Transcription In Vivo and In Vitro of the Histone-Encoding Gene hmfB from the Hyperthermophilic Archaeon * Methanothermus fervidus*

M. THOMM,† K. SANDMAN, G. FREY, G. KOLLER, AND J. N. REEVE

Lehrstuhl für Mikrobiologie, Universität Regensburg, D-8400 Regensburg, Germany, and Department of Microbiology, The Ohio State University, Columbus, Ohio 43210-1292

Received 7 November 1991/Accepted 24 March 1992

Immediately upstream of the hmfB gene, in a DNA fragment cloned from Methanothermus fervidus, are two identical tandemly repeated copies of a 73-bp sequence that contain the sequence 5'TTTATATA, which conforms precisely to the consensus TATA box element proposed for methanogen promoters. By using this duplicated region as the template DNA and a cell-free transcription system derived from Methanococcus thermolithotrophicus, transcription in vitro was found to initiate at two identical sites 73 bp apart, each 25 bp downstream from a TATA box, thus providing strong evidence for the functional conservation of this transcriptional signal in two phylogenetically very diverse methanogens. Transcription of the hmfB gene in vivo in *M. fervidus* was found to occur at only one of these sites, and consistent with this observation, reeling and sequencing of this intergenic region after its amplification by the polymerase chain reaction demonstrated that the gene of *M. fervidus* contains only one copy of the 73-bp sequence upstream of the hmfB gene. Since the second copy of the 73-bp sequence, presumably generated artifically during the original hmfB cloning, functioned equally well as a promoter in the *M. thermolithotrophicus* transcription system, all information needed by the heterologous RNA polymerase to initiate transcription accurately in vitro must be present within this sequence. The hmfB gene encodes HMf-2, one of the two subunits of HMf, an abundant DNA binding protein in *M. fervidus* which binds to DNA molecules in vitro, forming nucleosomelike structures. Cell-free transcription was inhibited by adding HMf or eucaryotic core histones at protein-to-DNA mass ratios of 0.3:1 and 1:1, respectively, whereas the archaeal histonelike protein HTa from Thermoplasma acidophilum inhibited transcription in vitro only at much higher protein-to-DNA mass ratios and the bacterial histonelike protein HU from Escherichia coli had no detectable effect on transcription.

The TATA box was initially recognized because of its conservation upstream of genes cloned from a wide range of methanogens (2, 23, 25), suggesting that a TATA box element from one methanogen might be recognized and used to initiate transcription by RNA polymerases from a different methanogen. The discovery of a pair of identical 73-bp sequences, each of which contained a consensus TATA box region, led to the conclusion that a TATA box region is present and used as a promoter in vivo. As this duplicated sequence was located immediately upstream of a cloned gene, namely, the hmfB gene, which encodes the HMf-2 subunit of the abundant DNA binding protein HMf in *M. fervidus* (11, 14), it strongly suggested that hmfB transcription in vivo must be directed by two identical promoters. It was therefore somewhat of a surprise that although 16 of 17 bases in the TATA box region immediately upstream of the closely related hmfB gene, cloned from Methanobacterium thermosautotrophicum D1H, were perfectly conserved (20), there was no duplicated region upstream of this gene. We have therefore investigated transcription of the hmfB gene both in vivo and in vitro and demonstrate here that although both TATA boxes associated with the cloned hmfB gene are recognized and used as promoters in vitro by RNA polymerases from *Methanococcus thermolithotrophicus* (10, 21), only one TATA box region is present and used as a promoter in vivo in *M. fervidus*. The duplication apparently occurred as an artifact during the original hmfB cloning. HMf binds tightly to DNA in vitro, forming nucleosomelike structures in which the DNA molecule is constrained in a positive toroidal...
supercoil (12, 14). It was therefore also important to determine whether HMf itself had an effect on transcription, and we demonstrate here that HMf is a potent inhibitor of transcription in vitro.

MATERIALS AND METHODS

Growth of methanogens and isolation of nucleic acids. Cultures of *M. thermolithotrophicus* and *M. fervidus* were grown anaerobically at 65 and 83°C, respectively (18, 21, 24). Nucleic acids were prepared as previously described from exponentially growing *M. fervidus* cells that were rapidly frozen in liquid N$_2$ and ruptured by grinding with a pestle and mortar (24).

Construction and preparation of template DNAs. The hmfB gene was cloned into pUC19 as part of a 3.2-kb HindIII-PstI restriction fragment. The sequences of the cloned hmfB gene and its flanking regions were determined (14, 15), and the overall structure of the resulting plasmid, pKS303, is shown in Fig. 1. Subcloning of the HindIII-Rsal fragment indicated in Fig. 1 into pUC19 generated pKS304, which was used as a template for in vitro transcription studies and as a substrate to obtain deletions lacking DNA from the 3' end of hmfB. Plasmid pKS304 was digested with both BamHI and KpnI, and, as the 3' single-stranded extensions generated by KpnI digestion are not substrates for exonuclease III digestion, subsequent exonuclease III digestion resulted in the loss of DNA from pKS304 only in the direction toward hmfB. The termini of the molecules so obtained were made double stranded by digestion with S1 nuclease and by filling with DNA polymerase (Klenow fragment). After ligation and transformation of *Escherichia coli* DH5a, the precise extents of the deletions produced were determined by DNA sequencing. Plasmid pKS304A16 (Fig. 1) was found to contain the first 51 bp of hmfB followed by 6 bp of vector DNA and then the EcoRI site of the pUC19 vector. Plasmid pKS304A16 DNA linearized by EcoRI digestion was used in vitro as the template in runoff transcription experiments.

Purification of RNA polymerase and transcription factors. Components of the cell-free system from *M. thermolithotrophicus* were purified either (i) by phosphocellulose (PC) chromatography as described previously (21), in which case the 0.35 M KCl PC eluate contained the RNA polymerase activity (0.25 mg/ml) and the 0.6 M KCl PC eluate contained a transcription factor (7), or (ii) by sequential S-100, DEAE-cellulose, heparin-cellulose, and Mono Q (fast protein liquid) chromatography, in which case the preparations obtained were ~80% RNA polymerase and were then combined with the PC transcription factor. *M. thermolithotrophicus* RNA polymerase prepared by either procedure generated the same transcripts of the hmfB gene in vitro.

In vitro transcription reactions. Purified, linear template DNA fragments (0.01 µg/ml) were incubated with RNA polymerase (20-µl PC fraction or 5-µl Mono Q fraction) and PC-purified transcription factor (2 µl) in vitro transcription reactions as described previously (7), except that the KCl concentration was increased from 50 to 110 mM.

SI nuclease mapping. The single-stranded probe used for S1 mapping experiments (1) was prepared from plasmid pKS304. Supercoiled DNA (18 µg) was denatured by alkali treatment and annealed with 20 pmol of 32P-end-labeled PE2 primer (Fig. 2A). The primer was extended with the Klenow fragment of DNA polymerase I (10 U) for 30 min at 37°C. To obtain a probe with a uniform 3' end, the DNA was cleaved with BstXI at position -122 relative to the site of transcription initiation directed by the upstream TATA box. The resulting single-stranded DNA fragment, 283 nucleotides in length, was purified from an alkaline agarose gel, and the 5' ends of in vitro transcripts were identified by the S1 mapping procedure (1) modified as previously described (9, 24). RNA purified from *M. fervidus* cells (12 µg) was annealed with the end-labeled probe (50,000 cpm) in a buffer containing 40 mM PIPES (piperazine-N,N'-bis(2-ethanesulfonic acid); pH 6.4), 80% formamide, 0.4 M NaCl, and 1 mM EDTA for 3 h at 40°C, and S1 nuclease digestions were performed for 1 h at 30°C.

Primer extension analyses. Primer extension and sequencing of transcripts synthesized in vitro were performed with 0.1 pmol (30,000 cpm) of the 32P-end-labeled primers designated PE1 and PE2 in Fig. 2A. Unlabeled RNA was isolated from the in vitro transcription reaction mixtures as described previously (7). The same protocols were employed to identify sites of transcription initiation in vivo by using 10 µg of total cellular RNA extracted from *M. fervidus* cells (24) and 0.03 pmol of the end-labeled PE1 primer.

Polymerase chain reaction (PCR) amplification. *M. fervidus* genomic DNA (100 ng) was denatured at 94°C for 5 min and mixed with 100 pmol of each of the primers PCR1 and PCR2 (Fig. 2A) in a 100-µl reaction mixture containing 2.5 U of AmpliTaq and the PCR buffer supplied by Perkin-Elmer-Cetus (Emeryville, Calif.). Amplification was allowed to proceed through 30 cycles at 94°C (1 min), 40°C (3 min), and 72°C (3 min). The amplified product obtained (125 bp) was digested with EcoRI and sequenced after its cloning into EcoRI-digested pUC19.

HMf purification. HMf was purified from *M. fervidus* cells as previously described (11). Core histones, purified from chicken erythrocytes, were generously provided by M. T. Muller (The Ohio State University). HU and HTa were purified by previously described procedures (4, 5).

RESULTS

Runoff transcription of the hmfB gene in vitro. RNA purified from *M. thermolithotrophicus* was used to transcribe pKS304A16 DNA linearized by digestion with EcoRI (Fig. 1A). Transcription must have been initiated in both duplicated regions upstream of the *M. fervidus* hmfB gene, as two transcripts (I and II [Fig. 1B]) which differed in length by 73 bases were synthesized. These transcripts formed two bands with approximately equal intensities on low-resolution gels. When separated on DNA sequencing gels, both transcripts could be further resolved into two bands which differed in length by 1 base. On the basis of their sizes (154 and 155 bases for transcript I and 81 and 82 bases for transcript II), transcription was initiated 24 and 25 bp downstream from each of the duplicated TATA boxes at the CG dinucleotide within the duplicated boxB-like sequence 5'TCGT (Fig. 2A).

Inhibition of in vitro transcription by HMf binding to the template DNA. Binding HMf to the pKS304A16 template before adding this DNA to the in vitro transcription system inhibited the synthesis of both transcripts (Fig. 1B). Adding HMf to an already actively transcribing reaction mixture immediately inhibited any further transcription. This inhibition was not, however, specific to DNA templates containing the hmfB gene or to methanogen-derived transcription systems. Transcription in vitro of the *M. fervidus* 7S RNA-encoding gene (8) by the *M. thermolithotrophicus* system and transcription by coliphage T7 RNA polymerase of cloned into T7 transcription vectors were also inhibited by HMf binding to similar HMf-to-DNA ratios (results not shown). HMf is most closely related to eucaryal histones (14), and
FIG. 1. Template construction and in vitro runoff transcription. (A) A 3.2-kb HindIII (H)-PstI (P) restriction fragment containing the hmfB gene and an adjacent but oppositely oriented open reading frame (ORF285) were cloned into pUC19, generating pKS303 (14). The sequences of the cloned hmfB and of the upstream, directly repeated 73-bp regions (indicated by arrows) containing the consensus TATA box, 5'TTTATATA (black boxes), have been published elsewhere (14) (see also Fig. 2A). The amino acid sequence encoded by ORF285 indicates a common evolutionary ancestry with arginases and agmatinas (15). Linearization of pKS303Δ16 by EcoRI (E) digestion resulted in the template used in vitro in the runoff transcription experiments shown in panels B and C. R, RsaI; Hc, HincII; B, BamHI; K, Kpnl; E, EcoRI. (B) Radioactively labeled transcripts synthesized in vitro in the absence (track C) and presence of HMf and/or core chicken histones were separated by electrophoresis through 6% polyacrylamide gels and detected by autoradiography. The RNA molecules synthesized, indicated by I and II, have sizes consistent with their being transcripts I and II as indicated in panel A. In tracks 1 and 2, HMf was prebound to the template DNA at HMf-to-DNA mass ratios of 1:6 and 1:3, respectively, before transcription was initiated. Similarly, in tracks 3 and 4, chicken histones were prebound to the DNA at protein-to-DNA mass ratios of 1:2 and 1:1, respectively, before transcription was initiated. In tracks 5 and 6, both HMf and the histones were bound to the DNA before transcription was initiated. In both cases, the HMf-to-DNA mass ratio was 1:6, with the histone-to-DNA mass ratios being 1:2 and 1:1 in tracks 5 and 6, respectively. (C) As in panel B, except that in tracks 1 through 4 HU protein from E. coli (4) was added at protein-to-DNA mass ratios of 2:1, 10:1, 20:1, and 40:1, respectively, and in tracks 5 through 7 HTa protein from T. acidophilum (5) was added at protein-to-DNA mass ratios of 1.6:1, 3:1, and 6:1, respectively.
FIG. 2. Sequence of the region upstream of the cloned hmfB gene and determination of the sites of initiation of hmfB transcription in vivo and in vitro. (A) The sequence and location of the duplicated 73-bp regions are shown, with the TATA box promoter sequences boxed. The adjacent nucleotides identified as the sites of initiation of transcription of hmfB in vivo are indicated by the arrow labeled I. The oligonucleotides used in primer extension experiments (PE1 and PE2), for RNA sequencing (PE1), and for PCR amplification (PCR1 and PCR2) are indicated. An inverted repeat sequence located immediately upstream of the duplication is indicated by converging dashed arrows. (B) The adjacent nucleotides at which hmfB transcription was initiated in vivo, determined as shown by S1 nuclease protection, are identified in the sequencing ladder to the left of the gel. Track K contained no M. fervidus RNA, and tracks 1, 2, and 3 show results obtained with S1 nuclease at concentrations of 100, 200, and 300 U per reaction mixture, respectively. (C) Sequence of the RNA synthesized in vitro from the upstream promoter region. Primer PE1 was used with reverse transcriptase to determine the 5' sequence of the products of an in vitro transcription reaction (7). As shown, the sequence obtained was that predicted for the complement of transcripts initiated at the CG dinucleotide indicated by the arrowhead labeled I in panel A. The control track (K) did not contain in vitro-synthesized RNA.
binding of chicken histones to the pKS304Δ16 template DNA was also found to inhibit transcription in vitro, although about fourfold more protein was required for the same level of inhibition. Adding both HMf and the chicken histones to the *M. thermolithotrophicus* transcription system resulted in an additive effect on the inhibition of transcription in vitro (Fig. 1B). Adding the histonelike protein HTa, isolated from the thermophilic but nonmethanogenic archaean *Thermoplasma acidophilum* (5), also partially inhibited transcription in vitro; however, >10-fold more HTa was needed to obtain the same levels of inhibition observed with HMf. Adding HU from *E. coli* (4) had no detectable effect on transcription, even at very high (40:1) protein-to-DNA ratios (Fig. 1C).

**Site of transcription initiation of the hmfB gene in vivo.** S1 nuclease protection experiments using RNA isolated from growing *M. fervidus* cells indicated that hmfB transcription in vivo was initiated only at the CG dinucleotide 23 and 24 bp upstream of the hmfB gene (Fig. 2B). These are the same nucleotides that are used in vitro by the *M. thermolithotrophicus* RNA polymerase to initiate synthesis of the runoff transcripts II. The S1 nuclease protection procedure gave no evidence for longer transcripts in vivo which might have been initiated at the upstream promoter and been equivalent to the transcript I product synthesized in vitro. This result was confirmed by primer extension experiments using PE2 (Fig. 2A) as the primer. As predicted by S1 mapping, transcripts that had been initiated at the CG dinucleotide closest to the hmfB gene were identified in preparations of RNA from *M. fervidus* cells by primer extension, but there was no indication of longer transcripts (results not shown). Primer extension experiments using primer PE1 (Fig. 2A) corroborated this result. Transcripts initiated at the upstream TATA box or completely upstream of the duplicated region would have been identified by hybridization to PE1, but despite considerable effort, such transcripts were never detected in RNA preparations from *M. fervidus* cells. To confirm that synthesis of transcript I in vitro was initiated within the upstream duplicated sequence as predicted in Fig. 2A, PE2 was used with reverse transcriptase to obtain the sequence complementary to the 5′ sequence of this transcript. The sequence obtained (Fig. 2C) was identical to the DNA sequence determined previously for this region (14).

**PCR amplification, recloning, and sequencing of the intergenic region.** The absence in vivo of hmfB transcripts initiated at the upstream promoter, coupled with the discovery that there was no duplication upstream of the closely related hmfB gene cloned from *M. thermoautotrophicum* ΔH (20), suggested that a 73-bp duplication might have occurred during the hmfB cloning. To investigate this possibility, the oligonucleotide primers designated PCR1 and PCR2 in Fig. 2A were used in a PCR to amplify the intergenic region upstream of the hmfB gene directly from *M. fervidus* genomic DNA. One dominant product which had an electrophoretic mobility indicating a length of ~125 bp was generated by the PCR. As indicated in Fig. 2A, PCR1 and PCR2 were designed to introduce 1-bp deletions to create flanking EcoRI sites, and after the PCR product was digested with EcoRI, it was cloned into EcoRI-digested pUC19. The sequence of this cloned DNA fragment was then determined and found to be the same as that shown in Fig. 2A but with only one copy of the 73-bp sequence.

**DISCUSSION**

Hausner et al. (9), using an homologous in vitro system in which both the RNAP and tRNA<sup>Val</sup> template gene were from *M. vannei*, showed that the TATA box sequence was essential for transcription initiation. Their work, combined with previous comparative studies (3, 7, 13, 25), predicted that TATA boxes conforming to the consensus sequence 5′ TTATATATA are likely to direct transcription initiation of many methanogen genes in a wide range of methanogens. We chose to test this prediction by using as the template DNA a cloned *M. fervidus* DNA sequence that contained two consensus TATA boxes within tandemly arranged, identical 73-bp sequences and RNAP isolated from the very distantly related methanogen *M. thermolithothrophicus*. As predicted, the in vitro transcription system prepared from *M. thermolithotrophicus* recognized and employed both *M. fervidus*-derived TATA boxes. Two transcripts which differed in length by 73 bases and which were initiated at the anticipated pyrimidine-purine dinucleotide downstream from the two TATA boxes were synthesized in approximately equal amounts. As transcription in vivo in *M. fervidus* was found to occur only at the dinucleotide closest to the hmfB gene, it appeared that there was a difference between promoter use in vivo and in vitro. Subsequent investigations, however, revealed that the 73-bp duplication did not exist in vivo in the genome of *M. fervidus*. A 73-bp duplicating event must have occurred aberrantly during the initial cloning of the hmfB gene into *E. coli*. Regardless of its origin, the efficient use of the duplicated sequence as a promoter in vitro clearly demonstrates that it must contain all the sequence information required by *M. thermolithothrophicus* RNA polymerase to initiate transcription accurately in vitro. This demonstration of conservation of promoter structure in very different methanogens has important phylogenetic implications and could be very important, in practical terms, for future cloning and heterologous gene expression studies of methanogens. The small PCR-generated EcoRI fragment (Fig. 2A) that contains the hmfB promoter could also be very useful for in vitro genetic constructions as an easily moveable methanogen promoter.

HMf is a major constituent of *M. fervidus* cells (11, 19). Presumably, it plays an important role in the topology of the genome of this hyperthermophile (12, 14), and by analogy with other histones (6, 17) and histonelike proteins (16), it is probably also involved in regulating gene expression and in recombination. The potent inhibition of transcription in vitro by HMf (Fig. 1B) obviously raises the question of how transcription in vivo avoids this inhibition. It is well established that in eucaryotic cells transcriptionally active DNA is depleted of nucleosomes and that nucleosome positioning can regulate promoter availability for transcription initiation (6, 17). There is sufficient HMf in *M. fervidus* cells to sequester ~25% of the genome in nucleosome-like structures (12, 14, 19), and therefore, in vivo, this binding must be localized and transitory. It has been shown previously that if protein-free DNA is added to preformed HMf-DNA complexes, HMf molecules rapidly migrate from the complexes to the newly added DNA (11). Adding protein-free competitor DNA to HMf-inhibited in vitro transcription reactions, such as those shown in Fig. 1B, restored transcription (results not shown). Since ~75% of the *M. fervidus* genome should always be available in vivo for HMf binding, HMf inhibition of transcription and DNA replication in vivo could be avoided by continually relocating HMf molecules from
sites of activity to temporarily inactive regions of the genome.

**ACKNOWLEDGMENTS**

This work was supported by contract N00014-86-K-0211 from the Office of Naval Research, a grant from the Deutsche Forschungsgemeinschaft, the Fonds der Chemischen Industrie, and NATO Scientific Affairs collaborative grant 0148/85. We thank K. O. Stetter for supporting this study with funds from the Leibniz-Preis. We also thank M. T. Muller for the gift of chicken histones and an anonymous reviewer for suggesting that we determine whether the 73-bp duplication was a cloning artifact.

**REFERENCES**


8. **Haas, E. S., J. W. Brown, C. J. Daniels, and J. N. Reeve.** 1990. Genes encoding the 7S RNA and tRNA<sup>5S</sup> are linked to one of the two rRNA operons in the genome of the extremely thermophilic archaeabacterium *Methanothermus fervidus*. Gene 90:51–59.


15. **Sandman, K., V. J. Steigerwald, and J. N. Reeve.** Unpublished results.


