

Identification and Analysis of “Extended –10” Promoters from Mycobacteria

MURALI D. BASHYAM† AND ANIL K. TYAGI*

Department of Biochemistry, University of Delhi South Campus, New Delhi-110021, India

Received 1 October 1997/Accepted 28 February 1998

Earlier studies from our laboratory on randomly isolated transcriptional signals of mycobacteria had revealed that the –10 region of mycobacterial promoters and the corresponding binding domain in the major sigma factor are highly similar to their *Escherichia coli* counterparts. In contrast, the sequences in –35 regions of mycobacterial promoters and the corresponding binding domain in the major sigma factor are vastly different from their *E. coli* counterparts (M. D. Bashyam, D. Kaushal, S. K. Dasgupta, and A. K. Tyagi, *J. Bacteriol.* 178:4847–4853, 1996). We have now analyzed the role of the TGN motif present immediately upstream of the –10 region of mycobacterial promoters. Sequence analysis and site-specific mutagenesis of a *Mycobacterium tuberculosis* promoter and a *Mycobacterium smegmatis* promoter reveal that the TGN motif is an important determinant of transcriptional strength in mycobacteria. We show that mutation in the TGN motif can drastically reduce the transcriptional strength of a mycobacterial promoter. The influence of the TGN motif on transcriptional strength is also modulated by the sequences in the –35 region. Comparative assessment of these extended –10 promoters in mycobacteria and *E. coli* suggests that functioning of the TGN motif in promoters of these two species is similar.

Transcription is the first and often the principal step in regulation of gene expression in bacteria. Comparisons of promoter sequences recognized by the major forms of RNA polymerases of several bacterial species have identified two conserved 6-bp canonical sequences located approximately 10 and 35 bp upstream from the transcription start point. Genetic analysis and protein-DNA interaction studies have confirmed that these two hexameric sequences (called the –10 and –35 regions) are necessary for initiation of transcription. Another region, an AT-rich up element (a target for the C-terminal domain of the α subunit of RNA polymerase), has been reported to enhance the promoter activity 2- to 20-fold (25). An interesting exception to this classical structure of promoters was elucidated by studies on the λ *PRE* promoter and the *Escherichia coli galP1* promoter. It was shown that RNA polymerase could initiate transcription (albeit suboptimally) from promoters lacking a functional –35 region sequence, provided an “extended –10” motif was present in these promoters (6, 14, 16, 23). The extended –10 motif comprises the sequence TGN present immediately upstream of the –10 region. It was also shown that the thermal-energy requirement for open complex formation in an extended –10 promoter was less than that for a conventional –10/–35 promoter (5). Other promoters that have been shown to function as extended –10 promoters include the *E. coli cysG* promoter (3), promoter of the *DpnII* operon of *Streptococcus pneumoniae* (26), the ferredoxin gene promoter of *Clostridium pasteurianum* (11), and the *amyP* gene promoter of *Bacillus subtilis* (29). Kenney and Churchward have reported that the TGN motif present upstream of the –10 hexamer can play a role in the activity of the *rpsL* promoter of *Mycobacterium smegmatis* (15).

We had earlier initiated studies on randomly isolated transcriptional signals of mycobacteria (7) and shown that the –10 region of these promoters and the corresponding binding domain in the principal sigma factor of mycobacteria (σ^A) are almost identical to those of *E. coli* (2). However, the –35 region of mycobacterial promoters and the corresponding binding domain in the principal sigma factor were found to be vastly different from those of *E. coli* (2). We also reported that the –35 region of mycobacterial promoters can tolerate a greater variety of sequences compared to other bacterial promoters, owing to the presence of multiple constitutive sigma factors having different or overlapping binding specificities for the –35 region of promoters (2). We have carried out analysis of mycobacterial promoters that belong to the class of extended –10 promoters. We show that the TGN motif located upstream of the –10 region in mycobacterial promoters is an important determinant of transcriptional activity. Our studies suggest that the TGN motif may play similar roles in the initiation of transcription in mycobacteria and *E. coli*.

Mycobacterial promoters containing the TGN motif immediately upstream of the Pribnow box. We have analyzed the sequences of 59 mycobacterial promoters (for which the transcriptional start points have been experimentally determined in our laboratory or elsewhere) for the presence of the TGN motif immediately upstream of the –10 hexamer. These included 16 *M. smegmatis* promoters and 12 *M. tuberculosis* promoters from our mycobacterial promoter library (2). The other promoters included in the analysis belonged to the *mpb70* gene of *M. bovis* BCG (20), the *rpsL* gene of *M. smegmatis* (15), the *recA* genes of *M. tuberculosis* and *M. smegmatis* (21, 22), the DNA gyrase genes of *M. smegmatis* and *M. tuberculosis* (18, 19), the *purC* and *purL* genes of *M. tuberculosis* (13), the 18-kDa gene of *M. leprae* (8), the *oxyR* and *ahpC* genes of *M. leprae* (9), and the *repA* gene of plasmid pAL5000 from *M. fortuitum* (27). The remaining 16 promoters have been listed earlier (2). Thirteen of the 59 promoters analyzed in this study contained the TGN motif (Fig. 1). These included six promoters from our *M. smegmatis* promoter library and two from our *M. tuberculosis* promoter library in addition to the *hsp60* P2

* Corresponding author. Mailing address: Department of Biochemistry, University of Delhi South Campus, Benito Juarez Rd., New Delhi-110021, India. Phone: 91-11-6881970. Fax: 91-11-6885270 or 6886427. E-mail: AKT@DUSC.ernet.in.

† Present address: Eukaryotic Gene Expression Laboratory, National Institute of Immunology, Aruna Asaf Ali Marg, New Delhi-110067, India.

A Promoter		TSP
	-50 -40 -30 -20 -10	
S5	AACAAGATTCCGTTAATCGTCTGGTGGAGCTGGTGGTAAAGCTTGATCC	G
S6	GATCGATTTTAAATTTTGATAGAGTGCARAATA	A
S16	TCCACCGGAACCCCTTCGGCGTCCCGCTTTTCCCTGGTTATAATCGGC	G
S19	TTTGATGTAGCCAAAGGCTCTCACCACCTGAGCCATGATAGTATCCATCC	C
S21	ACATGGCATTTTTCATTTAAACAGGACTCAGGTGGTATCGTTGACATCG	A
S119	GATCAAGAATCCAAATGATTTGTTAAACCAATTAAT	G
T101	AGCGATCGCAGCCGACGTGATACCTGACCGTGTGTGATAGTGTGGCGGC	A
T129	ACTCGCGGCAGATTACGCCGACGGTTCCTGGCGTGGTTCAATATCGCCG	T
B		
rpsL	CGGTGCAAGAGTTTGGTTCTGCGCGTTCGGCCCTGGTATTTGTTGGGATC	G
hsp60P2	CGGTGCGGGCTTCTTGCACTCGGCATAGGCGAGTGCATAGAAATAACGGT	G
repA	GAGCTCGTGTGGACCATACACCGGTGATTAATCGTGGTCTACTACCAAG	C
18kDa	CTGTCTATCACAACTTGCAATCAATATCGACCAGTCTATATCAAACTG	A
rzn	AGTCAACCGGGAGTTGATCTCTCTGCGATCTGTTAATCTGGCTG	G

FIG. 1. Alignment of putative extended -10 promoters of mycobacteria. Fifty-nine mycobacterial promoters for which the TSP had been experimentally determined were analyzed. The sequences of the 13 promoters which contained the TGN motif are shown. (A) Promoters from our mycobacterial promoter library. S5, S6, S16, S19, S21, and S119 are from *M. smegmatis*, and T101 and T129 are from *M. tuberculosis*. (B) *rpsL* gene promoter from *M. smegmatis*, *hsp60* P2 promoter from *M. bovis*, *repA* promoter from the plasmid pAL5000 of *M. fortuitum*, and promoters of the 18-kDa gene and the rRNA operon of *M. leprae*. The extended -10 sequences are underlined.

promoter of *M. bovis* BCG (28), the *rpsL* promoter of *M. smegmatis* (15), the *repA* promoter from the *M. fortuitum* plasmid pAL5000 (27), the rRNA gene promoter of *M. leprae* (30), and the 18-kDa gene promoter of *M. leprae* (8). Thus, based on a small sample size of 59, 22% of mycobacterial promoters contain the TGN motif. Analysis of 183 promoters from various species of gram-positive bacteria (11, 12, 26) reveals that frequency of occurrence of the TGN motif in these promoters is around 60%. In *E. coli* promoters, the TGN motif occurs with a frequency of about 16% (16).

Functional analysis of the S16 promoter of *M. smegmatis*. We chose the S16 promoter for further analysis because (i) it contained the perfect extended -10 motif (TGTTATAAT), (ii) it represented one of the strongest promoters in our library, and (iii) it exhibited suboptimal but significant activity in *E. coli*, an organism in which the importance of the TGN motif has been well demonstrated. We have earlier shown that the -10 region in mycobacterial promoters is essential for initia-

tion of transcription (2). In order to determine its importance in the putative extended -10 promoters of mycobacteria, we synthesized two truncated derivatives of S16—S16.2 and S16.4—and cloned them separately into pSD9, generating pS16.2 and pS16.4, respectively (Table 1). The cloning strategy is depicted in Fig. 2A. The S16.2 derivative harbored both the -10 and the -35 regions, whereas the S16.4 derivative did not harbor the -10 region. In addition, three bases were changed to incorporate an *SphI* restriction site at the -18 position in both derivatives. Since (i) for mycobacterial promoters including S16, the same transcription start point (TSP) is used in mycobacteria and *E. coli* (2), (ii) the conserved hexameric sequence TATAAT is located 10 bp upstream of the TSP in both *E. coli* and mycobacterial promoters (2), and (iii) the -10 binding domain in the sigma factor of *E. coli* is identical to the corresponding domain in the mycobacterial sigma factor (2), the effects of deletion of the -10 region should be similar in the two hosts. *M. smegmatis* and *E. coli* were transformed with pS16.2 and pS16.4 separately as described earlier (7). Deletion of the -10 region completely abolished promoter activity in both mycobacteria and *E. coli* (pS16.4 in Fig. 2B). This result confirmed our earlier observation that the -10 region is an important functional determinant of promoter activity in mycobacteria as in other eubacteria. In order to determine the role of the TGN motif in mycobacterial promoters, we decided to study the effect of mutating this motif. For this purpose, pS16.2(TG⁻) was generated by using synthetic oligonucleotides in which TG of the extended -10 motif of pS16.2 was replaced by CC. *M. smegmatis* and *E. coli* were separately transformed with the constructs pS16.2 and pS16.2(TG⁻), and the chloramphenicol acetyltransferase (CAT)-specific activities supported by each of these constructs were determined as described earlier (7). As shown in Fig. 3, mutation of the TGN motif resulted in a fourfold reduction in the CAT activity supported by the promoter in mycobacteria. In *E. coli*, there was an eightfold decrease in the activity when the TGN motif was mutated (Fig. 3). These results suggested that the TGN motif may be an important determinant of transcriptional strength in mycobacteria as is the case in *E. coli*. To determine whether the effect of mutation of the TGN motif would be different for different sequences in the -35 region, six modified pS16.2 constructs (pS16.2A to pS16.2F) and six modified pS16.2(TG⁻) constructs [pS16.2(TG⁻)A to pS16.2(TG⁻)F] were generated as described in Table 1. These six pS16.2 con-

TABLE 1. Promoter constructs used in this study

Plasmid(s)	Description
pS16.2.....	Synthetic S16 promoter, representing promoter sequences from -42 to +5, cloned into pSD9
pS16.4.....	Synthetic S16 promoter, representing sequences from -42 to -17, cloned into pSD9
pS16.2(TG ⁻).....	pS16.2 with CC instead of TG 1 bp upstream of the -10 region of the promoter
pS16.2A to -F and pS16.2(TG ⁻)A to -F.....	Contain various fragments (44 to 276 bp) derived from the multiple cloning region of plasmid pGEM5Zf(+) cloned into the <i>SphI</i> restriction site of pS16.2 [or pS16.2(TG ⁻)], giving rise to six mosaic pS16.2 [or pS16.2(TG ⁻)] constructs containing identical -10 but different -35 regions
pT125(TG ⁻).....	Original T125 promoter from pSD7.T125, which does not contain the TGN motif, cloned into pSD9
pT125(TG ⁻).2.....	Construct containing the T125 promoter sequence till the -29 position and lacking the -10 region
pT125.....	Synthetic T125 promoter fragment, containing the sequence TG instead of GA 1 bp upstream of the -10 region sequence, cloned into pT125(TG ⁻) as described in the text
pT125A to -F and pT125(TG ⁻)A to -F.....	Contain various fragments (44 to 276 bp) derived from the multiple cloning region of plasmid pGEM5Zf(+) cloned into the <i>ApaI</i> restriction site of pT125 [or pT125(TG ⁻)], giving rise to six mosaic pT125 [or pT125(TG ⁻)] constructs containing identical -10 but different -35 regions

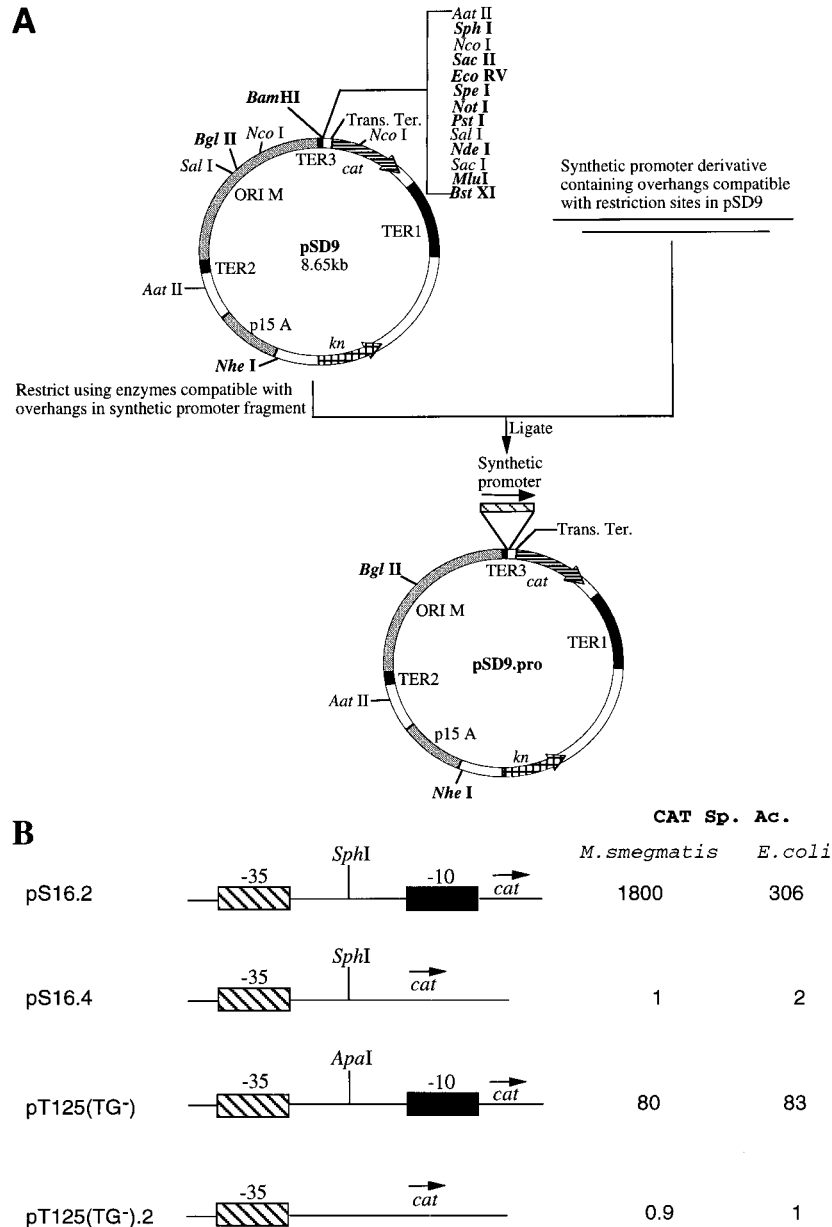


FIG. 2. (A) Strategy for cloning of various modified promoter derivatives in pSD9. pSD9 was generated by cloning the end-repaired *ApaI*-*NsiI* fragment derived from the multiple cloning region of pGEM5Zf(+) into the *PstI* restriction site present upstream of the reporter gene encoding CAT in pSD7; the other *PstI* restriction site present in the inessential region downstream to TER1 in pSD7 was inactivated prior to this cloning. The synthetic promoter fragments were annealed in 1× NEB 2 buffer (10 mM Tris-HCl [pH 7.9], 10 mM MgCl₂, 50 mM NaCl, 1 mM dithiothreitol) and cloned into pSD9 restricted with appropriate enzymes. Trans. Ter., transcription terminus. (B) Functional dissection of the S16 promoter of *M. smegmatis* and the T125 promoter of *M. tuberculosis*. The construct pS16.2 represents the synthetic S16.2 promoter, containing sequences from -42 to +5, cloned into pSD9, pS16.4 represents the synthetic S16.4 promoter, containing sequences from -42 to -17 cloned into pSD9, pT125(TG⁻) represents the original T125 promoter from pSD7.T125 cloned into pSD9, and pT125(TG⁻).2 represents the T125 promoter harboring only the -35 region cloned into pSD9. The CAT specific activity (Sp. Ac.) supported by each construct in *M. smegmatis* and *E. coli* is expressed as nanomoles of chloramphenicol converted into its acetylated derivatives per minute per milligram of protein.

structs as well as the six pS16.2(TG⁻) constructs contained identical -10 regions but different -35 regions. *M. smegmatis* and *E. coli* were transformed by using each of these 12 constructs separately. The CAT specific activities supported by the pS16.2 constructs were compared to those of the corresponding pS16.2(TG⁻) constructs to determine the contribution of the TGN motif in the context of different sequences in the -35 region. The results are given in Fig. 3. All the pS16.2(TG⁻) constructs supported lower levels of CAT activity in mycobac-

teria compared to their corresponding pS16.2 derivatives. There was a 13-fold reduction in the case of pS16.2C, whereas in the case of pS16.2E the decrease was twofold. Similar results were obtained when the constructs were tested in *E. coli* (Fig. 3). In each case, there was a decrease in the activity when the TGN motif was mutated. The level of reduction varied, ranging from about eightfold for pS16.2, pS16.2A, and pS16.2C to about twofold for pS16.2E. These results further substantiate the importance of the TGN motif in initiation of transcription

Promoter construct	Diagram	CAT Sp. Act.		
		TG ⁺	TG ⁻	
pS16.2		<i>M. sm.</i>	1800	400
		<i>E. coli</i>	306	37
pS16.2A		<i>M. sm.</i>	2000	500
		<i>E. coli</i>	716	86
pS16.2B		<i>M. sm.</i>	4157	1070
		<i>E. coli</i>	96	16
pS16.2C		<i>M. sm.</i>	914	17
		<i>E. coli</i>	223	28
pS16.2D		<i>M. sm.</i>	1028	148
		<i>E. coli</i>	199	29
pS16.2E		<i>M. sm.</i>	2474	96
		<i>E. coli</i>	96	29
pS16.2F		<i>M. sm.</i>	500	60
		<i>E. coli</i>	962	133

FIG. 3. A comparative assessment of CAT activities supported by various modified pS16.2 constructs with and without the TGN motif in mycobacteria and *E. coli*. The unique *Sph*I restriction site in pS16.2 and pS16.2(TG⁻) was used to generate the various modified promoter derivatives as described in Table 1. The extended -10 sequence (identical for all constructs) and the -35 region (different for each construct) are indicated. The CAT specific activity (Sp. Act.) was determined as described elsewhere (7) and is expressed as nanomoles of chloramphenicol converted into its acetylated derivatives per minute per milligram of protein. *M. sm.*, *M. smegmatis*.

in mycobacteria. Additionally, in both mycobacteria and *E. coli*, the levels of reduction following the mutation in the TGN motif were different for different constructs depending on the sequences generated in the -35 region due to the various insertions.

Functional analysis of the *M. tuberculosis* promoter T125. In the case of the S16 promoter of *M. smegmatis*, we had mutated the TGN motif to evaluate its importance in transcription initiation. Following this, we proceeded to determine the effect of introducing the TGN motif in a conventional -10/-35 promoter that does not harbor the TGN motif. This was done to ensure that the results obtained with the *M. smegmatis* S16 promoter were applicable to other promoters of mycobacteria. For this purpose, we chose the *M. tuberculosis* promoter T125 because it contained a near-perfect -10 region sequence (TA TTAT) and provided a unique *Apa*I restriction site between the -10 and -35 regions of the promoter for carrying out manipulations (2). In order to facilitate modifications in the promoter, the T125 promoter fragment from pSD7.T125 was cloned into pSD9. This promoter construct lacking the TGN motif was designated pT125(TG⁻) (Table 1). We first determined the effect of deletion of the -10 region. The *Apa*I restriction site was used to delete promoter sequences downstream of the -29 position, and the resulting construct was designated pT125(TG⁻).2 (Table 1). As shown in Fig. 2B, deletion of the -10 region completely abolished promoter activity in both mycobacteria and *E. coli*. For introducing the TGN motif, synthetic oligonucleotides containing the sequence TG instead of GA 1 bp upstream of the -10 region were used, resulting in the generation of pT125 (Table 1). *M. smegmatis* and *E. coli* were transformed with pT125 and pT125(TG⁻)

separately, and the CAT activity supported by each construct was determined. Introduction of the TGN motif resulted in threefold and fivefold enhancements in CAT activity in *M. smegmatis* and *E. coli*, respectively (Fig. 4). These results suggest that the TGN motif may be an important determinant of promoter activity in mycobacteria. In order to determine the role of the TGN motif vis-à-vis different -35 regions in the promoter T125, six modified pT125 constructs (pT125A to pT125F) and six modified pT125(TG⁻) constructs [pT125(TG⁻)A to pT125(TG⁻)F] containing identical -10 regions but different -35 regions were generated (Table 1). *M. smegmatis* and *E. coli* were separately transformed with all these mosaic promoter constructs. The CAT activities supported by each construct in *M. smegmatis* and *E. coli* are listed in Fig. 4. Introduction of the TGN motif into various modified pT125(TG⁻) constructs significantly raised the levels of CAT activity in *M. smegmatis* (Fig. 4). The level of enhancement varied from 75-fold in the case of pT125E to about 2-fold in the case of pT125A. Similar results were obtained when the CAT activities supported by the various constructs were determined for the respective *E. coli* transformants. The level of enhancement for different constructs in *E. coli* varied from 45-fold in the case of pT125E(TG⁻) to about 4-fold in the case of pT125B(TG⁻) and pT125D(TG⁻) (Fig. 4). In fact, the CAT activity supported by pT125E(TG⁻) in mycobacteria was extremely low, comparable to the basal activity supported by pSD9 (1 nmol/min/mg of protein). Introduction of the TGN motif upstream of the -10 region raised the activity to 188 nmol/min/mg of protein. Therefore, as observed with the S16 promoter, the effect of the TGN motif on transcriptional ac-

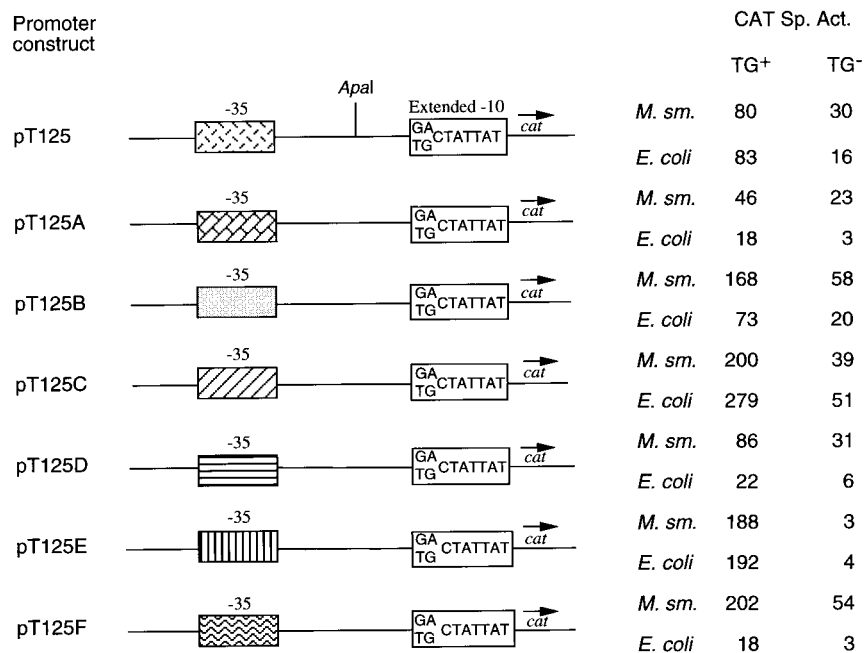


FIG. 4. A comparative assessment of CAT activities supported by various modified pT125 constructs with and without the TGN motif in mycobacteria and *E. coli*. The unique *Apa*I restriction site in pT125 and pT125(TG⁻) was used to generate the various modified promoter derivatives as described in Table 1. The extended -10 sequence (identical for all constructs) and the -35 region (different for each construct) are indicated. The CAT specific activity (Sp. Act.) was determined as described elsewhere (7) and is expressed as nanomoles of chloramphenicol converted into its acetylated derivatives per minute per milligram of protein. *M. sm.*, *M. smegmatis*.

tivity of a promoter varies with different sequences in the -35 region, in both *M. smegmatis* and *E. coli*.

Concluding remarks. We have screened 59 mycobacterial promoters for which the TSPs have been experimentally determined and carried out a functional analysis of the S16 promoter of *M. smegmatis* and the T125 promoter of *M. tuberculosis*. Our results suggest that the TGN motif may be an important determinant of transcriptional strength in mycobacteria as in other bacterial species. Therefore, the mechanism of basal transcription in mycobacteria appears similar to that of conventional bacterial promoters. Presently, it seems unlikely that the mycobacterial extended -10 promoters are responsible for transcription of a specific set of genes. Genes like *rpsL* (which encodes the S12 ribosomal protein) of *M. smegmatis*, *hsp60* of *M. bovis*, *repA* from pAL5000, the rRNA gene of *M. leprae*, and the 18-kDa gene of *M. leprae* do not belong to any particular category. However, presence of the extended -10 promoter may be required in particular regions of the bacterial chromosome having sequence constraints, where it may be difficult to maintain two specific hexameric sequences (e.g., within an open reading frame), as has been already suggested for the *E. coli* *cysG* promoter (3). Another role of extended -10 promoters could be to maintain a basal level of transcription in the case of promoters that contain a weak -35 region and are regulated by protein-DNA interactions in the -35 region (as in the case of the *galP1* promoter of *E. coli*) (6). One possible advantage of the TGN motif could be to facilitate transcription initiation at cold temperatures or when the sigma factor is proteolytically cleaved (under stress conditions). As has been suggested earlier, the present transcription machinery may have evolved from a primitive apparatus containing just the extended -10 motif in the promoters and the -10 binding domain of the σ factor along with a catalytic subunit constituting the RNA polymerase (4). One may need to look at the -10 region not as a hexamer but as a nonamer, and the

presence of perfect nucleotides in all the nine positions (as opposed to six) enables RNA polymerase to undertake transcription initiation due to the extended contact it makes with DNA in the -10 region (nine bases equal almost one full turn of the DNA helix), even though it may make poor contacts with the -35 region. Kumar and coworkers have actually shown that RNA polymerase can initiate transcription from an extended -10 promoter in the absence of the -35 binding domain of the sigma factor (16). Transcription initiation based on such an extended -10 region alone, however, may not represent the optimal strength of the promoter. As shown in the present study, the sequences present in the -35 region modulate initiation of transcription even in the presence of a TGN motif. Chan and Busby as well as Keilty and Rosenberg have shown that -35 region sequences can apparently modulate overall promoter strength in the λ *PRE* extended -10 promoter (6, 14). With suppression genetics, it has been shown that the region of σ^{70} immediately downstream of the region 2.4 (termed region 2.5) is actually responsible for contacting the TGN motif (1). We have compared the amino acid sequence of region 2.5 of σ^{70} with that of the corresponding region of the *M. smegmatis* sigma factors σ^A and σ^B (Fig. 5) (24). The two amino acids, glutamic acid and histidine, implicated in interaction with the TGN motif in *E. coli* (1) are also conserved in σ^A and σ^B . The region of mycobacterial σ factors corresponding to the 2.5 region of *E. coli* σ^{70} (spanning 22 amino acids) contains only four nonconserved substitutions. Also, the amino acids that exhibited a high frequency of occurrence, as deduced from a comparison of 30 σ factor sequences (4), are conserved in both the σ factors of *M. smegmatis*. The amino acid sequences in the 2.5 regions of σ^A and σ^B of *M. tuberculosis* and *M. smegmatis* are also identical (10). These observations further substantiate the similarities between the transcriptional machineries of *E. coli* and mycobacteria, as far as the extended -10 region is concerned. Foot-

	1	2	3	4	5	6	7	8	9	10	11
RpoD	V	H	M	I	E	T	I	N	K	L	N
MysA	-	-	-	-	-	V	-	-	-	-	G
MysB	-	-	L	-	-	Q	V	-	-	-	-
Consensus	V	H	.	V	E	.	.	N	.	.	.
	12	13	14	15	16	17	18	19	20	21	22
RpoD	R	I	S	R	Q	M	L	Q	E	M	G
MysA	-	-	Q	-	E	-	-	-	D	L	-
MysB	-	-	K	-	E	-	H	-	N	L	-
Consensus	R	.	.	R	.	L

FIG. 5. Comparison of the amino acid sequences in the region 2.5 of the sigma factors of *M. smegmatis* (σ^A and σ^B) and the principal sigma factor of *E. coli* (RpoD). The two amino acids previously implicated in making specific contacts with the TGN motif (filled circles), identical amino acids (dashes), conserved substitutions (plain letters), and nonconserved substitutions (bold-faced letters) are indicated. The amino acid sequences are from reference 17 for RpoD of *E. coli* and from reference 24 for σ^A and σ^B of *M. smegmatis*. The consensus sequence is from reference 4, wherein the amino acids that exhibit a high-frequency occurrence (based on a compilation of 30 sigma factors) are indicated. Conserved substitutions are defined as the following groups: I, L, M, and V; A and G; S and T; K, H, and R; D, E, N, and Q; F, Y, and W; C; and P. The numbers do not represent actual amino acid positions.

printing analysis and studies on methylation and ethylation interference in combination with site-specific mutagenesis would further enhance our understanding of the role of the TGN motif in initiation of transcription in mycobacteria and help in the development of tools for high-level expression of genes in mycobacteria.

We thank Manisha Sharma for excellent technical assistance and J. S. Tyagi and Shruti Jain for critical reading of the manuscript.

This work was supported by a financial grant from the Department of Biotechnology of India. M.D.B. is thankful to the Council of Scientific and Industrial Research and the Department of Biotechnology for financial assistance.

REFERENCES

- Barne, K. A., J. A. Brown, S. J. E. Busby, and S. D. Minchin. 1997. Region 2.5 of the *Escherichia coli* RNA polymerase σ^{70} subunit is responsible for the recognition of the 'extended -10' motif at promoters. *EMBO J.* **16**:4034-4040.
- Bashyam, M. D., D. Kaushal, S. K. DasGupta, and A. K. Tyagi. 1996. A study of the mycobacterial transcriptional apparatus: identification of novel features in promoter elements. *J. Bacteriol.* **178**:4847-4853.
- Belyaeva, T., L. Griffiths, S. Minchin, J. Cole, and S. Busby. 1993. The *Escherichia coli* *cysG* promoter belongs to the 'extended -10' class of bacterial promoters. *Biochem. J.* **296**:851-857.
- Bown, J. A., K. A. Barne, S. D. Minchin, and S. J. W. Busby. 1997. Extended -10 promoters. *Nucleic Acids Mol. Biol.* **11**:41-52.
- Burns, H., and S. Minchin. 1994. Thermal energy requirement for strand separation during transcription initiation: the effect of supercoiling and extended protein-DNA contacts. *Nucleic Acids Res.* **22**:3840-3845.
- Chan, B., and S. Busby. 1989. Recognition of nucleotide sequences at the *Escherichia coli* galactose operon P1 promoter by RNA polymerase. *Gene* **84**:227-236.
- DasGupta, S. K., M. D. Bashyam, and A. K. Tyagi. 1993. Cloning and assessment of mycobacterial promoters by using a plasmid shuttle vector. *J. Bacteriol.* **175**:5186-5192.
- Dellagostin, O. A., G. Esposito, L.-J. Eales, J. W. Dale, and J. McFadden. 1995. Activity of mycobacterial promoters during intracellular and extracellular growth. *Microbiology* **141**:1785-1792.
- Dhandayuthapani, S., M. Mudd, and V. Deretic. 1997. Interactions of OxyR with the promoter region of the *oxyR* and *ahpC* genes from *Mycobacterium leprae* and *Mycobacterium tuberculosis*. *J. Bacteriol.* **179**:2401-2409.
- Doukhan, L., M. Predich, G. Nair, O. Dussurget, I. Mandic-Mulec, S. T. Cole, D. R. Smith, and I. Smith. 1995. Genomic organization of the mycobacterial sigma gene cluster. *Gene* **165**:67-70.
- Graves, M. C., and J. C. Rabinowitz. 1986. *In vivo* and *in vitro* transcription of the *Clostridium pasteurianum* ferredoxin gene. *J. Biol. Chem.* **261**:11409-11415.
- Helman, J. 1995. Compilation and analysis of *Bacillus subtilis* σ A-dependent promoter sequences: evidence for extended contact between RNA polymerase and upstream promoter DNA. *Nucleic Acids Res.* **23**:2351-2360.
- Jackson, M., F.-X. Berthet, I. Otal, J. Rauzier, C. Martin, B. Gicquel, and C. Guilhot. 1996. The *Mycobacterium tuberculosis* purine biosynthetic pathway: isolation and characterization of the *purC* and *purL* genes. *Microbiology* **142**:2439-2447.
- Keilty, S., and M. Rosenberg. 1987. Constitutive function of a positively regulated promoter reveals new sequences essential for activity. *J. Biol. Chem.* **262**:6389-6395.
- Kenney, T. J., and G. Churchward. 1996. Genetic analysis of the *Mycobacterium smegmatis* *rpsL* promoter. *J. Bacteriol.* **178**:3564-3571.
- Kumar, A., R. A. Malloch, N. Fujita, D. A. Smillie, A. Ishihama, and R. S. Hayward. 1993. The minus 35-recognition region of *Escherichia coli* sigma 70 is essential for initiation of transcription at an "extended minus 10" promoter. *J. Mol. Biol.* **232**:406-418.
- Lonetto, M., M. Gribskov, and C. A. Gross. 1992. The σ^{70} family: sequence conservation and evolutionary relationships. *J. Bacteriol.* **174**:3843-3849.
- Madhusudan, K., and V. Nagaraja. 1995. *Mycobacterium smegmatis* DNA gyrase: cloning and overexpression in *Escherichia coli*. *Microbiology* **141**:3029-3037.
- Madhusudan, K., V. Ramesh, and V. Nagaraja. 1994. Molecular cloning of *gyrA* and *gyrB* genes of *Mycobacterium tuberculosis*: analysis of nucleotide sequence. *Biochem. Mol. Biol. Int.* **33**:651-660.
- Matsuo, T., S. Matsumoto, N. Ohara, H. Kitaura, A. Mizuno, and T. Yamada. 1995. Differential transcription of the MPB70 genes in two major groups of *Mycobacterium bovis* BCG substrains. *Microbiology* **141**:1601-1607.
- Movahedzadeh, F., M. J. Colston, and E. O. Davis. 1997. Determination of DNA sequences required for regulated *Mycobacterium tuberculosis* RecA expression in response to DNA-damaging agents suggests that two modes of regulation exist. *J. Bacteriol.* **179**:3509-3518.
- Papavinasundaram, K. G., F. Movahedzadeh, J. T. Keer, N. G. Stoker, M. J. Colston, and E. O. Davis. 1997. Mycobacterial *recA* is cotranscribed with a potential regulatory gene called *recX*. *Mol. Microbiol.* **24**:141-153.
- Ponnambalam, S., C. Webster, A. Bingham, and S. Busby. 1986. Transcription initiation at the *Escherichia coli* galactose operon promoters in the absence of the normal -35 region sequences. *J. Biol. Chem.* **261**:16043-16048.
- Predich, M., L. Doukhan, G. Nair, and I. Smith. 1995. Characterization of RNA polymerase and two sigma-factor genes from *Mycobacterium smegmatis*. *Mol. Microbiol.* **15**:355-366.
- Ross, W., K. Gosink, J. Solomon, K. Igarashi, C. Zou, A. Ishihama, K. Severinov, and R. Gourse. 1993. A third recognition element in bacterial promoters: DNA binding by the α subunit of RNA polymerase. *Science* **262**:1407-1413.
- Sabelnikov, A. G., B. Greenberg, and S. A. Lacks. 1995. An extended -10 promoter alone directs transcription of the *DpnII* operon of *Streptococcus pneumoniae*. *J. Mol. Biol.* **250**:144-155.
- Stolt, P., and N. G. Stoker. 1996. Protein-DNA interactions in the ori region of the *Mycobacterium fortuitum* plasmid pAL5000. *J. Bacteriol.* **178**:6693-6700.
- Stover, C. K., V. F. de la Cruz, T. R. Fuerst, J. E. Burlein, L. A. Benson, L. T. Bennett, G. P. Bansal, J. F. Young, M. H. Lee, G. F. Hatfull, S. B. Snapper, R. G. Barletta, W. R. Jacobs, Jr., and B. R. Bloom. 1991. New use of BCG for recombinant vaccines. *Nature (London)* **351**:456-460.
- Voskuil, M. I., K. Voepel, and G. H. Chambliss. 1995. The -16 region, a vital sequence for the utilization of a promoter in *Bacillus subtilis* and *Escherichia coli*. *Mol. Microbiol.* **17**:271-279.
- Yuan-en, J., M. J. Colston, and R. A. Cox. 1994. Nucleotide sequence and secondary structures of precursor 16S rRNA of slow-growing mycobacteria. *Microbiology* **140**:123-132.