J. Mol. Biol. (1991) **222**, 495–508

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*

Winfried Hausner, Gerhard Frey and Michael Thomm†

Lehrstuhl für Mikrobiologie, Universität Regensburg Universitätsstr. 31 D-8400 Regensburg, Germany

(Received 19 December 1990; accepted 13 August 1991)

To identify the DNA sequences required for initiation of transcription in archaea, the 5′-flanking region of the $tRNA^{Val}$ gene of *Methanococcus vannielii* was modified by deletions, restructuring and site-directed mutagenesis, and the tRNA encoding sequence was replaced by a fortuitous *Escherichia coli* sequence. The effects of these mutations on promoter function were tested in an homologous cell-free transcription system. The DNA region from position −35 to +9 relative to the transcription start site was sufficient for maximal initiation of cell-free transcription. Removal of the DNA region between −35 and −30 reduced initiation by a factor of 2. Deletions extending to position −24 almost completely abolished specific transcription. Analysis of 16 site-specific mutations in the region from −33 to +2 provided evidence that a conserved A+T-rich sequence (TATA box), centered at −25, is essential for initiation of transcription. Single point mutations in six positions of the TATA box reduced initiation of transcription from 0·2 to 0·01 of wild-type levels. A second conserved motif at the transcription start site (consensus ATGC) could be replaced by some sequences containing a pyrimidine-purine dinucleotide but appeared necessary for a maximal rate of gene transcription. Mutations altering the spacing between the two conserved elements demonstrated that initiation occurs at a strictly defined distance of 22 to 27 base-pairs downstream from the TATA box. Our results support the conclusion that the TATA box is the major DNA region mediating promoter recognition, influencing the efficiency of transcription and specifying the site of transcription initiation. This *Methanococcus* promoter element closely resembles in structure and function the TATA box of promoters of eukaryotic protein-encoding genes transcribed by RNA polymerase II.

Keywords: archaea; TATA box; promoter; cell-free transcription; site-directed mutagenesis

1. Introduction

Although the DNA elements promoting gene transcription in bacteria and higher organisms have been extensively characterized, their counterparts in the domain of archaea (formerly archaebacteria; Woese *et al.*, 1990) are poorly understood. Comparison of DNA sequences in the upstream region of a large variety of archaeal genes has revealed the presence of conserved elements: Wich *et al.* (1986) identified two sequences upstream from stable RNA genes of the genus *Methanococcus*, the box A motif extending from position −40 to −20 relative to the transcription initiation site and box B at the transcription start site. Statistical analyses of further genes revealed that a part of the box A sequence (TATA box) and a modified box B version are conserved in the 5′ flanking regions of most constitutively expressed archaeal genes (Thomm & Wich, 1988; Reiter *et al.*, 1988; Brown *et al.*, 1989). However, no evidence is available demonstrating the importance of these conserved sequences in promoter function.

One possibility of defining a promoter in a functional assay is to mutagenize the putative recognition sequence and to test the effect of the alteration of the template on the rate of cell-free transcription (for a review, see Breathnach & Chambon 1981; Siebenlist *et al.*, 1980). Recently, *in vitro* transcrip-

† Author to whom all correspondence should be addressed at: Institut für Allgemeine Mikrobiologie, Christian-Albrechts-Universität Kiel, Olshausenstr. 40, 2300 Kiel, Germany.

0022-2836/91/230495-14 $03.00/0

tion systems have been developed that synthesize specific transcripts from tRNA genes of *Methanococcus vannielii* (Frey *et al.*, 1990), rRNA genes of *Sulfolobus shibatae* (Hüdepohl *et al.*, 1990) and a protein-encoding gene of *Methanobacterium thermoautotrophicum* (Knaub & Klein, 1990). The *Methanococcus* and *Sulfolobus* systems initiate at the same site *in vitro* as *in vivo* at a distance of approximately 25 bp† downstream from the TATA box. The ability of these enzymes to utilize the correct start site was shown to be strictly dependent upon transcription factors (Frey *et al.*, 1990; Hüdepohl *et al.*, 1990). In contrast, the *Methanobacterium* enzyme initiates at a distance of five nucleotides downstream from the TATA box, 21 bp upstream from the *in vivo* initiation site (Knaub & Klein, 1990). These cell-free systems were incapable of specific initiation when the complete 5′-flanking sequences approximately upstream from position −20 had been deleted from these templates, suggesting that the promoter signals of archaeal genes reside in the DNA region upstream from the transcription start site (Thomm *et al.*, 1990; Hüdepohl *et al.*, 1990; Knaub & Klein, 1990). However, the DNA sequences mediating initiation of transcription have not been identified.

Here, we used a homologous cell-free transcription system to identify the DNA sequences required for transcription of the $tRNA^{Val}$ gene of *M. vannielii*. The 5′-flanking region of this gene contains the sequence 5′-TTTATATA-3′ (consensus TTTA T/A ATA; Thomm & Wich, 1988) between position −30 and −24 and 5′-TTGC-3′ (consensus ATGC; Thomm *et al.*, 1989) at the transcription start site (GTP is the initiator nucleotide; Frey *et al.*, 1990). We report on experiments locating the 5′ boundary of this promoter and demonstrate that internal sequences are not required for transcription initiation. Furthermore, we analyze the effects of single point mutations in the DNA region from −33 to +2 on the rate of cell-free transcription. From these studies we propose a general model for the structural determinants of an archaeal promoter.

2. Materials and Methods

(a) *Bacterial strains and plasmids*

M. vannielii was grown as described (Frey *et al.*, 1990). *E. coli*, strain JM109 (Yanisch-Perron *et al.*, 1985) was used for amplification of plasmids. The plasmids pIC-19H and pIC-20H (Marsh *et al.*, 1984) were used as cloning vectors for template constructions.

(b) *Enzymes and oligonucleotides*

Restriction endonucleases and other DNA enzymes were purchased from Boehringer-Mannheim or New England Biolabs. Radioactive nucleotides were from Dupont, NEN Research Products. The synthetic oligonucleotides used for the construction of promoter mutants (Table 1) were synthesized on a gene assembler (Applied Biosystems).

† Abbreviation used: bp, base-pair(s).

(c) *Preparation of extracts for* in vitro *transcription*

An S-100 extract from 2 g cells of *M. vannielii* was prepared as described (Frey *et al.*, 1990) and applied to a 30 ml phosphocellulose column (2 cm × 10 cm) equilibrated with TK buffer (50 mM-Tris·HCl (pH 8.0), 100 mM-KCl, 20% (v/v glycerol). The column was washed with 3 vol. TK buffer. The adsorbed proteins were eluted in 3 steps with TK buffer containing 0·35 M, 0·6 M and 1·0 M-KCl. One transcription factor activity was eluted at 1·0 M-KCl (Frey *et al.*, 1990). The RNA polymerase was purified from the S-100 extract by DEAE-cellulose chromatography. An extract of 2 g cells was applied to a 30 ml DEAE-cellulose column (2 cm × 10 cm) equilibrated with TMK buffer (50 mM-Tris·HCl, (pH 8·0), 10 mM-$MgCl_2$, 50 mM-KCl, 20% (v/v) glycerol). After washing the column with 3 vol. TMK buffer, the RNA polymerase activity was eluted with 0·35 M-KCl.

(d) *Transcriptional analysis*

RNA polymerase (20 μl DEAE fraction) and transcription factor (20 μl phosphocellulose eluate) were incubated in transcription reactions under the conditions described (Frey *et al.*, 1990). The DNA concentration was 30 μg/ml and 2 μg/ml when circular plasmids or purified restriction fragments were used as template, respectively. After electrophoresis, the RNA products were quantified by liquid scintillation counting of excised radioactive transcripts. The experiments were repeated at least 4 times and the percentage mean activity was determined relative to the clone pIC31/1 harboring the wild-type DNA sequence up to position −590, or to the construct pIC31/30 (see section (h), below).

(e) *Determination of the initiatior nucleotide by primer extension*

Primer extension experiments were done as described (Frey *et al.*, 1990). The primer used is shown in Fig. 1(a).

(f) S_1 *mapping*

Unlabeled *in vitro* RNA from one cell-free transcription reaction was purified by phenol treatment and precipitated with ethanol. The nucleic acids were resuspended in 50 μl of buffer containing 50 mM-Tris·HCl, (pH 7·5), 10 mM-$MgCl_2$, 0·1 mM-dithiothreitol. After addition of 25 units of DNase I (RNase free; Boehringer) the template was hydrolyzed for 30 min at 37 °C. The RNA was purified again by extraction with phenol and chloroform and stored at −20 °C. S_1 mapping was performed by a modification of the procedure described by Wich *et al.* (1986). Appropriate amounts of RNA (see Fig. 6(d)) and the corresponding end-labeled probes (30,000 cts/min; see Fig. 6(a)) were denatured for 10 min at 93 °C in a total volume of 30 μl of hybridization buffer (50 mM-Tris·HCl (pH 7·5), 200 mM-KCl, 2·5 mM-EDTA). Hybridization was carried out for 3 h at 65 °C. After addition of 300 μl of S_1 buffer (280 mM-NaCl, 50 mM-sodium acetate (pH 4·6), 4·5 mM-$ZnSO_4$, 20 μg single-stranded carrier DNA/ml) and 100 units of nuclease S_1, the reaction mixture was incubated for 1 h at 30 °C. The reaction was stopped by the addition of 40 μl of 4 M-LiCl. S_1-resistant hybrids were recovered by precipitation with ethanol and analyzed by polyacrylamide/urea gel electrophoresis.

(g) *Construction of templates*

The 721 bp *Cla*I fragment of plasmid pMT31 (Wich *et al.*, 1986) harboring a $tRNA^{Val}$ gene of *M. vannielii* was

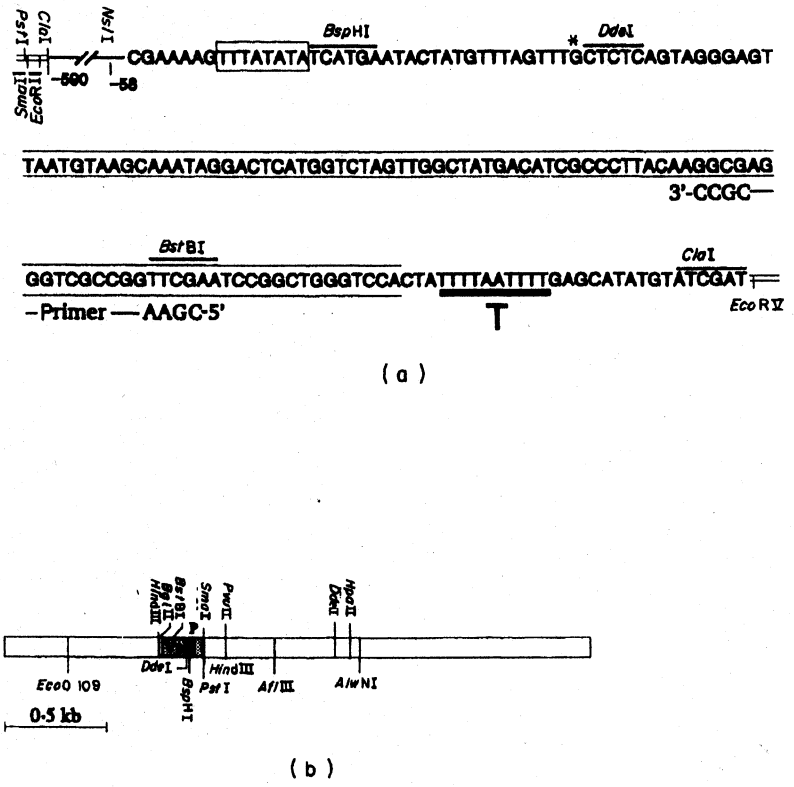

Figure 1. DNA sequence of the $tRNA^{Val}$ gene of *M. vannielii* and restriction map of plasmid pIC31/1 and pIC31/2. (a) The sequence from −37 to 16 bp downstream from the transcription termination site is shown. DNA regions encoding the mature $tRNA^{Val}$ are marked by 2 parallel bars above and below the sequence. The consensus promoter octanucleotide is boxed and the transcription start site labeled by an asterisk. One putative terminator element (Wich *et al.*, 1986) is underlined and indicated by a T. The synthetic oligonucleotide used for start site mapping experiments is indicated below the complementary DNA sequence. Two parallel bars in close vicinity at the 5′ and 3′ ends of the gene indicate sequences of the polylinker site of plasmid pIC-19H (Marsh *et al.*, 1984). Some restriction sites relevant for the construction of mutants are labeled by a heavy line above the *Methanococcus* DNA sequence. The DNA region from −590 to −58 (bold line) does not contain a promoter detectable by *in vitro* transcription. (b) Partial restriction site map of plasmid pIC31/2. The orientation of the $tRNA^{Val}$ gene of *M. vannielii* inserted in the polylinker site of plasmid pIC-19H is shown. Two parallel bars indicate vector sequences, the shaded area the $tRNA^{Val}$ gene. The $tRNA^{Val}$ gene promoter (P) is indicated by a filled bar. Some restriction sites used for the constructions of promoter mutants are shown.

inserted into the *Cla*I site of the cloning vector pIC-19H (Marsh *et al.*, 1984). The resulting plasmid pIC31/1 (Figs 1(a) and 2) harbors the wild-type *Methanococcus* DNA sequence from position −590 to +131 relative to the transcription initiation site of the $tRNA^{Val}$ gene. To remove the upstream region from −590 to −58, plasmid pIC31/1 was digested with the restriction endonucleases *Eco*RI and *Nsi*I (Fig. 1(a)). The purified large fragment of 2886 bp was treated with the Klenow fragment of DNA polymerase I and circularized by ligation. The resulting plasmid pIC31/2 contains the DNA region from −58 to +131 (Figs 1(b) and 2).

To define the 5′ end of the promoter region more precisely, a set of unidirectional deletion clones was prepared. After digestion of plasmid pIC31/2 with *Pst*I and *Sma*I (Fig. 1(b)), the large fragment of 2869 bp was purified and treated with exonuclease III and nuclease S_1 according to the procedure described by Henikoff (1984). The 5′ end of the *Methanococcus* DNA in various deletion clones was determined by dideoxy sequencing (Sanger *et al.*, 1977). The sequences of 4 clones selected for this study, pIC31/8, pIC31/9, pIC31/10 and pIC31/11, are shown in Fig. 2.

To replace the tRNA coding region by a random sequence, a 2349 bp *Afl*III–*Bst*BI fragment of pIC31/2 (Fig. 1(b)) containing most of the vector sequence and the terminator region of the $tRNA^{Val}$ gene, was purified. This fragment was ligated with a *Dde*I–*Afl*III fragment of pIC31/2 harboring the DNA sequence from position −58 to +5, and a 78 bp *Dde*I–*Hpa*II fragment from the vector pIC-19H (see Fig. 6(a)). In the resulting clone (pIC31/61) the tRNA-encoding DNA sequence from position +9 to +82 is replaced by a non-coding sequence of *E. coli* DNA (Fig. 6(a)).

(h) *Construction of single-point mutations in the promoter region*

Point mutations of the $tRNA^{Val}$ gene were constructed using complementary oligonucleotides as adaptors. A 589 bp *Eco*O109I–*Bsp*HI fragment containing the tRNA gene without the promoter and a small part of the vector were isolated from pIC31/2 (Fig. 1(b)). The second part of the vector was isolated in a 2233 bp *Pst*I–*Eco*O109I fragment from pIC31/2 (Fig. 1(b)). In a ligation reaction equivalent amounts of molecules of both fragments were incubated with the oligonucleotides 31/30A and 31/30B (Table 1). Usually, oligonucleotides were added in a 100-fold excess. The 2 oligonucleotides hybridized in the reaction and the overhanging ends of the adaptor are on one

Table 1
List of oligonucleotides used for the construction of single point mutations in the promoter region of the $tRNA^{Val}$ gene

Oligonucleotides used (5′ to 3′)	Resulting plasmid
31/30A: AAAAGTTTATATAT 31/30B: CATGATATATAAACTTTTTGCA	pIC31/30
31/51A: AAGAGTTTATATAT 31/51B: CATGATATATAAACTCTTTGCA	pIC31/51
31/50A: AAAATTTTATATAT 31/50B: CATGATATATAAAATTTTTGCA	pIC31/50
31/31A: AAAAGGTTATATAT 31/31B: CATGATATATAACCTTTTTGCA	pIC31/31
31/32A: AAAAGTGTATATAT 31/32B: CATGATATATACACTTTTTGCA	pIC31/32
31/35A: AAAAGGGTATATAT 31/35B: CATGATATATACCCTTTTTGCA	pIC31/35
31/33A: AAAAGTTGATATAT 31/33B: CATGATATATCAACTTTTTGCA	pIC31/33
31/43A: AAAAGTTTGTATAT 31/43B: CATGATATACAAACTTTTTGCA	pIC31/43
31/44A: AAAAGTTTAGATAT 31/44B: CATGATATCTAAACTTTTTGCA	pIC31/44
31/45A:P AAAAGTTTATGTAT 31/45B: CATGATACATAAACTTTTTGCA	pIC31/45
31/34A: AAAAGTTTATAGAT 31/34B: CATGATCTATAAACTTTTTGCA	pIC31/34
31/46A: AAAAGTTTATATGT 31/46B: CATGACATATAAACTTTTTGCA	pIC31/46
31/47A: AAAAGTTTATATAG 31/47B: CATGCTATATAAACTTTTTGCA	pIC31/47
31/48A: AAAAGTTTATATATCGTGAATACTATGTTTAGTTTGCTC 31/48B: TGAGAGCAAACTAAACATAGTATTCACGATATATAAACTTTTTGCA	pIC31/48
31/49A: CATGGATACTATGTTTAGTTTGCTC 31/49B: TGAGAGCAAACTAAACATAGTATC	PIC31/49
31/57A: CATGAATACTATGTTTAGTTTGATC 31/57B: TGAGATCAAACTAAACATAGTATT	pIC31/57
31/53A: CATGAATACTATGTTTAGTTTTCTC 31/53B: TGAGAGAAAACTAAACATAGTATT	pIC31/53
31/56A: CATGAATACTATGTTTAGTTAGCTC 31/56B: TGAGAGCTAACTAAACATAGTATT	pIC31/56
31/59A: AAAAGTTTATATAAATCTAGAATTCTATAGTTATTGCTC 31/59B: TGAGAGCAATAACTATAGAATTCTAGATTTATATAAACTTTTTGCA	pIC31/59
31/54A: CATGAATATACTCTATGTTTAGTTTGCTC 31/54B: TGAGAGCAAACTAAACATAGAGTATATT	pIC31/54
31/55A: CATGAATATACTTACTATGTTTAGTTTGCTC 31/55B: TGAGAGCAAACTAAACATAGTAAGTATATT	pIC31/55
31/60A: CATGAAATGTTTAGTTTGCTC 31/60B: TGAGAGCAAACTAAACATTT	pIC31/60

side complementary to the 3′ extension of the *Pst*I site of the *Pst*I–*Eco*O109I fragment. On the opposite side of the fragment the ends are compatible to the extension of the *Bsp*HI site of the second fragment. Therefore, an intact plasmid could be generated by the ligation of 2 larger fragments with 2 double stranded synthetic oligonucleotides. The ligated plasmids were transformed into *E. coli* JM109. The resulting plasmid pIC31/30 contains the wild-type *Methanococcus* DNA sequence from −35 to +131 and served as a wild-type reference (Fig. 3). The clones pIC31/51, pIC31/50, pIC31/31, pIC31/32, pIC31/35, pIC31/33, pIC31/43, pIC31/44, pIC31/45, pIC31/34, pIC31/46 and pIC31/47 (Figs 3 and 4) were constructed in the same manner by using the corresponding oligonucleotides indicated in Table 1. The sequences of all clones containing single point mutations were verified by dideoxy sequencing.

To create clones pIC31/48 and pIC31/59 (Figs 4 and 5) a 2650 bp *Pst*I–*Bgl*II fragment and a 150 bp *Dde*I–*Bgl*II fragment were isolated from pIC31/2 (Fig. 1(b)). The large fragment served as vector and the small fragment contained the $tRNA^{Val}$ gene from position +4 to the 3′ end. These fragments were ligated with oligonucleotides 31/48A, 31/48B or 31/59A, 31/59B to construct plasmid 31/48 and 31/59, respectively. The ends of the hybridized oligonucleotides are compatible with the *Pst*I and *Dde*I restriction site extensions.

For constructing plasmid pIC31/49 (Fig. 4), 3 DNA restriction fragments were purified. The 152 bp *Dde*I–*Hin*dIII fragment containing the $tRNA^{Val}$ gene without the promoter and the 1865 bp *Hin*dIII–*Alw*N I fragment harboring about 2/3 of the vector were isolated from pIC31/2 (Fig. 1(b)). An *Alw*NI–*Bsp*HI fragment containing the *Methanococcus* promoter DNA region from −35 to −20 and the residual third of the vector was purified from pIC31/30. Equivalent amounts of molecules of the 3 fragments were incubated in a ligation reaction in the presence of oligonucleotides 31/49A and 31/49B (Table 1). The ends of the hybridized oligonucleotides are compatible with the *Bsp*HI and *Dde*I restriction site extensions. After transforming into *E. coli* JM109, the introduction of single point mutations was verified by dideoxy sequencing. The clones pIC31/57, pIC31/53, pIC31/56, pIC31/54, pIC31/55 and pIC31/60 (Figs 4 and 5) were constructed in the same manner by using the corresponding oligonucleotides (Table 1).

Plasmid DNA was purified by repeated centrifugation in CsCl density gradients as described (Thomm & Wich, 1988).

3. Results

(a) *Location of the 5′ end of the promoter*

The cell-free transcription system used for this study was reconstituted from two protein fractions purified from a soluble extract (S-100) of *M. vannielii* by phosphocellulose and DEAE-cellulose chromatography (Materials and Methods). When plasmid pIC31/1 harboring the wild-type DNA sequence up to position −590 upstream from the transcription initiation site was used as a template in an *in vitro* transcription reaction an RNA product of 117 nucleotides was observed (Fig. 2, lane 2). This is the size expected for a transcript corresponding to the pre-tRNA synthesized in *Methanococcus* cells (Wich *et al.*, 1986). No specific transcript could be detected when the vector DNA

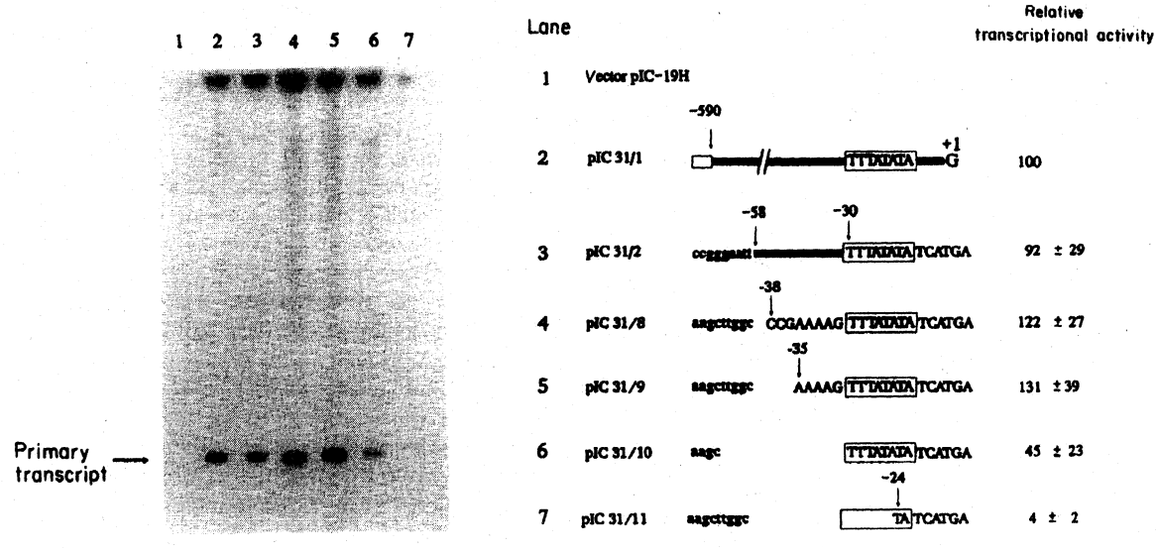

Figure 2. Localization of the upstream boundary of the $tRNA^{Val}$ gene promoter. The DNA upstream regions of the various 5′ deletion clones and their template activities are shown. The filled bar represents *Methanococcus* wild-type DNA and 2 parallel bars, vector DNA. The 5′ ends of the various deletion clones are labeled by arrows. Lower case letters indicate vector DNA sequences, upper case letters *Methanococcus* DNA. The position of the first nucleotide of the TATA box is indicated in lane 3. The other symbols are as for Fig. 1. The relative transcriptional activities of the various clones were determined after polyacrylamide/urea gel electrophoresis of *in vitro* RNA (Frey *et al.*, 1990). The labeled RNA products were excized from the gel and quantitated by Čerenkov counting. The numbers on top of the gel correspond to the numbers designating the corresponding DNA sequences at the right side of the Figure. The amount of radioactivity not entering the gel is proportional to the amount of specific transcripts. This high molecular weight RNA was absent when run-off transcripts were analyzed (data not shown). Therefore, these RNA products appear to be caused mainly by read-through at the primary terminator.

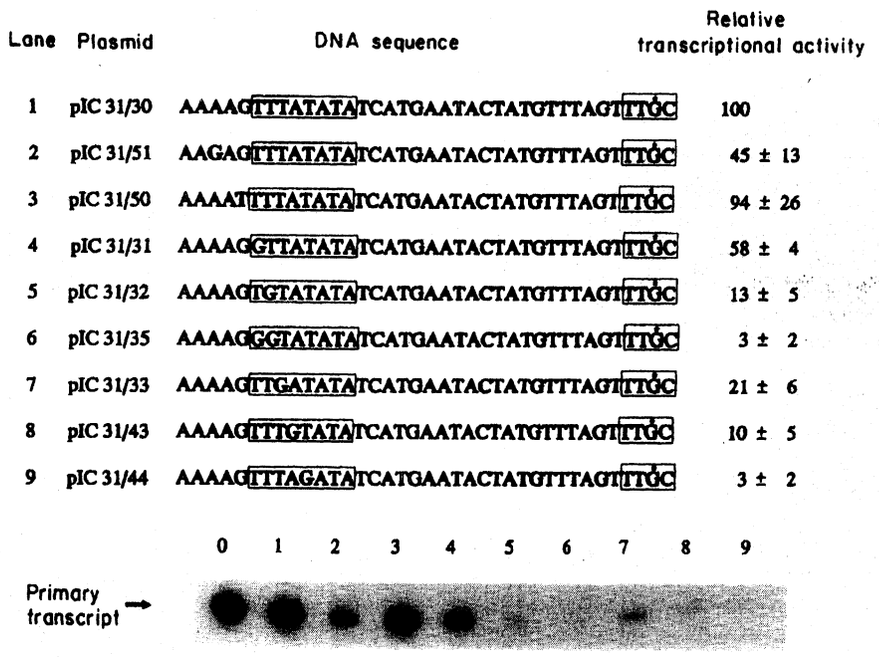

Figure 3. Effects of site-specific mutations in the DNA region from −33 to −26 on promoter function. The wild-type DNA sequence of the promoter region of the $tRNA^{Val}$ gene between −35 and +2 is shown (lane 1). The various single point mutations and the double mutation (lane 6) are indicated by bold characters. The TATA box and the initiator element are boxed. The transcriptional efficiencies are indicated on the right. A gel electrophoretic analysis of the RNA products from the corresponding templates is shown below the sequence. Only the section of the gel relevant for the quantitation of transcripts is shown. Lane 0 shows the transcripts from pIC31/2 containing the wild-type sequence up to position −58, for comparison.

not containing a tRNA gene was employed as a template (Fig. 2, lane 1). A primer extension experiment (Fig. 5(b), left side, lane 1) and sequence analysis (data not shown) of *in vitro* transcripts from pIC31/1 revealed that the RNA initiates at the same site as *in vivo*. These results indicate that this reconstituted *Methanococcus* system initiates and terminates transcription faithfully *in vitro* as well as the RNA polymerase fraction purified from the S-100 extract by glycerol gradient centrifugation (Frey *et al.*, 1990).

To identify the DNA region required for initiation of transcription in a methanogen, a series of deletions was introduced into the DNA 5′ to the $tRNA^{Val}$ gene of *M. vannielii* by cleavage of plasmid pIC31/1 with restriction enzymes (Fig. 2, lanes 2 and 3) or digestion with exonuclease III (Fig. 2, lanes 4 to 7). The deletion mutants were ligated at the 5′ end with DNA sequences originating from the vector. The template activity of circular plasmids was determined relative to the activity of pIC31/1 in cell-free transcription experiments. When the wild-type DNA between position −590 and −58 was deleted the resulting plasmid was transcribed with a similar efficiency as pIC31/1 (Fig. 2, lanes 2 and 3). This finding indicates that the region upstream from position −58 is not essential for initiation of transcription. After removal of the sequence between −58 and −35 the efficiency of tRNA expression was slightly increased (lanes 4 and 5), suggesting that this region contains a negative regulatory element. The corresponding plasmids pIC31/2 and pIC31/8 differ also in the vector sequence 5′ to the *Methanococcus* DNA (Fig. 2, lanes 3 and 4). Therefore, we cannot exclude the possibility that the increase in promoter activity is caused by some stimulatory effect of the substituting sequence. A 5′ deletion extending to position −30 retained levels of specific transcription that were about 50% of wild-type levels (lane 6). An almost complete inactivation of transcription was observed when six (lane 7) or seven nucleotides (data not shown) of the TATA box were removed. Mutant DNAs with deletions extending to position +3 were also inactive (data not shown). Thus, the sequences between −30 and −24 appears to be essential for accurate initiation of transcription, suggesting that the TATA box is a major element of the *Methanococcus* promoter. Although not absolutely required for cell-free transcription, the DNA sequence from −35 to −30 also contributes to promoter function.

(b) *Analysis of single point mutations in the promoter region*

To provide unequivocal evidence for the significance of the TATA box as a promoter signal and to define the internal organization of the $tRNA^{Val}$ promoter, we introduced 16 single point mutations in the DNA region between −33 and +2 (Figs 3 and 4). Unless otherwise indicated, an A or T

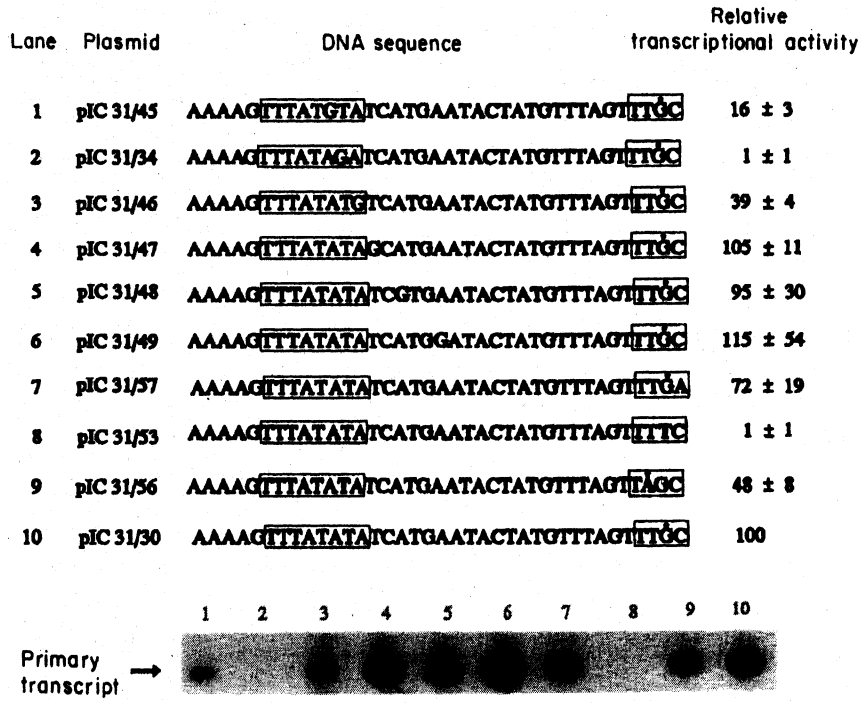

Figure 4. Effects of site-specific mutations in the DNA region from −25 to +2 on promoter function. The location of single point mutations is indicated and the transcriptional activities of the various site-specific mutants were determined as indicated for Fig. 3.

residue of the wild-type sequence was replaced by a G. To facilitate the introduction of this alteration, the wild-type DNA sequence from position −35 to +3 was synthesized by the use of oligonucleotides and ligated to the tRNA encoding region (see Materials and Methods). The transcripts of the resulting clone (pIC31/30) were initiated with comparable efficiency to that for those from pIC31/2 harboring the wild-type upstream region extending to position −58 (compare lanes 0 and 1 in Fig. 3). For purposes of this study we refer to pIC31/30 as wild-type. The various clones indicated in Figures 3 and 4 differ from the wild-type sequence in the region from −35 to position +131 relative to the transcription initiation site only in a single base-pair.

When the A residue at position −33 was substituted by a G residue, the rate of transcription was reduced to 45% (Fig. 3, lane 2), confirming that the DNA region directly adjacent to the 5′ end of the TATA box also augments the initiation process. The replacement of the G by a T, at position −31 and mutations of some positions downstream from the TATA box did not lead to a pronounced effect on the promoter activity of the template (Fig. 3, lane 3; Fig. 4, lanes 4 to 6). A mutation of the first and last positions of the TATA box octanucleotide decreased the tRNA expression rate only to about 60% and 40% of wild-type levels (Fig. 3, lane 4; Fig. 4, lane 3). However, single point mutations in all other positions of the TATA box dramatically reduced or completely abolished the activity of this promoter (Fig. 3, lanes 5 and 7 to 9; Fig. 4, lanes 1 and 2). A double mutation of both the first and second positions of the TATA box reduced the promoter activity to about 0·01 of wild-type levels (Fig. 3, lane 6). The T → G transversions at position 5 and 7 of the TATA box had the most severe effect on promoter function (no significant residual activity). However, the A → G transition at position 4 led to a stronger reduction in the transcription rate (residual activity 10%) than the T → G transversion at position 3 (20%). This finding suggests that the various positions of the TATA box differ in their contribution to promoter strength. The effect exerted by the type of substituting base-pair seems to be less important.

To investigate the role of the second conserved element (ATGC) at the transcription start site (Wich *et al.*, 1986; Reiter *et al.*, 1988; Thomm *et al.*, 1989) in promoter function, three positions of this element were altered. When the initiator nucleotide G was replaced by a T the template activity was completely abolished (Fig. 4, lane 8). Substitution of the C at position +2 and of the T at position −1 by an A residue resulted in a milder reduction of the rate of transcription. The remaining activities were 70% and 50% of wild-type levels, respectively (Fig. 4, lanes 7 and 9). The initiator nucleotides of transcripts from pIC31/57 and pIC31/56 were identified by a primer extension experiment (Fig. 5(b), right part). Transcription initiates at the G of the TTGA sequence on pIC31/57 (lane 7) and at the A of the TAGC sequence on pIC31/56 (lane 6). This finding suggests that initiation of transcription preferably occurs at a purine preceded by a pyrimidine residue.

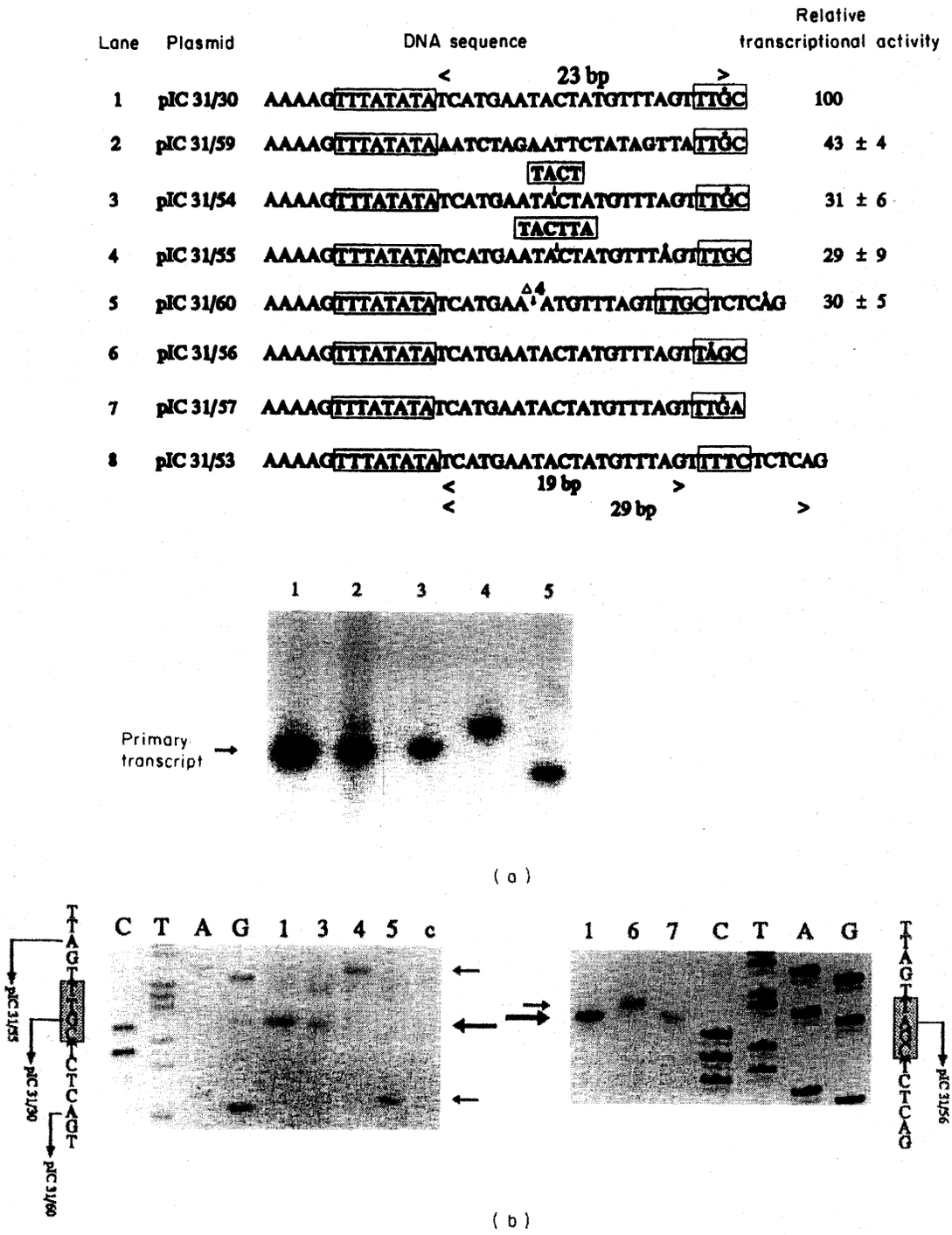

Figure 5. Effects of mutations in the spacer region between the TATA box and the transcription start site on the efficiency of cell-free transcription and start site selection. (a) Clones with 14 single point mutations (lane 2), insertions of 4 (lane 3) and 6 bp (lane 4) and a deletion of 4 bp (lane 5) were transcribed *in vitro* and the template activities compared with transcripts from the wild-type DNA sequence (lane 1). The inserted nucleotides are shown boxed above the sequence; the site of insertion is labeled by an arrow. Δ4 designates a deletion of 4 nucleotides. Line 8. extended DNA sequence of the initiator nucleotide point mutant pIC31/53 analyzed in Fig. 4, lane 8. The distance from the purine residues next to the initiator nucleotide to the TATA box is designated to illustrate that the promoter-down phenotype of this mutant is probably due to the lack of purines at the correct distance to the TATA box. (b) Mapping of the 5′ end of transcripts by primer extension analysis. Unlabeled *in vitro* RNA transcribed from plasmid pIC31/30, the spacer mutants indicated in lanes 3 to 5 of (a) and the single-point mutants of the initiator element (lanes 6 and 7) were annealed with a single-stranded 5′ end-labeled primer (Fig. 1(a)) and extended with reverse transcriptase under the conditions described (Frey *et al.*, 1990). Left part: analysis of spacer mutants; right part: analysis of single point mutants. The lane labeled c is a control reaction without ribonucleoside triphosphates during the transcription reaction. The 5′ end of the reverse transcriptase products from the wild-type, pIC31/54 and pIC31/57 are labeled by a large arrow and those from the mutants pIC31/55, pIC31/60 and pIC31/56 by small arrows. The same primer was used for the DNA sequencing reactions shown as standard. C, T, A and G indicate DNA sequencing lanes. The DNA sequence complementary to the sequence of the reverse transcriptase product is shown to the left and the right of the correponding panels. The transcription start sites at some mutant promoters are indicated within the sequence.

(c) *Analysis of the spacer region and the role of the distance between the TATA box and the initiator element*

Figure 4 shows that single point mutations in the spacer region between the TATA box and the initiation start site have no pronounced effect on cell-free transcription (lanes 4 to 6). Even a template containing 14 mutations in this region retained levels of transcription that were 40% of wild-type levels (Fig. 5(a), lane 2). These findings suggest that the sequence of the spacer between the conserved elements may modulate promoter strength, but is not essential for faithful initiation of transcription.

To address the question whether the distance of the TATA box relative to the ATGC motif is important for a functional promoter we introduced and deleted sequences in the spacer region (Fig. 5). After insertion of 4 bp at position −13, the rate of transcription was reduced to 30%. The RNA product showed the same electrophoretic mobility as the primary transcript from the wild-type template (Fig. 5(a), lanes 1 and 3). This finding suggests that the transcript from pIC31/54 initiates at the wild-type initiation site. This was confirmed by a primer extension experiment (Fig. 5(b), left side, compare lanes 1 and 3). Therefore, the distance of the initiator nucleotide from the last position of the TATA box can vary from 22 bp (pIC31/56) to 27 bp (pIC31/54), although the efficiency of transcription is lower when the distance deviates from the wild-type situation. However, when 6 bp were inserted (Fig. 5(a), lane 4) or 4 bp were deleted (Fig. 5(a), lane 5) the corresponding transcripts were longer and shorter, respectively (Fig. 5(b), lanes 4 and 5), indicating that initiation of transcription occurred at different sites. A primer extension experiment revealed that the RNAs transcribed from the insertion clone pIC31/55 and the deletion clone pIC31/60 initiate at adenosine residues located five nucleotides upstream and six nucleotides downstream from the wild-type initiator nucleotide, respectively (Fig. 5(b), left side, lanes 4 and 5). The distance of the initiation site of these transcripts from the TATA box is 24 bp and 25 bp, respectively. These results show that initiation of transcription occurs outside the ATGC element when it is placed 19 or 29 bp downstream from the TATA box. The alternate initiation sites TTAG and TCAG share the presence of two pyrimidine residues upstream from the initiator purine. The latter sequence shows homology to the wild-type sequence TTGC only in a single position. This finding argues against the importance of the ATGC motif. On the other hand, no initiation of transcription was detectable when the G at position +1 was replaced by a T (pIC31/53; Fig. 4, lane 8). The transcriptional machinery apparently is unable to initiate at the pyrimidine residues located at the correct distance. Furthermore, it cannot utilize the G residue 19 bp or the A residues 29 bp downstream from the TATA box as alternate initiator nucleotide (Fig. 5(a), lane 8). These data support the conclusion that the minimal requirement for a functional promoter is the presence of a pyrimidine-purine dinucleotide at a distance of 22 to 27 bp downstream from the TATA box. However, at the *nif*H1 promoter of *M. thermolithotrophicus* initiation of transcription can occur at a pyrimidine residue located at a distance of 23 bp downstream from the TATA box when a purine is adjacent to it (H. P. Gohl, W. Hausner & M. Thomm, unpublished results).

(d) *The role of internal DNA sequences*

To investigate a possible promoter function of internal DNA sequences, the DNA region extending from position +6 to +88 of the $tRNA^{Val}$ gene was deleted and replaced by a DNA fragment of 78 bp originating from the *E. coli* plasmid pIC-19H (Fig. 6(a); Materials and Methods). Analysis of cell-free transcripts from this template revealed the presence of two RNA products (Fig. 6(b), lane 2). The intensities of these RNA bands were weak when compared with transcripts from the wild-type template pIC31/2 (lane 1). In the experiments described in Figures 2 to 5(a) we analyzed transcripts from circular templates. Under these conditions, the synthesis of distinct RNA products was dependent upon both correct initiation and termination of transcription. The DNA sequences mediating termination of transcription in archaea have not yet been identified. Also in plasmid pIC31/61 DNA sequences close to the 3′ end of the $tRNA^{Val}$ gene have been replaced which might be essential for termination of transcription. Thus, the lower intensities of the RNA bands transcribed from plasmid pIC31/61 (Fig. 6(b), lane 2) might be caused by inefficient initiation or termination of transcription. To discriminate between these possibilities we analyzed runoff transcripts from linear DNA restriction fragments of pIC31/2 and pIC31/61 (Fig. 6(a), lower part, left side). In a second set of experiments, initiation of transcription at both templates was measured by a quantitative S_1 mapping procedure.

When purified DNA fragments of plasmids pIC31/2 and pIC31/61 containing the promoter but cleaved upstream from the transcription termination site were used as templates in cell-free transcription reactions, distinct RNA products were synthesized (Fig. 6(c), lanes 3 and 4). The size of these transcripts was 89 and 47 nucleotides, respectively, as predicted from the length of the DNA fragments. This finding indicates that these RNAs initiate at the promoter and terminate at the end of the linearized template. These runoff transcripts from pIC31/2 and pIC31/61 showed comparable intensities, indicating that initiation of transcription occured at similar rates on both templates.

To investigate whether transcripts from supercoiled pIC31/2 and pIC31/61 initiate at a similar rate, an S_1 mapping procedure was used. When varying amounts of unlabeled *in vitro* transcripts from these templates were added to an excess of the corresponding ^{32}P-end-labeled probes (Fig. 6(a), lower part, right side), the band intensities of S_1

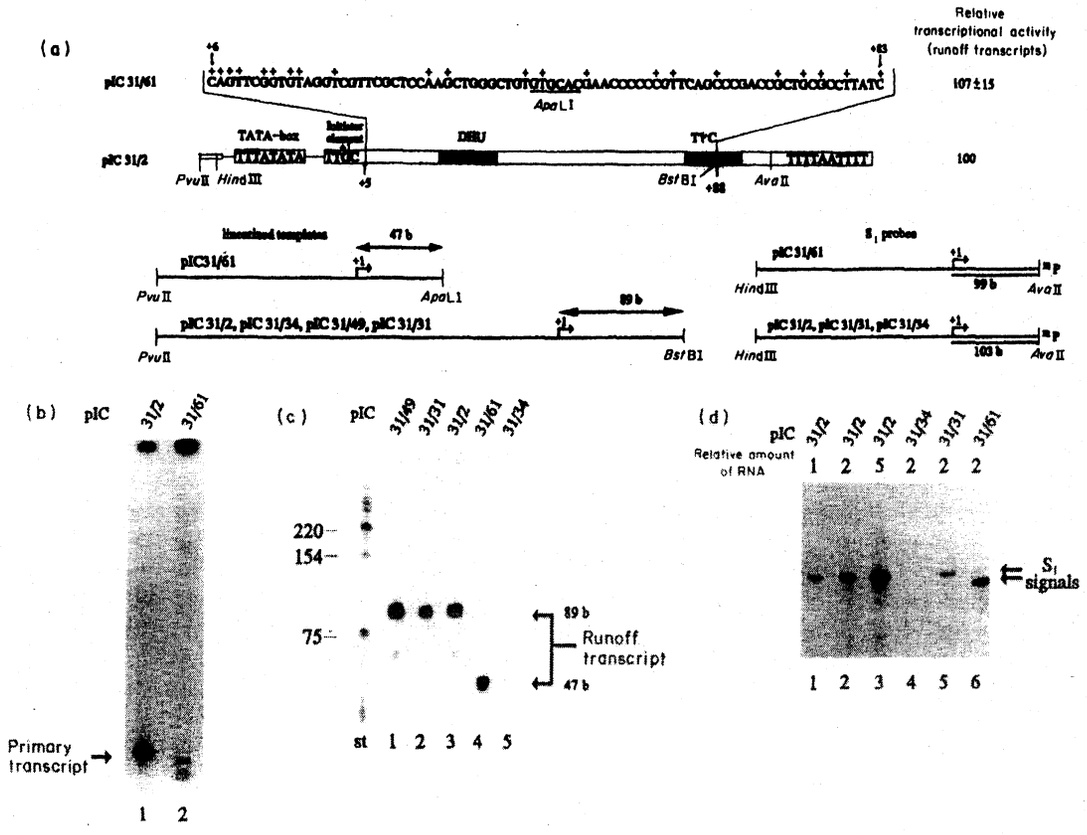

Figure 6. The role of internal DNA sequences. (a) Templates depleted of tRNA coding sequences and S_1 probes. The DNA region from position +6 to +87 of pIC31/2 has been replaced by a fragment of 78 bp derived from pIC-19H (Materials and Methods). This *E. coli* DNA sequence and a map of the resulting clone pIC31/61 is shown at the top. The DNA sequences encoding the DHU and TΨC stem and loop required for initiation of transcription at eukaryotic tRNA genes are indicates by shaded boxes; + indicates nucleotides which are by chance identical to the $tRNA^{Val}$ coding sequence (note that the first 3 and the last nucleotide of the replacing sequence are identical to the wild-type tRNA sequence). The other symbols are as for Fig. 1. The relative activities of runoff transcripts from pIC31/2 and pIC31/61 are shown to the right. The value for pIC31/61 was corrected according to the different U content of the shorter runoff transcript from this template (14 compared to 22 uridine residues; the RNA was labeled with [α-^{32}P]UTP). Left side, lower part: DNA fragments used as templates for runoff transcription assays. The length of the corresponding transcripts initiating at the transcription start site (+1) and terminating at the end of the DNA fragment are indicated. Right side, lower part: the ^{32}P-end-labeled probes for S_1 analysis derived from pIC31/2 and pIC31/61 are shown. 99 b and 103 b indicate the lengths (bases) of the corresponding fragment protected from S_1 digestion by *in vitro* RNA. The probes were used to quantitate transcripts from the clones indicated in the Figure. (b) Direct analysis of cell-free transcripts. Labeled transcripts from supercoiled pIC31/2 (lane 1) and pIC31/61 (lane 2) were analyzed by polyacrylamide/urea gel electrophoresis. The transcripts in lane 2 are shorter since the DNA sequence from +6 to +87 has been replaced by a fragment of 78 bp in pIC31/61. As shown in (c) and (d), transcription initiates at both templates with similar efficiency. The deletion of internal sequences appears to affect the process of termination of transcription. The presence of 2 weak transcripts of similar size from pIC31/61 is indicative for inaccurate termination. Note furthermore, that the amount of high molecular weight RNA not entering the gel is increased in lane 2, which might be caused by read-through at the terminator. This finding, in addition, is consistent with the assumption that pIC31/61 lacks sequences necessary for efficient termination of transcription. (c) Analysis of transcripts from linearized templates. Purified DNA restriction fragments ((a) lower part, left side) of the various plasmids shown at the top of the Figure were used as templates in cell-free transcription reactions. Lane 1, molecular weight standard (^{32}P-end-labeled *Hin*fI fragments of pBR322), the size of some standards and of runoff transcripts from the corresponding templates (89 and 47 nucleotides, respectively) are shown to the left and the right. (d) S_1 mapping of *in vitro* transcripts from supercoiled templates. ^{32}P-end-labeled probes derived from pIC31/2 and pIC31/61 (a) lower part, to the right) were annealed with unlabeled *in vitro* RNA transcribed from the templates indicated at the top of the corresponding lanes and treated with nuclease S_1 (Materials and Methods). The S_1-resistant hybrids were analyzed on a 6% (w/v) polyacrylamide/urea gel. The amount of RNA added to the probe was 1, 2 and 5% of a transcription reaction as indicated on top of the corresponding lanes.

signals were directly proportional to the amount of RNA (Fig. 6(d), lane 1 to 3, and data not shown). At a given RNA concentration, the amount of probe protected from S_1 digestion was similar for pIC31/2 and pIC31/61 (Fig. 6(d), lanes 2 and 6). This finding supports the conclusion that both templates show comparable promoter activities. Hence, internal DNA sequences are not required for transcription initiation. The low amount of specific RNA products transcribed from pIC31/61 (Fig. 6(b),

lane 2) thus appear to be caused by inefficient termination of transcription.

Both the runoff and S_1-assay were used to re-examine the effects of some promoter mutations on the rate of transcription initiation. When transcripts from pIC31/49, pIC31/34 and pIC31/31 were analyzed by these techniques these mutants showed a mild promoter-up (Fig. 6(c), lane 1), a severe promoter-down (Fig. 6(c), lane 5; (d), lane 4) and a medium promoter-down (Fig. 6(c), lane 2; (d), lane 5) phenotype, as expected. These results further substantiate the conclusions derived from a direct analysis of *in vitro* RNA products shown in Figures 2 to 5(a).

4. Discussion

We have identified two elements controlling initiation of transcription of the $tRNA^{Val}$ gene of *M. vannielii*. Several lines of evidence indicate that the TATA box octanucleotide centered at −25 is the major structural determinant of an archaeal promoter. First, we (Thomm *et al.*, 1990) and others (Knaub & Klein, 1990; Hüdepohl *et al.*, 1990) have shown that archaeal genes depleted of the DNA region upstream from position −20 show no or reduced template activity in cell-free transcription experiments. Furthermore, the results of this study demonstrate that deletions progressively approaching the promoter from the 5′ end have no dramatic effect on the rate of transcription until they reach the sequence between −35 and −24 (Fig. 2). Finally, we show here that site-specific mutagenesis of six positions of the sequence located between position −29 and −23 upstream from the $tRNA^{Val}$ gene decreases the rate of transcription drastically. This finding indicates that these nucleotides are essential for promoter function (Figs 3 and 4; the results are summarized in Fig. 7(b)).

The role of the second conserved element, located at the transcription start site, in promoter function is less clear. Point mutations in this region clearly decrease promoter activity (Fig. 7(b)). However, our data suggest that it is not absolutely required for correct initiation of transcription. Analysis of the spacer deletion and insertion clones pIC31/60 and pIC31/55 demonstrates (Fig. 5, lanes 4 and 5) that the components of the transcriptional machinery selecting the transcription start site are not primarily influenced by the location of the ATGC sequence. Transcription initiates at a strictly defined distance 22 to 27 bp downstream from the TATA box. Thus, the TATA box and not the ATGC motif is the major structural DNA element specifying the transcription start site. However, the severe promoter-down phenotype of the wild-type initiator mutant pIC31/53 (Fig. 4, lane 8) indicates that a second sequence downstream from the TATA box is necessary for promoter activity. Our data suggest that a strong contact in the TATA box region directs the RNA polymerase and/or transcription factors to the promoter. A second weak contact between the RNA polymerase and the initiator motif or a pyrimidine-purine dinucleotide possibly anchors the active center of the enzyme at the site where the polymerization reaction starts.

Analyses of transcripts from the spacer mutant clones pIC31/55, pIC31/60 and of the initiator nucleotide point mutation pIC31/53 demonstrate that the transcriptional machinery of *Methanococcus* cannot utilize the purine residues located at a distance of 19 and 29 bp downstream from the TATA box as initator nucleotides. Our results suggest that the potential of the *Methanococcus* RNA polymerase to initiate transcription is strictly limited to the DNA region from 22 to 27 bp downstream from the TATA box, with an optimum initiation rate at a distance of 23 bp. Among wild-type promoters of methanogens the spacer between the TATA box and the initiation site varies in length from 18 to 24 bp (Fig. 7(a)). These data support the conclusion that the correct distance between these two promoter elements is also very critical for promoter function in other genera of archaea. Contradictory to this is the finding that the *Methanobacterium* RNA polymerase initiates transcription *in vitro* 5 bp downstream from the TATA box of the methyl CoM reductase operon (Knaub & Klein, 1990). However, the presence of a transcription factor in the *Methanobacterium* RNA polymerase preparation was not demonstrated. Thus, a possible rationale for this unusual start site could be provided by the observation of Hüdepohl *et al.* (1990) that the semi-specific initiation by the purified RNA polymerase of *Sulfolobus* at ATGC-like sequences can be repressed by the addition of a transcription factor. In light of this and our data it appears likely that the *Methanobacterium* system, initiating at an ATGA sequence, could be lacking a component selecting the correct start site.

All the single point mutations decreasing homology to the consensus sequence at −25 and at the initiation start site confer a severe, or an at least medium, promoter-down phenotype (Fig. 7). Thus, homology to the consensus, especially in positions 2 to 7 of the TATA box appears to be correlated with promoter strength. An argument supporting this conclusion is that strong promoters of methanogens, e.g. those for genes encoding stable RNAs, display high homology to the canonical TATA box and initiator motif (Fig. 7(a)). Among archaea, promoter sequences show an unexpectedly high degree of homology between phylogenetically quite distinct genera, such as *Methanococcus* and *Sulfolobus* (Thomm & Wich, 1988; Reiter *et al.*, 1988). All promoters hitherto mapped in methanogens do not contain a G residue at positions essential for promoter function in the *Methanococcus* system (Fig. 7(a)). However, one distinct difference can be recognized: the promoters of extreme halophiles have a G residue instead of an A at position 6 of the consensus (Thomm *et al.*, 1989). A G residue at this position would not allow initiation in the *Methanococcus* system. However, this difference in the consensus can be correlated with the G+C content of the cellular DNA, which is 60% in

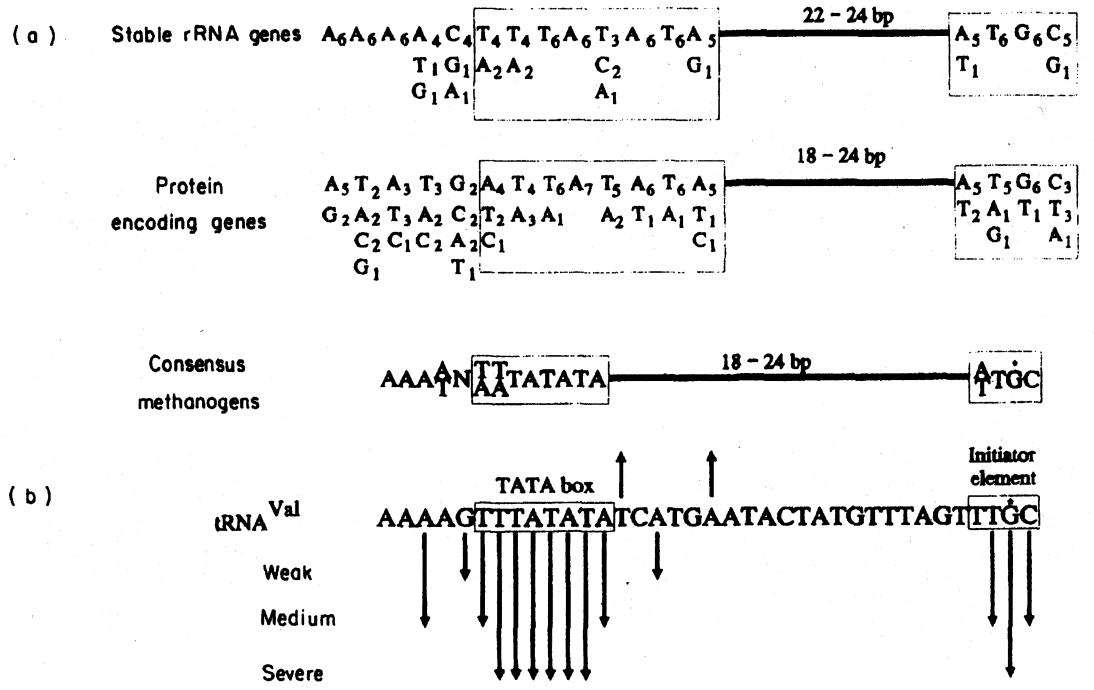

Figure 7. Consensus promoter sequences in methanogens and effects of single point mutations on promoter function. (a) Consensus sequences of protein-encoding and stable RNA genes of methanogens. Sequences upstream from mapped promoters for stable RNA (1st line) and protein-encoding genes (2nd line) were aligned to yield maximum homology. The frequencies of a particular base at each position of the consensus is indicated by subscripts. Note the high homology of both the TATA box and the initiator element to the consensus and the constant spacing of these elements among strong promoters of tRNA/rRNA genes. The consensus shown in line 3 is based on comparison of DNA sequences upstream from 14 transcription start sites which have been mapped *in vivo* (Brown *et al.*, 1989; Souillard & Sibold, 1989; Baier *et al.*, 1990; Eggen *et al.*, 1990; Weil *et al.*, 1989) or *in vitro* (Frey *et al.*, 1990). The sequence of an rRNA promoter of *Methanobacterium thermoautotrophicum* proposed by Ostergaard *et al.* (1987) was not included in this comparison since this gene is preceded by a 7 S RNA gene that had been noticed later (Haas *et al.*, 1990). Therefore, the authors may have mapped a processing site. The variation of the distance between the TATA box and the transcription start site is indicated. The TATA element and the ATGC motif are boxed. (b) Summary of the effects of single point mutations on the template activity of the $tRNA^{Val}$ gene in cell-free transcription experiments. The sequence of the promoter region of the $tRNA^{Val}$ gene of *M. vannielii* is shown below the consensus sequence derived for methanogens (a). Single point mutations conferring a weak, medium and severe promoter-down phenotype (residual activity > 70%, > 30% and < 30%, respectively) are indicated by short, medium and long arrows directed downwards. Arrows directed upwards indicate weak promoter-up mutations.

Halobacterium, and approximately 30% in *Methanococcus* and *Sulfolobus* (Boone & Mah, 1989; Grant & Larsen, 1989). Hence, a structural analysis based on more promoter sequences may reveal genus-specific variations of the *Methanococcus* consensus theme, especially in organisms with high G+C content of the DNA. However, the basic structural features of constitutively expressed promoters appear to be similar in the genera of archaea inspected so far.

A+T-rich sequences also constitute a major element of most promoters of *E. coli* and a number of promoters directing transcription of pre-mRNA by RNA polymerase II. Curiously, the most important part of the archaeal promoter identified in this study is a hexameric sequence

$$5'\text{-}\begin{matrix}\text{T}\\\text{A}\end{matrix}\text{TATAT-}3'$$

that shows homology to the Pribnow box of *E. coli* promoters, 5'-TATAAT-3' (Hawley & McClure, 1983), centered at position −10. However, the location of the archaeal consensus sequence at −25 resembles much more that of the TATA box in higher eukaryotes (Corden *et al.*, 1980; Hu & Manley, 1981). Furthermore, in accord with the *Methanococcus* enzyme, RNA polymerase II initiates at a constant distance of 25 to 29 bp downstream from the TATA box when a linker is inserted between this sequence and the transcription start site (Dierks *et al.*, 1983). Thus, in both systems the TATA box is the dominant element in determining the transcription start site. The DNA region required for optimal initiation of transcription by RNA polymerase II is larger when assayed *in vivo* rather than in cell-free transcription experiments (McKnight *et al.*, 1981). We have to await the development of efficient transformation systems in *Methanococcus* to address the question as to whether the promoter sequences identified in this study are sufficient for maximal initiation of transcription *in vivo*. In contrast to transcription of eukaryotic

tRNA genes by RNA polymerase III (Geiduschek & Tocchini-Valentini, 1988) internal sites are not essential for the promoter function of the $tRNA^{Val}$ gene of *M. vannielii*. Both analyses of runoff transcripts from linear fragments of pIC31/61 and a quantitative S_1 study of transcripts from a circular template lacking tRNA coding sequences demonstrate that the DNA sequences downstream from position +9 are not necessary for accurate and efficient initiation. Thus, the DNA sequences encoding the DHU and TΨC stem and loop of $tRNA^{Val}$, which correspond to the A box and B box element of RNA polymerase III promoters, are clearly not a part of this *Methanococcus* promoter. However, our data suggest that internal sequences are required for efficient termination of transcription.

The components of the transcriptional machinery of *Methanococcus* interacting with the elements of the $tRNA^{Val}$ promoter remain to be identified. Beside the RNA polymerase at least two transcription factors appear to be involved in initiation of transcription at *Methanococcus* promoters (Frey *et al.*, 1990; our unpublished results). The binding region of the purified RNA polymerase extends from position −30 to +20 (Thomm & Wich, 1988; Brown *et al.*, 1988; Thomm *et al.*, 1988). Hence, the promoter-down mutations in the TATA box region may exert their effect by disturbing chemical contacts between the enzyme and this DNA region. However, the TATA box independent initiation of transcription of the *Sulfolobus* RNA polymerase at ATGC-like sequences provides evidence that archaeal RNA polymerases can also recognize the initiator element. One of the *Methanococcus* transcription factors also appears to interact primarily with the promoter region (our unpublished results). Thus, a more complex assembly of transcriptional components at the promoter may precede the initiation process in archaea. The set of mutants derived for this study may prove useful in investigating the sequence and nature of the interactions of these polypeptides with the promoter.

We appreciate the excellent technical assistance of Sigrid Bogner. We thank Drs G. R. Hartmann and G. Sawers for comments on the manuscript. This work was funded by a grant from the Deutsche Forschungsgemeinschaft and the Fonds der Chemischen Industrie. We thank Dr K. O. Stetter for supporting this work by funds from the Leibniz Preis.

References

Baier, G., Piendl, W., Redl, B. & Stöffler, G. (1990). Structure, organization and evolution of the L1 equivalent ribosomal protein gene of the archaebacterium *Methanococcus vannielii*. *Nucl. Acids Res.* **18**, 719–724.

Boone, D. R. & Mah, R. A. (1989). Methanogenic archaeobacteria. In *Bergey's Manual of Systematic Bacteriology* (Staley, J. T., Bryant, M. P., Pfennig, N. & Holt, J. G., eds), vol. 3, sect. 25, Williams & Wilkins, Baltimore.

Breathnach, R. & Chambon, P. (1981). Organization and expression of eukaryotic split genes coding for protein. *Annu. Rev. Biochem.* **50**, 349–383.

Brown, J. D., Thomm, M., Beckler, G., Frey, G., Stetter, K. O. & Reeve, J. N. (1988). An archaebacterial RNA polymerase binding site and transcription initiation of the *his*A gene in *Methanococcus vannielii*. *Nucl. Acids Res.* **16**, 135–150.

Brown, J. D., Daniels, C. & Reeve, J. N. (1989). Gene structure, organization and expression in archaebacteria. *Crit. Rev. Microbiol.* **16**, 287–338.

Corden, J., Wasylyk, B., Buchwalder, A., Sassone-Corsi, P., Kedinger, C. & Chambon, P. (1980). Promoter sequences of eukaryotic protein-coding genes. *Science*, **209**, 1406–1414.

Dierks, P., van Ooyen, A., Cochran, M. D., Dobkin, C., Reiser, J. & Weissmann, C. (1983). Three regions upstream from the Cap site are required for efficient and accurate transcription of the rabbit β-globin gene in mouse 3T6 cells. *Cell*, **32**, 695–706.

Eggen, R., Harmsen, H. & de Vos W. M. (1990). Organization of a ribosomal RNA gene cluster from the archaebacterium *Methanothrix soehngenii*. *Nucl. Acids Res.* **18**, 1306.

Frey, G., Thomm, M., Brüdigam, B., Gohl, H. P. & Hausner, W. (1990). An archaebacterial cell-free transcription system. The expression of tRNA genes from *Methanococcus vannielii* is mediated by a transcription factor. *Nucl. Acids Res.* **18**, 1361–1367.

Geiduschek, P. & Tocchini-Valentini G. P. (1988). Transcription by RNA polymerase III. *Annu. Rev. Biochem.* **57**, 873–914.

Grant, W. D. & Larsen, H. (1989). Extremely halophilic archaeobacteria. In *Bergey's Manual of Systematic Bacteriology* (Staley, J. T., Bryant, M. P., Pfennig, N. & Holt, J. G., eds), vol. 3, sect. 25, Williams and Wilkins, Baltimore.

Haas, E. S., Brown, J. W., Daniels, C. J. & Reeve, J. N. (1990). Genes encoding the 7 S RNA and a $tRNA^{Ser}$ are linked to one of the two rRNA operons in the genome of the extremely thermophilic archaebacterium *Methanothermus fervidus*. *Gene*, **90**, 51–59.

Hawley, D. K. & McClure, W. R. (1983). Compilation and analysis of *Escherichia coli* promoter DNA sequences. *Nucl. Acids Res.* **11**, 2237–2255.

Henikoff, S. (1984). Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. *Gene*, **28**, 351–359.

Hu, S. & Manley, J. L. (1981). DNA sequence required for initiation of transcription *in vitro* from the major late promoter of adenovirus 2. *Proc. Nat. Acad. Sci., U.S.A.* **78**, 820–824.

Hüdepohl, U., Reiter, W. D. & Zillig, W. (1990). *In vitro* transcription of two rRNA genes of the archaebacterium *Sulfolobus sp.* B12 indicate a factor requirement for specific transcription. *Proc. Nat. Acad. Sci., U.S.A.* **87**, 5851–5855.

Knaub, S. & Klein, A. (1990). Specific transcription of cloned *Methanobacterium thermoautotrophicum* transcription units by homologous RNA polymerase *in vitro*. *Nucl. Acids Res.* **18**, 1441–1446.

Marsh, J. L., Erfle, M. & Wykes, E. J. (1984). The pIC plasmid and phage vectors with versatile cloning sites for recombinant selection by insertional inactivation. *Gene*, **32**, 481–489.

McKnight, S. L., Gavis, E. R., Kingsbury, R. & Axel, R. (1981). Analysis of transcriptional regulatory signals of the HSV thymidine kinase gene: identification of an upstream control region. *Cell*, **25**, 385–398.

Ostergaard, L., Larsen, N., Leffers, H., Kjems, J. & Garret, R. (1987). A ribosomal RNA operon and its flanking region from the archaebacterium *Methanobacterium thermoautotrophicum*. *Syst. Appl. Microbiol.* **9**, 199–209.

Reiter, W. D., Palm, P. & Zillig, W. (1988). Analysis of transcription in the archaebacterium *Sulfolobus* indicates that archaebacterial promoters and eukaryotic RNA polII promoters are of the same type. *Nucl. Acids Res.* **16**, 1–19.

Sanger, F. Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Nat. Acad. Sci., U.S.A.* **74**, 5463–5467.

Siebenlist, U., Simpson, R. B. & Gilbert, W. (1980). *E. coli* RNA polymerase interacts homologously with two different promoters. *Cell*, **20**, 269–281.

Souillard, N. & Sibold, L. (1989). Primary structure, functional organization and expression of the nitrogenase structural genes of the thermophilic archaebacterium *Methanococcus thermolithotrophicus*. *Mol. Microbiol.* **3(4)**, 541–551.

Thomm, M. & Wich, G. (1988). An archaebacterial promoter element for stable RNA genes with homology to the TATA box of higher eukaryotes. *Nucl. Acids Res.* **16**, 151–163.

Thomm, M., Sherf, B. & Reeve, J. (1988). RNA polymerase binding and transcription initiation sites upstream of the the methyl reductase operon of *Methanococcus vannielii*. *J. Bacteriol.* **170**, 1958–1961.

Thomm, M., Wich, G., Brown, J., Frey, G., Sherf, B. & Beckler, G. S. (1989). An archaebacterial promoter sequence assigned by RNA polymerase binding experiments. *Canad. J. Bacteriol.* **35**, 30–35.

Thomm, M., Frey, G., Hausner, W. & Brüdigam, B. (1990). An archaebacterial *in vitro* transcription system. In *Microbiology and Biochemistry of strict Anaerobes involved in Interspecies Hydrogen Transfer* (Belaich, J. P., Bruschi, M. & Garcia, J.-L., eds), part 2.3., Plenum Press, New York.

Weil, C., Sherf, B. & Reeve J. (1989). A comparison of the methyl reductase genes and gene products. *Canad. J. Bacteriol.* **35**, 101–108.

Wich, G., Hummel, H., Jarsch, M., Bär, U. & Böck, A. (1986). Transcription signals for stable RNA genes in *Methanococcus*. *Nucl. Acids Res.* **14**, 2459–2479.

Woese, C. R., Kandler, O. & Wheelis, M. L. (1990). Towards a natural system of organisms. Proposal of the domains archaea, bacteria and eucarya. *Proc. Nat. Acad. Sci., U.S.A.* **87**, 4576–4579.

Yanisch-Perron, C., Vieira, J. & Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene*, **33**, 103–119.

Edited by M. Yaniv