Transcription initiation and a RNA polymerase binding site upstream of the *purE* gene of the Archaeabacterium *Methanobacterium thermoautotrophicum* strain ΔH

James W. Brown 1,* and John N. Reeve 1,2

1 Molecular, Cellular and Developmental Biology Program, 2 Department of Microbiology, The Ohio State University, Columbus, OH 43210, U.S.A.

Received 28 December 1988
Revision received 2 March 1989
Accepted 4 March 1989

Key words: Archaeabacterium promoter; Filter-binding; Methanogen; DNA-dependent RNA polymerase

1. SUMMARY

DNA-dependent RNA-polymerase (RNAP) purified from the thermophilic archaeabacterium *Methanobacterium thermoautotrophicum* strain ΔH has been shown to bind specifically to DNA in the intergenic region upstream of the *purE* gene cloned from this species. The RNAP binding site has been limited to a 41 bp region of DNA which contains the Box A archaeabacterial promoter sequence, 5'ATTAAATA. Transcription of the *purE* gene appears to initiate in vivo at three locations 20–22 bp downstream of the Box A sequence and 27–29 bp upstream of the ATG translation initiation codon of the *purE* gene.

2. INTRODUCTION

All DNA-dependent RNA-polymerases (RNAP) so far isolated from archaeabacteria are very different from eubacterial RNAPs [1–4] and therefore the structure(s) of archaeabacterial promoters is currently a topic of much interest and speculation [4–6]. Archaeabacterial promoters have been identified primarily as consensus sequences located in the vicinity of the sites at which transcription initiation has been shown to occur in vivo using primer extension and S1-nuclease protection procedures [6–10]. Only for the mesophilic methanogen, *Methanothermobacter vannielii*, has RNAP been purified and used to correlate the sites of RNAP-binding in vitro with the in vivo sites of transcription initiation [11–13]. We have now extended this approach using RNAP purified from a very different, thermophilic methanogen, *Methanobacterium thermoautotrophicum* ΔH.

3. MATERIALS AND METHODS

3.1. M. thermoautotrophicum RNAP purification

RNAP was purified anaerobically from strains of *M. thermoautotrophicum* by a modification of the published procedure ([1,3]; see Fig. 1). *M. thermoautotrophicum* cell paste (40 g wet weight of exponentially growing cells) was resuspended in
pressure cell at 20,000 p.s.i. The resulting lysate was cleared by centrifugation (30,000 × g, 4°C, 10 min), adjusted to 6% PEG 6000 and stirred on ice for 1 h. The supernatant, obtained after centrifugation (29,000 × g, 4°C, 30 min), was dialyzed against TMK and mixed with 100 ml DEAE-cellulose (pre-equilibrated in TMK) by stirring for 30 min. The slurry was poured into a 2.5 × 20 cm column, washed with 300 ml TMK and bound proteins eluted with 800 ml of a 50 to 800 mM KCl gradient (in TMK). Fractions containing RNAP activity were pooled, dialyzed against TMK and applied to a 50 ml double-stranded DNA-cellulose column. After washing with 3 column volumes of TMK, bound proteins were eluted with 500 ml of a 50 to 800 mM KCl gradient (in TMK). Active fractions were pooled and concentrated by ultrafiltration to a volume of 400 μl, then loaded onto a 10–30% w/v sucrose; 5–10% v/v glycerol gradient (in TMK with 500 mM NH₄Cl) for centrifugation at 286,000 × g for 24 h at 18°C (SW41 rotor, 41,000 RPM, 1.6 × 10¹² rad²/s). Aliquots from the fractionated gradient were assayed for RNAP activity, protein concentration and characterized by SDS-PAGE. Fractions with the highest activity were pooled, dialyzed against TMK and stored as small aliquots under liquid nitrogen (Fig. 1).

### 3.2 Filter binding assays

Construction and sequencing of plasmid pET441, a pUC8 derivative containing a 1.6 kbp fragment of *M. thermoautothrophicum* strain ΔH DNA which encodes the entire purE gene, the amino-terminus of a downstream ORF (ORF₆) and the amino-terminus of an oppositely-oriented upstream ORF (ORF₇), have been described [14].

---

**Fig. 1.** SDS-PAGE analysis of DNA-dependent RNA polymerases (RNAP). Electrophoretic separation of the polypeptide subunits of the RNAPs purified from *M. thermoautothrophicum* strains ΔH [18] and Winter [W; [19]] are shown adjacent to the separated subunits of RNAP obtained from *E. coli* (E.c.). The differences observed in the mobilities of the subunits of the RNAPs from the two strains of *M. thermoautothrophicum* support the previous conclusion that these are not very closely related strains [20].

70 ml TMK buffer (50 mM Tris (pH 7.5), 50 mM KCl, 10% v/v glycerol, 10 mM MgCl₂) containing 2 M KCl and passed twice through a French

**Fig. 2.** Binding of pET441 restriction fragments to nitrocellulose filters by *M. thermoautothrophicum* RNAP. [³²P]-end-labeled restriction fragments of pET441 DNA were allowed to bind *M. thermoautothrophicum* ΔH RNAP (see MATERIALS AND METHODS) and the mixtures passaged through nitrocellulose filters. Following filtration, DNA fragments in the filtrates (no RNAP used) and bound to the nitrocellulose filters (with RNAP) were visualized by autoradiography following agarose gel electrophoresis [11]. The results obtained for an *AvaII* digest of pET441 are shown with an arrow indicating the *AvaII* restriction fragment (fragment D) which is preferentially retained on the filter by the presence of *M. thermoautothrophicum* ΔH RNAP. The alphabetical designations of the *AvaII* restriction fragments are given to the left of the autoradiogram. The relative amounts of *AvaII* and *HinII* restriction fragment of pET441 which bound to filters in the presence of RNAP are shown as histograms above *HinII* and *AvaII* restriction maps of pET441 (lower right). The overall organization of pET441 is given (upper right) with the intergenic region separating purE and ORF₆ [14] expanded to show the location of the *AvaII* (A) and *HinII* (H) restriction sites which delineate the RNAP binding region. The *AvaII* restriction fragment D and *HinII* restriction fragment H, which specifically bind RNAP, are emphasized as thick lines.
pET441 Ava II

DNA fragments which bind RNAP
Ava II
Hinf I
RNAP binding region

Relative binding
Hinf I restriction map of pET441

Relative binding
Ava II restriction map of pET441
Fig. 3. Identification of sites of transcription initiation by S1-nuclease protection analysis. Double-stranded single $^{32}$P-end-labeled DNA probes were obtained [11] from the intergenic region of pET44 as shown. The probes were labeled only at the 5' end of the restriction sites marked with stars. Restriction sites are identified as DdeI(D), HinfI(H), AvaII(A), and RsaI(R). The DNA molecules derived from the *DdeI-RsaI probe which were protected from S1-nuclease digestion by transcripts of purE are indicated by arrows in the track of the autoradiogram labeled 'S1.' The *RsaI-DdeI probe was completely digested by S1-nuclease (not shown) indicating that transcripts of ORFc were not present in the RNA preparations obtained from *M. thermoautotrophicum* ΔH. The lane labeled 'G' contains the products of a G-specific Maxam and Gilbert sequencing reaction [17] and the lane labeled 'R' the products of a purine-specific Maxam and Gilbert sequencing reaction of the probe DNA.

The DNA sequence shown in the lower part of the figure is that of the intergenic region between the translation initiation ATG codons (boxed) of the divergent purE and ORFc reading frames. Putative ribosome-binding sequences (RBS) and the Box A consensus promoter sequence are boxed. The apparent sites of transcription initiation (Tnn start) of the purE gene are indicated and the RNAP-binding region, defined as the overlapping region of AvaII fragment D and HinfI fragment H (Fig. 2), is indicated as a stippled bar beneath the sequence.
Restriction fragments of pET441 generated by AvaII or HindI digestion (Fig. 2) were 5'-[^32]P-end-labeled (approximately 1 x 10^4 DPM) and mixed with 1 to 10 μg RNAP in a 100 μl reaction mixture containing 100 mM KCl, 10 mM MgCl₂, 0.1 mM Na₂ EDTA, 0.1 mM DTT, 5 μg BSA and 10 mM Tris (pH 8) [11]. Complexes were allowed to form at 55°C for 5 min; ATP, GTP and CTP were added to final concentrations of 167 μM and incubation continued for 5 min. The reaction mixtures were then cooled on ice and filtered slowly through preboiled nitrocellulose filters. The filters were washed with 500 μl ice-cold buffer and DNA retained by the filters eluted by incubation at 37°C for 1 h in 400 μl 0.2% w/v SDS, 20 mM Tris (pH 8), extracted with 500 μl 50:48:2 phenol:chloroform:isoamyl alcohol, precipitated with ethanol, lyophilized, redissolved in 20 μl gel sample buffer and visualized by autoradiography following electrophoresis through non-denaturing 5% polyacrylamide gels. Exposures of autoradiograms, measured using a Zenith scanning laser densitometer, were quantitated using the GELSCAN (Pharmacia-LKB Biotechnology Inc. Piscataway, NJ) program run on an Apple Ile computer.

3.3. RNA purification and S1-nuclease protection experiments

Crude extracts of M. thermoautotrophicum strain ΔH were obtained from 10 g of cell paste resuspended in 40 ml RNA isolation buffer (80 mM Tris, pH 7.5; 10 mM MgCl₂; 10 mM 2-mercaptoethanol; 0.5% SDS; 10 mM EDTA; 2 mM 1,10-phenanthroline; 0.2 mg heparin/ml) [15] by passage through a French pressure cell at 20000 p.s.i. These extracts were deproteinized twice with 50% phenol:48% chloroform:2% isoamyl alcohol and the nucleic acids precipitated from the resulting aqueous phase by addition of ethanol. RNA and DNA were separated and purified by CsCl:ethidium bromide centrifugation [16]. RNA pellets were dissolved in H₂O, mixed with an equal weight of CsCl and ethidium bromide removed by repeated extraction with 2-propanol (saturated with 10 mM Tris (pH 8), 1 mM EDTA and 5 mM NaCl). After 3-fold dilution with H₂O, RNA was precipitated with 3 volumes of ethanol, redissolved at a concentration of 10 mg/ml and used in S1-nuclease protection experiments [11]. Hybridization reaction mixtures (30 μl) contained 125 μg RNA and singly 5'-[^32]P-end-labeled probe DNAs (approximately 3 x 10^4 DPM; Fig. 3) in 40 mM PIPES, 1 mM Na₂ EDTA, 400 mM NaCl, 8% v/v deionized formamide (pH 6.4) were incubated at 80°C for 10 min, cooled slowly to 47°C and incubated at this temperature for 16 h. S1-nuclease solution (300 μl containing 100-500 units of S1-nuclease in 280 mM NaCl, 4.5 mM ZnSO₄, 50 mM Na acetate (pH 4.6), 6 μg ssDNA) was added and incubation continued at 37°C for 30 min. Radioactively-labeled DNA molecules, protected from S1-nuclease digestion by hybridization, were visualized by autoradiography following their separation by electrophoresis through 6% polyacrylamide sequencing gels.

4. RESULTS

4.1. Localization of RNAP binding sites

Binding of M. thermoautotrophicum ΔH RNAP to AvaII and HindI digests of pET441 resulted, in each case, in the retention on nitrocellulose filters of predominantly one restriction fragment (Fig. 2). These retained fragments overlap by 41 bp within the intergenic region which separates the purE gene from the upstream ORFₖ. This 41 bp sequence must therefore contain sufficient information to direct the binding of M. thermoautotrophicum ΔH RNAP.

4.2. Localization of the site(s) of transcription initiation in vivo

ORFₖ and purE are located on opposite strands of the DNA [14] and if ORFₖ is a bona-fide gene then the two genes would be transcribed in opposite directions, possibly from divergent promoters. S1-nuclease protection experiments were undertaken, using the probes shown in Fig. 3, to locate sites of transcription initiation within the intergenic region which separates ORFₖ and purE. Transcription in the direction of ORFₖ was not detected. Transcription initiation in the direction of purE was located at three adjacent sites 27, 28 and 29 bp upstream of the ATG translation ini-
5. DISCUSSION

Fig. 3 gives the sequence of the intergenic region separating ORF_C and purE [14], indicates the location of the RNAp binding region and the sites detected for transcription initiation upstream of purE. Three consensus sequences, Box A, Box B [4,5,7,8,11,13] and Box II [6] have been implicated as important elements in archaeobacterial promoters. Box A (consensus sequence 5'-TITAA/ATA) sequences are located approximately 25 bp upstream of the site(s) of transcription initiation [4,5,7,8,11,13] and an excellent Box A sequence (ATTAAATA) is found at the expected location upstream of purE, actually within the RNAp binding region. Box B (consensus sequence 5'-TGT/ATA[10]) sequences are found at the site(s) of transcription initiation [4,5,7,8,10,11,13] however the sequence at this location for the purE transcript is TATA. Box II (GTG) sequences are located approximately 17 bp downstream of the site(s) transcription initiation [6] and there is a GTG sequence at this location within the purE transcription, however this is within the sequence AGGTGA which presumably is the ribosome binding sequence (RBS) used to initiate translation of the purE transcript [5,14].

The results of this study support the importance of Box A sequences as a key feature of archaeobacterial promoters. Site specific mutagenesis of the 41 bp AvaiI-HinfI restriction fragment containing the purE Box A sequence (Fig. 3) should delineate exactly which of these base pairs direct M. thermoautotrophicum ΔH RNAp binding to DNA.

ACKNOWLEDGEMENTS

This study was supported by research grant DE-FG02-87ER13731A and contract NOOO14-86-K-02111 from the Office of Naval Research. The close collaboration and expert advice of M. Thomm, facilitated by NATO Collaborative Research grant 0148/85, made this project possible.

REFERENCES