subunits are expressed from a transcription unit, namely the genes rpoB and rpoC. They code for the subunits ß and ß' corresponding to the B'' and A subunits in the methanogens, respectively. The rpoD gene coding for the σ subunit in E. coli and corresponding to the C gene in the methanogens is not cotranscribed with the rpoB and rpoC genes but located distant from these genes. This may have functional reasons. since σ is a specificity factor and can be replaced by other subunits depending on the physiological state of the cell.

The surprising finding of two small open reading frames in the presumptive methyl reductase transcription unit in Mc. voltae is of special interest to us. The reductase has very low activity in the purified form which is composed of the three mentioned subunits α, ß and γ together with the necessary nickel containing cofactor. The detection of further components might help define a reason for the functional deficiency of the enzyme in vitro. It must be stressed, however, that the organization of genes in a common transcription unit need not mean that their products are parts of a protein complex.

The comparison of the two methyl reductase α genes from Mb. thermoautotrophicum and Mc. voltae has elicited strong homology of the proteins. Most of the base exchanges in the structural genes are consistent with adjustment to the different overall GC contents of the DNA in the two organisms. The concervation of the polypeptide sequences in these organisms might reflect constraints due to essential functional domains. Specific conclusions must, however, await the elucidation and comparison of all enzyme subunits.

REFERENCES

1 C.R. Woese, L.J. Magrum and G.E. Fox, Archaebacteria, J. Mol. Evol., 11 (1978) 245-252.
2 W. Zillig, K.O. Stetter, R. Schnabel, J. Madon and A. Gierl, Transcription in archaebacteria, Zbl. Bakt. Hyg. I. Abt. Orig. C, 3 (1982) 218-227.
3 G. Schmid, T. Pecher and A. Böck, Properties of the translational apparatus of archaebacteria, Zbl. Bakt. Hyg. I. Abt. Orig. C, 3 (1982) 209-217.
4 W.E. Balch, G.E. Fox, L.J. Magrum, C.R. Woese and R.S. Wolfe, Methanogens: Reevaluation of a unique biological group, Microbiol. Rev., 43 (1979) 260-296.
5 A. Kiener, C. Hollinger and T. Leisinger, Analogue-resistant and auxotrophic mutants of Methanobacterium thermoautotrophicum, Arch. Microbiol., 139 (1984) 87-90.
6 J.E. Harris and P.A. Pinn, Bacitracin-resistant mutants of a mesophilic Methanobacterium species, Arch. Microbiol., 143 (1985) 151-153.

7 H. Hummel and A. Böck, Mutations in Methanobacterium formicicum conferring resistance to anti-80S ribosome targeted antibiotics, Mol. Gen. Genet., 198 (1985) 529-533.
8 A. Kiener, T. Rechsteiner and T. Leisinger, Mutation to pseudomonic acid resistance of Methanobacterium thermoautotrophicum leads to an altered isoleucyl-tRNA synthetase, FEMS Microbiology Letters, 33 (1986) 15-18.
9. E.S. Haas, L.A. Hook and J.N. Reeve, Antibiotic resistance caused by permeability changes of the archaebacterium Methanococcus vannielii, FEMS Microbiology Letters, 33 (1986) 185-188.
10 P. Gernhardt and A. Klein, Analogue resistant mutants of Methanococcus voltae, in: O. Kandler (Ed.), Archaebacteria, Gustav Fischer Verlag, Stuttgart, 1986, in press.
11 K.O. Stetter, J. Winter and R. Hartlieb, DNA-dependent RNA-polymerase of the archaebacterium Methanobacterium thermoautotrophicum, Zbl. Bakt. Hyg. I. Abt. Orig. C, 1 (1980) 201-214.
12 K. Strebel, E. Beck, K. Strohmaier and H. Schaller, Characterization of foot and mouth disease virus gene products with antisera against bacterially synthesized fusion proteins, J. Virol., 57 (1986) 983-991.
13 U. Konheiser, G. Pasti, C. Böllschweiler and A. Klein, Physical mapping of genes coding for two subunits of methyl CoM reductase component C of Methanococcus voltae, Mol. Gen. Genet., 198 (1984) 146-152.
14 E.M. Southern, Detection of specific sequences among DNA fragments, J. Mol. Biol., 98 (1975) 503-517.
15 T.V. Huynh, R.A. Young and R.W. Davis, Construction and screening cDNA libraries in λgt10 and λgt11, in: D.M. Glover (Ed.), DNA cloning a practical approach, Vol. I, IRL Press, Oxford, 1985, pp. 49-78.
16 R.D. Young and W.D. Davis, Efficient isolation of genes using antibody probes, Proc. Natl. Acad. Sci. USA, 80 (1983) 1194-1198.
17 R.M. Mitchell, L.A. Loeblich, L.C. Klotz and A.R. Loeblich, DNA organization of Methanobacterium thermoautotrophicum, Science, 204 (1979) 1082-1084.
18 A. Klein and M. Schnorr, Genome complexity of methanogenic bacteria, J. Bacteriol., 158 (1984) 628-631.
19 W. Zillig, R. Schnabel, K. Stetter, M. Thomm, F. Gropp and W.D. Reiter, The evolution of the transcription apparatus, in: K.H. Schleifer and E. Stackebrandt (eds.), The evolution of the prokaryotes, Academic Press, New York, 1985, pp. 45-72.
20 A.M. Maxam and W. Gilbert, Sequencing end-labeled DNA with base-specific chemical cleavage, Methods in Enzymology, 65 (1980) 499-559.
21 L.A. Allison, M. Moyle, M. Shales and C.J. Ingles, Extensive homology among the largest subunits of eukaryotic and prokaryotic RNA polymerases. Cell, 42 (1985) 599-610.
22 P.L. Hartzell and R.S. Wolfe, Comparative studies of component C from the methylreductase system of different methanogens, System. Appl. Microbiol., 7 (1986) 376-382.
23 R. Allmansberger, C. Bollschweiler, U. Konheiser, B. Müller, E. Muth. G. Pasti and A. Klein, Arrangement and expression of methyl CoM reductase genes in Methanococcus voltae, System. Appl. Microbiol., 7 (1986) 13-17.
24 M. Jarsch and A. Böck, DNA sequence of the 16S rRNA/23S rRNA intercistronic spacer of two rDNA operons of the archaebacterium Methanococcus vannielii. Nucl. Acids Res., 21 (1983) 7537-7544.
25 C. Bollschweiler, R. Kühn and A. Klein, Non-repetitive AT-rich sequences are found in intergenic regions of Methanococcus voltae DNA, EMBO J. 4 (1985) 805-809.
26 D. Cue, G.S. Beckler, J.N. Reeve and J. Konisky, Structure and sequence divergence of two archaebacterial genes, Proc. Natl. Acad. Sci. USA, 82 (1985) 4207-4211.
27 H. Matzura, Regulation of biosynthesis of the DNA-dependent RNA polymerase in Escherichia coli, Curr. Top. Cell. Regulation, 17 (1980) 89-136.

We thank H. Bestgen and H. Steinebach for their competent assistance.
This work was supported by the Deutsche Forschungsgemeinschaft.