T7 promoter contacts essential for promoter activity *in vivo*

Richard A. Ikeda, Cathleen M. Ligman and Sakuntala Warshamana

School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, GA 30332-0400, USA

Received January 24, 1992; Revised and Accepted April 7, 1992

ABSTRACT

T7 RNA polymerase promoters consist of a highly conserved 23 base-pair sequence that spans the site of the initiation of transcription (+1) and extends from -17 to +6. To determine the bases within the T7 consensus promoter that are essential for promoter function a library of mutant T7 promoters was constructed, and the *in vivo* activity of the mutant promoters was correlated to their sequence. The library of mutant promoters was created by randomly mutagenizing the T7 ϕ 10 promoter between positions -22 and +6 during the synthesis of oligonucleotides containing the ϕ 10 promoter. The mutagenized oligonucleotides were then ligated to a promoterless chloramphenicol acetyl transferase gene creating a plasmid (pCM-X#) that can potentially express chloramphenicol acetyl transferase in the presence of T7 RNA polymerase. *E. coli* containing pCM-X# and a second compatible plasmid carrying T7 gene 1 (T7 RNA polymerase) were screened for chloramphenicol resistance or chloramphenicol sensitivity. The point mutations that were found to inactivate a T7 promoter are a C to A or G substitution at -7, a T to A substitution at -8, a C to A, T, or G substitution at -9, and a G to T substitution at -11.

INTRODUCTION

Bacteriophage T7 encodes its own highly specific RNA polymerase that initiates transcription of the T7 genome from 17 phage promoters. Comparison of the 17 T7 RNA polymerase promoters yields a highly conserved 23 base pair consensus sequence that spans the site of the initiation of transcription (+1) and extends from -17 to +6 (1–4). Functionally, these 17 T7 promoters fall into three classes: replication promoters, class II promoters, and class III promoters. The two replication promoters are apparently involved in T7 replication, while the ten class II and the five class III promoters are responsible for transcription of the T7 middle and late genes, respectively (1,2). The five class III promoters and the replication promoter ϕ OR are identical to the T7 consensus promoter and show additional homology from -22 to -18; conversely, the ten class II promoters and the replication promoter ϕ OL are not identical to the T7 consensus promoter and contain from two to seven non-consensus bases (2).

Recognition of the T7 promoter by T7 RNA polymerase requires specific contacts between the promoter and the protein

(5–7). Simple footprinting studies have shown that T7 RNA polymerase protects the promoter from -21 to -3 (8–10), while high resolution footprinting has shown two protected regions across the minor grooves of one face of the promoter helix (11). The importance of both minor and major groove contacts within the T7 promoter has been demonstrated by studies of promoters containing modified nucleotides (12–14). These studies showed that base modifications within the major and minor groove of the promoter prevented binding and utilization of the promoter.

Studies of mutant T7 promoters have yielded even more detailed information about the relationship of structure to function in the T7 promoter, and the recognition of the promoter by T7 RNA polymerase. The activities of specific T7 promoter mutants suggest that the promoter is organized into two domains, an initiation domain from -4 to +5, and a binding domain from -5 to -12 (15,16). Single base changes in the binding domain of the promoter reduce the efficiency of promoter binding, but have little effect on the initiation of transcription. In contrast, single base changes in the initiation domain have little effect on promoter binding, but reduce the rate of the initiation of transcription. The remainder of the promoter from -12 to -17 seems to be almost dispensable. Much of this leftward section of the T7 promoter can be deleted or changed without abolishing promoter activity (17,18); however, optimal T7 promoter activity appears to require contacts within the -12 to -17 region (19).

These correlations of T7 promoter sequence to promoter activity show that some of the bases within the 23 base pair consensus promoter are more important for promoter activity than other bases. In an analysis of the T7 promoter sequence, Schneider and Stormo suggested that bases at positions -12, -9, -8, -7, and -5 of the T7 promoter carry much of the information encoded by the promoter (20); however, this analysis was indirect. Schneider and Stormo could not find strong T7 promoters with mutations at positions -12, -9, -8, -7, and -5; consequently, they concluded that these bases are important for promoter function.

The questions that remain are: whether specific bases within the conserved T7 consensus promoter are required for promoter activity, and what are the identity of these bases? To directly determine what bases within the T7 consensus promoter are necessary for promoter function we describe in the following work the construction of a library of mutant T7 promoters, the determination of the activities of the mutant promoters, and the compilation of the sequences of the mutant promoters.

EXPERIMENTAL PROCEDURES

Materials

Cell Strains. *E. coli* DH5 and *E. coli* DH5 α were obtained from Gibco/BRL, and *E. coli* JM101 was obtained from Stratagene.

Enzymes. Restriction endonucleases, T4 ligase, T4 polynucleotide kinase, calf intestine alkaline phosphatase, and large fragment *E. coli* DNA polymerase I were purchased from New England Biolabs, Boehringer Mannheim Biochemicals, and United States Biochemical. Sequenase Version 2.0 was purchased from United States Biochemical.

T7 RNA polymerase was purified by standard methods from *E. coli* HMS12 harboring plasmids pGP1-1 and pGP1-5 (21,22). The T7 RNA polymerase was estimated to have a specific activity of 483,000 units/mg (23).

Chemicals. Buffers, acrylamide, N,N'-methylene-bis-acrylamide, ammonium persulfate, N,N,N',N'-tetramethylethylenediamine, and urea were electrophoresis grade. X-gal, and isopropyl-1-thio- β -D-galactoside (IPTG) were molecular biology grade. Media was from Difco. Antibiotics were from Sigma, and all other chemicals were reagent grade.

Media. LB media consists of 10.0 g of tryptone, 5.0 g of yeast extract, and 10.0 g of NaCl per liter of media; the pH of the media was adjusted to 7.5 with NaOH. SOC media consists of 20 g/l tryptone, 5.0 g/l yeast extract, 0.6 g/l NaCl, 0.5 g/l KCl, 10 mM MgCl₂, 10 mM MgSO₄, and 20 mM glucose; the pH of the media was adjusted to 7.5 with NaOH.

Nucleoside 5'-triphosphates. [α - ^{35}S]-dATP (800 Ci/mmol) and [2,8- ^3H]-ATP (25–40 Ci/mmol) were purchased from Dupont/New England Nuclear Research Products. Deoxynucleoside triphosphates and dideoxynucleoside triphosphates were obtained from United States Biochemical.

Plasmids. Plasmids pKK232-8 (24) and pKK223-3 (25) were purchased from Pharmacia/LKB Biotechnology. Plasmids pGP1-5 and pGP1-1 (23) were a gracious gift from Dr. Stanley Tabor, Harvard Medical School and Dr. Charles C. Richardson, Harvard Medical School.

Oligodeoxyribonucleotides. See Table 1 for oligos B, T, P, C and G. *Bam*HI (decamer) and *Bg*III linkers were purchased from New England Biolabs. The pBR322 *Eco*RI sequencing primer was purchased from Promega.

Construction of the Promoter Assay Plasmids

A 50 nucleotide, double stranded, DNA fragment containing a wild type T7 class III promoter or a mutated T7 promoter was synthesized by a combination of chemical and enzymatic methods as described by Schneider and Stormo (20). The primer oligo was annealed to each of the six oligodeoxyribonucleotides, WT, B, T, P, C, and G (Table 1). Large fragment of *E. coli* DNA polymerase I and dNTP's were used to extend the primer and synthesize the DNA strand complementary to the oligonucleotides WT, B, T, P, C and G. The double stranded 50 nucleotide DNA fragments were then cleaved with *Hind*III and ligated between the *Hind*III and *Sma*I restriction sites of pKK232-8 (When the B oligonucleotide was cleaved with *Hind*III and ligated into pKK232-8 it appears that a single A was lost from the 3' end of the oligonucleotide in most of the pCM-B# clones. This A deletion does not affect the sequence of the promoter and does not affect the behavior of the clones.). The ligation mixtures were transformed into competent *E. coli* DH5, and the transformations were grown on LB plates containing 50 μ g/ml ampicillin.

For the WT oligo, ten independent colonies were chosen, and plasmid DNA was isolated from overnight cultures grown at 37°C in 4 ml of LB containing 50 µg/ml ampicillin. Insertion of the T7 ϕ 10 promoter into pKK232-8 was confirmed by *Asel* restriction maps of the plasmids. The cloning places a wild type

Table 1. Oligodeoxyribonucleotides.[illegible]

Oligonucleotides B, P, C, G, and Primer were purchased from National Biosciences. Oligo T was purchased from New England Biolabs, and Oligo WT was a gift from Dr. Keith McKenney, Center for Advanced Research in Biotechnology. In oligos B and T, the lower case italics are used to designate those bases in the oligonucleotide that are a mixture of 96.1% of the indicated base and 1.3% of each of the other 3 bases. In oligo P, the lower case italics are used to designate those bases in the oligonucleotide that are a mixture of 67.0% of the indicated base and 11% of each of the other 3 bases. The mutation in oligo C is an A to C (lower case italics) point mutation at -10, and the mutation in oligo G is an A to G (lower case italics) point mutation at -10. The primer oligo is complementary to the 3' end of oligos WT, B, P, T, C, and G.

T7 $\phi 10$ promoter just upstream of the promoterless chloramphenicol acetyl transferase gene of pKK232-8. This new plasmid is designated pCAT10-1 (Figure 1).

For the B, T, P, C, and G oligodeoxyribonucleotides, 400, 350, 329, 10, and 10 independent colonies, respectively, were chosen from the transformations. Plasmid DNA was isolated from each transformant from overnight cultures grown at 37°C in 4 ml of LB containing 50 μ g/ml ampicillin. Insertion of oligo B into pKK232-8 was confirmed by *Hind*III and *Pst*I restriction maps of the plasmids. Insertion of oligo T was confirmed by *Sca*I restriction maps, and insertion of oligos P, C, and G were confirmed by *Sca*I or *Ase*I restriction maps. This restriction mapping identified 252 B clones, 138 T clones, 121 P clones, 8 C clones, and 8 G clones with apparent insertions of the respective oligodeoxyribonucleotides. These clones are designated pCM-B#, pCM-T#, pCM-P#, pCM-C#, and pCM-G#, respectively (Figure 1), where # is the general designation for the number assigned to the isolated clone. In the remainder of the paper the entire family of clones will be referred to as pCM-X#, where X refers to the B, T, P, C, and G clones, collectively.

Subsequently, these clones were screened for T7 promoter activity (see below).

Construction of pKGP1-1

To screen the promoter assay plasmids for *in vivo* T7 promoter activity it is necessary to express T7 RNA polymerase in the cell strain used for screening. For this purpose we chose to construct a p15A origin plasmid that expresses T7 RNA polymerase from the *tac* promoter.

Plasmid pGP1-5 (2.0 μ g) (23) was linearized by cleavage with *Bgl*II, and the linearized plasmid DNA was incubated at 37°C with 0.3 units of *Bal*31 to delete the lambda P_L promoter from the DNA. Samples (0.4 μ g) were removed from the reaction after 5, 10, 20, and 40 minutes of incubation, and the ends of the *Bal*31-treated DNA were repaired with large fragment of *E. coli*

DNA polymerase I in the presence of excess dNTP's. *Bam*HI linkers were ligated to the DNA samples, the ligation mixtures were recut with *Bam*HI, and the *Bam*HI digested DNA was analyzed by gel electrophoresis on a 1% low melting point agarose gel. The 2.8 kilobase fragments from the 10 and 20 min. samples were cut out of the gel, and the fragments were isolated on an Elutip-d column (Schleicher and Schuell). These 2.8 kilobase *Bam*HI fragments that contain T7 gene 1 were ligated into the *Bam*HI site of pUC19, and competent *E. coli* DH5 α were transformed with the two ligation reactions. The transformations were grown on LB plates containing 50 μ g/ml ampicillin, 20 μ g/ml X-gal, and 0.5 mM IPTG. Ten white colonies from each of the two ligations were picked for further analysis. Plasmid DNA was isolated from 4.0 ml overnight cultures (37°C and 50 μ g/ml ampicillin in LB) of each of the twenty colonies, and the lengths of the DNA fragments inserted into the pUC vector were analyzed by restriction mapping with *Acc*I, *Eco*RI, *Bam*HI, *Bam*HI & *Sca*I, *Eco*RI & *Hind*III, and *Acc*I & *Kpn*I. One clone contained an insert with 337 bases removed from the *Bgl*II end of the *Bgl*II/*Bam*HI fragment of pGP1-5. This clone was designated pGP1-20B, and was chosen for further manipulation.

The 2.5 kilobase *Eco*RI/*Hind*III fragment from pGP1-20B was isolated on a 1% low melting point agarose gel and ligated into the *Eco*RI/*Hind*III site of plasmid pKK223-3. This places T7 gene 1 immediately downstream of the *tac* promoter of plasmid pKK223-3. This plasmid was designated pKK-gp1.

Plasmid pKK-gp1 was cut with *Nae*I, and *Bgl*II linkers were ligated onto the *Nae*I cleavage sites. The ligation mix was recut with *Bgl*II and *Bgl*II, and the 4362 base pair *Bgl*II/*Bgl*II(*Nae*I) fragment from pKK-gp1 was ligated to the 2892 base pair *Bam*HI/*Bgl*II fragment of pACYC177 (26,27). This ligates the *tac* promoter-T7 gene 1 fusion to the p15A origin of pACYC177 and generates a plasmid that also expresses both ampicillin and kanamycin resistance. To inactivate the ampicillin resistance gene, the p15A/gene 1 plasmid was recut with *Bgl*II and *Sca*I, repaired with T4 DNA polymerase and excess dNTP's, and recircularized with T4 DNA ligase. This deletes 360 bases from the middle of the ampicillin resistance gene and produces plasmid pKGP1-1 (Figure 2). In *E. coli* JM101 plasmid pKGP1-1 efficiently produces active T7 RNA polymerase in the presence of IPTG (data not shown).

Figure 1: Diagram of the family of promoter selection plasmids pCM-X#. 'Mutant T7 promoter' designates the location of the insertion of the mutated oligonucleotides as described in the Materials and Methods section. *rrnBT2T1/5S* are the transcription terminators of the *E. coli* 5S rRNA gene. Insertion of the WT oligo (the oligo containing an unmutated T7 $\phi 10$ promoter) into the location designated by 'Mutant T7 promoter' generates plasmid pCAT10-1

Figure 2: Diagram of plasmid pKGP1-1.

Screening for T7 Promoter Activity *In Vivo*

In the presence of T7 RNA polymerase, pCAT10-1 and the pCM-X# plasmids that contain sequences recognizable as T7 promoters will express chloramphenicol acetyl transferase. Consequently, *E. coli* harboring pKGP1-1 and pCAT10-1 or pCM-X# can be screened for chloramphenicol resistance or chloramphenicol sensitivity.

Table 2. Clone Selection

	B	Clone Family T	P
Total # of Clones	252	138	121
# Strong Clones	36	71	32
# Intermediate Clones	119	40	33
# Inactive Clones	97	27	56

B, T, and P refer to the families of clones generated from the oligonucleotides B, T, and P, respectively.

Competent *E. coli* JM101, 200 μ l, is cotransformed with pKGP1-1 (40 ng) and the promoter assay plasmid (1 μ l of a solution of DNA from a miniprep), either pCAT10-1 or a pCM-X# plasmid. SOC, 0.9 ml, is added to the cells, and the culture is grown at 37°C for one hour. Equal volumes of the transformation, 200 μ l, are spread on three different types of LB-Agar plates containing either (1) 50 μ g/ml kanamycin and 50 μ g/ml ampicillin, (2) 50 μ g/ml kanamycin, 50 μ g/ml ampicillin, and 1.0 mM IPTG, or (3) 50 μ g/ml kanamycin, 50 μ g/ml ampicillin, and 30 μ g/ml chloramphenicol. The plates are incubated at 37°C for 16 to 20 hours, and the number of colonies on each plate is counted.

Note: The *E. coli* JM101 must be positively *lacIQ*. The absence of a good *lacIQ* phenotype produces misleading results.

Sequencing the Mutated T7 Promoters

The pCAT10-1 and the pCM-X# plasmids were sequenced using the double stranded dideoxy termination method as described in the Sequenase Version 2.0 (United States Biochemical) protocols manual. The primer used for sequencing was the pBR322 *EcoRI* sequencing primer from Promega.

Table 3. Mutations that Reduce T7 Promoter Activity (Intermediate Substitution Mutants)

pCM-X#	Sequence	Changes	pCM-X#	Sequence	Changes
	-17	+6		-17	+6
WT	GAAATTAATACGACTCACTATAGGGAGA		WT	GAAATTAATACGACTCACTATAGGGAGA	
B64	C.....	1	T271	C.....T...	2
B93	A.....	1	T289	G.....T...	2
B100	C.....	1	T290	...C.....C	2
B121	C.....	1	T301	.G.....T.....	2
B139	C.....	1	T308	.C.....G.....	2
B156	T.....	1	T336	.C.....T.....	2
B186	G.....	1	T415	T.....A.....	2
B233	T.....	1	T444	T.....T..T	2
B241	G.....	1	T467	A.....T.....	2
B248	T.....	1	T482	.G.....A.....	2
B254	T.....	1	T493	...A..C.....	2
B282	A.....	1	P1260	.C.....T.....	2
T220	T.....	1	B33	.G..C.....G.....	3
T221	G.....	1	B36	C.....AG.....	3
T231	T.....	1	B39	T.....T.C.....	3
T272	G.....	1	B78	CT.....C.....	3
T324	C.....	1	B84	ACA.....	3
T418	G.....	1	B201	A.....T.G.....	3
T454	T.....	1	B268	C.....G.....T	3
T503	T.....	1	B271	.T.....T.G.....	3
P1198	G.....	1	B350	AT.....T.....	3
P1263	A.....	1	T280	C.....C..T	3
C2001	C.....	1	T284	G..T..G..	3
G3001	G.....	1	T323	A.....T...A.	3
B59	T...A.....	2	T269	...G..C..C.....	3
B166	C.....T.....	2	B38	C..C...GA.	4
B188	...T..T.....	2	B258	TC.GA.....	4
B226	A.....T.....	2	T236	T..G.....C.C.....	4
B277	..G..A.....	2	T508	..CCAG.....	4
B456	A.....C.....	2	B246	TACT.....A.....	5
T224	A.....G.....	2	T37	.CC.GG..T.....	5
T240	G.....A.....	2	T83	.CC.GG.....G.....	5
T244	T.....A.	2			

WT designates the T7 consensus promoter. The table lists the base changes observed in the mutants. Many of the pCM-B# clones were missing an A next to the *HindIII* site. This is a deletion produced during cloning; however, this deletion does not change the sequence of the promoter and does not affect the behavior of the promoter selection plasmids.

Measurement of Promoter Strength *In Vitro*

As previously described by Burgess and coworkers (15,16), promoter strength *in vitro* can be determined by measuring the production of RNA from a supercoiled plasmid containing a mutant promoter in comparison to the production of RNA from a supercoiled plasmid containing a consensus promoter (pCAT10-1). See 'Results' for the discussion of the assay conditions.

Plasmids pCAT10-1, pKK232-8, pCM-B14, pCM-B57, pCM-B93, pCM-B139, pCM-T448, pCM-T460, pCM-T270, and pCM-T286 were prepared by standard methods (28). The identity of the plasmids were confirmed by restriction mapping (28) and by phenotypic analysis, *in vivo* (see the section Screening for T7 Promoter Activity). Electrophoretic analysis of the plasmids showed only supercoiled DNA (data not shown). No linear or open circular forms could be detected. The plasmids were used in their natural state at natural superhelical densities.

Transcription assays were carried out in 60 μ l reactions containing 50 mM Tris-HCl, pH 7.8, 60 mM NaCl, 2.5 mM spermidine, 10 mM MgCl₂, 1 mM dithiothreitol, 400 μ M of each of the four rNTPs, 0.64 μ Ci [2,8-³H]-ATP, 30 μ g/ml bovine serum albumin, 8 nM promoter, and 80 nM T7 RNA polymerase. The reactions were equilibrated at 37°C, and RNA synthesis was initiated by addition of T7 RNA polymerase. At 6 min. and 12 min., 25 μ l samples were removed from the reactions and were spotted on DE81 filters. The filters were washed three times with 0.3 M ammonium formate, pH 8.0, and once with ethanol (29). The RNA retained on the filters was then quantitated by liquid scintillation. The specific activity of the [2,8-³H]-ATP was 41.6 cpm/pmol.

The relative strength of a promoter on a pCM-X # clone was determined by dividing the amount of RNA synthesized in a reaction containing the pCM-X # clone by the amount of RNA synthesized in a reaction containing pCAT10-1 consensus clone. This procedure defines the activity of pCAT10-1 to be 1.00.

RESULTS

Assaying T7 Promoter Activity *In Vivo*

The T7 ϕ 10 promoter and oligonucleotides carrying mutagenized T7 promoters were cloned into the promoter selection vector pKK232-8. The promoter selection plasmid containing the T7

ϕ 10 promoter was designated pCAT10-1, and the promoter selection plasmids containing mutagenized promoters were designated pCM-X #. The plasmid pCAT10-1 and the pCM-X # plasmids that contain sequences that can be recognized as T7 RNA polymerase promoters will express chloramphenicol acetyl transferase in the presence of T7 RNA polymerase. Consequently, *E. coli* harboring the T7 RNA polymerase expression plasmid pKGP1-1 and a promoter selection plasmid, either pCAT10-1 or a pCM-X # plasmid, can be screened for chloramphenicol resistance or chloramphenicol sensitivity.

When *E. coli* JM101 that contains pKGP1-1 and a promoter selection plasmid are spread on LB-agar plates containing: (1) 50 μ g/ml kanamycin (kan) and 50 μ g/ml ampicillin (amp), (2) 50 μ g/ml kanamycin, 50 μ g/ml ampicillin, and 1.0 mM IPTG, and (3) 50 μ g/ml kanamycin, 50 μ g/ml ampicillin, and 30 μ g/ml chloramphenicol (cam), three types of results are observed. If the promoter selection plasmid carries a strong T7 promoter (e.g. pCAT10-1) there are approximately 200 colonies on the kan/amp and kan/amp/cam plates and no colonies on the kan/amp/IPTG plate (A strong T7 promoter in the presence of large amounts of T7 RNA polymerase is lethal in *E. coli* (20).); these clones are designated as 'strong' clones. If the promoter selection plasmid carries a sequence that cannot be recognized or utilized by T7 RNA polymerase there are approximately 200 colonies on the kan/amp and kan/amp/IPTG plates and no colonies on the kan/amp/cam plate; these clones are designated as 'inactive' clones. Finally, if the promoter selection plasmid carries a sequence that can be weakly recognized or utilized by T7 RNA polymerase there are approximately 200 colonies on the kan/amp and kan/amp/IPTG plates, and between 25 and 200 colonies on the kan/amp/cam plate; these clones are designated as intermediate clones. Representative numbers of strong, inactive, and intermediate clones identified are tabulated in Table 2. As suggested by Table 2, the families of clones generated from the five different mutagenic oligonucleotides are designated with the same letter associated with the oligonucleotide used to construct the clones (Table 1 and Experimental Procedures).

The sequences of the 'promoter regions' of 76 intermediate B-clones and 40 intermediate T-clones were determined by double stranded dideoxy sequencing. Forty-seven of the intermediate B-clones were substitution mutants containing one to five base changes, while the remaining 29 intermediate B-clones were

Table 4. *In Vitro* Activities of Representative Mutant Promoters

Plasmid	Type	Mutation	Activity $\pm \sigma$	Net Activity
pCAT10-1	WT	Consensus	1.00	1.00
pKK232-8	Vector	No promoter	0.005 $\pm$ 0.0040	0.000 $\pm$ 0.008
pCM-B14	Str	+3G to A	0.74 $\pm$ 0.08	0.74 $\pm$ 0.08
pCM-B57	Str	-19A to T	0.72 $\pm$ 0.03	0.72 $\pm$ 0.03
pCM-B93	Int	-5C to A	0.009 $\pm$ 0.005	0.004 $\pm$ 0.009
pCM-B139	Int	-6A to C	0.027 $\pm$ 0.006	0.022 $\pm$ 0.01
pCM-T448	Int	-6A to T	0.13 $\pm$ 0.02	0.13 $\pm$ 0.02
pCM-T460	Int	-6A to C	0.028 $\pm$ 0.008	0.023 $\pm$ 0.01
pCM-T270	Inact	-9C to A	0.005 $\pm$ 0.004	0.000 $\pm$ 0.008
pCM-T286	Inact	-8T to A	0.007 $\pm$ 0.005	0.002 $\pm$ 0.009

The *in vitro* activity of the consensus T7 promoter in plasmid pCAT10-1 was defined as 1.00. WT, wild type; Str, strong; Int, intermediate; Inact, inactive. Mutations are listed as a change at the specified position from the consensus base to a mutant base. The values listed under activity are the fraction of activity observed in an *in vitro* transcription reaction in comparison to the activity observed with pCAT10-1. These values are the average of at least three different trials. The values listed under net activity are obtained by subtracting the background activity of pKK232-8 from the activity values. The errors associated with the net activity values were determined by standard methods for propagation of errors.

deletion or insertion mutants. Thirty-nine of the intermediate T-clones were substitution mutants containing one to five base changes, and the remaining intermediate T-clone was a deletion mutant. Deletion mutants are difficult to analyze with respect to the relationship of promoter sequence to promoter activity; consequently, only the substitution mutants were considered in this study. The unique substitution mutants are shown in Table 3.

High proportions of both the intermediate and inactive B-clones (see below) were deletion mutants. These deletions may have been caused by an oligonucleotide of poor quality or by recombination of the semi-repetitive sequences at the 5' end of oligonucleotide B. The sequences of oligonucleotides T, P, C, and G differ from oligonucleotide B at the extreme 5' end of the oligonucleotides. These differences are upstream of the T7 promoter sequences and reduce the repetitiveness of the 5' end of the oligonucleotide. Clones constructed from oligonucleotides T, P, C, and G showed very few deletion mutants. Furthermore, a control experiment showed that the *in vitro* strengths of identical mutants generated from the B oligo (pCM-B139) and the T oligo (pCM-T460) are indistinguishable (Table 4). This suggests that results with mutants generated from the B oligo are comparable to results with mutants generated from the other four oligonucleotides.

A compilation of the base changes observed in all of the B, T, P, C, and G intermediate substitution mutants (Figure 3) shows that many of the bases within a T7 promoter may be changed without inactivating T7 promoter activity. Furthermore, addition of the nonconsensus bases found in the T7 ϕ OL promoter and the 'weaker' T7 class II promoters (2) to the compilation of the substitution mutants shows that active T7 promoters can differ

from the T7 consensus sequence at almost any position within the promoter. Does this mean that there are no essential promoter contacts necessary for transcriptional activity?

The inactive promoter mutants were sequenced to determine whether single base substitutions could inactivate a T7 promoter. Seventy inactive B mutants and twenty-seven inactive T mutants were sequenced. Fifty-nine of the inactive B mutants were deletion or insertion mutants, while four of the inactive T mutants were deletion mutants. The remainder of the inactive B and T mutants were substitution mutants. None of the inactive B mutants and six of the inactive T mutants contained a single base substitution. The base changes observed in these independent clones were a C to G substitution at -7, a T to A substitution at -8, a C to A substitution at -9, and a G to T substitution at -11. The position and identities of the single base changes in these inactive promoter mutants generally correspond to the base changes that were not seen in the compilation of the intermediate mutant promoters and the nonconsensus T7 promoters; consequently, oligonucleotides P, C, and G were designed to generate the remaining single base substitutions between -6 and -11.

The single base substitution mutants obtained from oligonucleotides P, C, and G, confirmed that positions -7, -8, -9, and -11 are important for T7 promoter activity, but also showed that only seven of the twelve possible single base substitutions at these positions inactivates a T7 promoter (Table 5). The single base substitutions that inactivate a T7 promoter are a C to A or G substitution at -7, a T to A substitution at -8, a C to A, T, or G substitution at -9, and a G to T substitution at -11 (Figure 4).

Assaying T7 Promoter Activity *In Vitro*

Burgess and coworkers have previously reported that the *in vivo* activities of T7 promoters on plasmids most closely correlated with their *in vitro* activities when the promoters were assayed, *in vitro*, on naturally supercoiled templates in the presence of NaCl and spermidine (15,16). To attempt to define what level of promoter activity produces the strong, intermediate, and inactive phenotypes discussed in the previous section and to demonstrate that possible copy number variations of the pKGP1-1 and pCM-X# plasmids do not invalidate the screening results, representative strong, intermediate, and inactive clones were assayed for T7 promoter activity using the methods described by Burgess and coworkers (15,16). Table 4 shows that the relative

A. All Intermediate Mutations

-17 +6
WT GAAAT TAATACGACTCACTATAGGGAGA

t t c t a a t t a t t a t c t t t a t a c t t t g t t
a c g c g g c c c c a c c g c a c c g g c c a c
g g g g g g g g g g g a

B. Mutations Occuring in ϕ OL and Class II T7 Promoters

-17 +6

WT GAAAT TAATACGACTCACTATAGGGAGA

tttta ct tga g ataaagat
ccgcc c cc
aq gg

C. Intermediate plus ϕ OL and Class II Mutations

-17+6

WT GAAAT TAATACGACTCACTATAGGGAGA

t t t t a a t t a t a a t c t t a a t g t t t t g t t
c c c c c c c c c t c g g c t c c a c c a c c c
a q q q q q q q q q c q q q q q a a a

Figure 3: Mutations found in viable T7 promoters. A) A compilation of all the mutations that are seen in the intermediate clones listed in Table 2. B) A compilation of all the nonconsensus bases observed in ϕ OL and the class II T7 promoters. C) All mutations observed in either panel A or panel B.

Table 5. Single Mutations that Inactivate T7 Promoter Activity (Inactive Substitution Mutants)

pCM-X#	Sequence
	-17 +6
WT	GAAATTAATACGACTCACTATAGGGAGA
T270	A.....
T286	A.....
T297	T.....
P1031	A.....
P1087	T.....
P1160	G.....
P1208	G.....

WT designates the T7 consensus promoter. The table lists the base changes observed in the mutants. Generally, there was more than one independently isolated example of each inactive clone. Only the first clone isolated is specifically listed in the table.

activity of the mutant T7 promoters on the strong clones, pCM-B14 and pCM-B57, are 0.72 ± 0.03 and 0.74 ± 0.08 , respectively, while the relative activities of the mutant T7 promoters on the inactive clones, pCM-T270 and pCM-T286, are very nearly zero. The relative activities of the mutant T7 promoters on the intermediate clones, pCM-B93, pCM-B139, and pCM-T448, fall between the activities of the strong and inactive promoters and range from nearly inactive (0.004 ± 0.009 for pCM-B93) to moderately active (0.13 ± 0.02 for pCM-T448).

DISCUSSION

Although the relationship of T7 promoter sequence to promoter function has been studied extensively, the identity of the bases within the T7 promoter that are essential for promoter function *in vivo* are poorly defined. To determine what bases within the T7 consensus promoter are essential for promoter function we have constructed a library of mutant promoters and correlated the *in vivo* activity of the mutant promoters to their sequence.

Mutant T7 promoters were created by mutagenizing the T7 $\phi 10$ promoter between positions -22 and $+6$ during the synthesis of oligonucleotides containing the $\phi 10$ promoter (Table 1). The mutagenized oligonucleotide was then ligated to the chloramphenicol acetyl transferase gene of the promoter selection vector pKK232-8. This creates a plasmid (pCM-X#) that can potentially express chloramphenicol acetyl transferase in the presence of T7 RNA polymerase.

To provide T7 RNA polymerase *in vivo*, a second compatible plasmid (pKGP1-1) was constructed; this pACYC177 based plasmid carries T7 gene 1 ligated to the *tac* promoter. Although others have observed that plasmids carrying a T7 promoter are unstable in the presence of T7 RNA polymerase, plasmids pCM-X# and pKGP1-1 are sufficiently stable in *E. coli* JM101 (see the note in Experimental Procedures) to screen the bacteria that carry the plasmids for chloramphenicol resistance or chloramphenicol sensitivity.

E. coli that contained plasmids pCM-X# and pKGP1-1 were screened by spreading the bacteria on plates containing kanamycin (kan) and ampicillin (amp); kanamycin, ampicillin, and IPTG; and kanamycin, ampicillin, and chloramphenicol (cam). It has been previously noted that a plasmid carrying a strong T7

promoter is lethal to bacteria that express large amounts of T7 RNA polymerase (20). We also found this to be true. *E. coli* containing pKGP1-1 and pCAT10-1 (the promoter selection plasmid containing the T7 $\phi 10$ promoter) would not grow on the plates containing kan, amp, and IPTG; however, if expression of T7 RNA polymerase was not induced, pKGP1-1 and pCAT10-1 were compatible. *E. coli* containing pKGP1-1 and pCAT10-1 would grow on plates containing kan and amp but lacking IPTG.

In the uninduced state the *tac* promoter is not silent. Basal transcription from the *tac* promoter allows expression of T7 RNA polymerase in the absence of IPTG. This small amount of T7 RNA polymerase allows chloramphenicol acetyl transferase to be expressed if the selection plasmid contains a recognizable T7 promoter. It was observed that *E. coli* JM101 containing pKGP1-1 and pCAT10-1 was resistant to chloramphenicol, and grew equally well on plates containing either kan, amp, and cam or kan and amp. Promoter selection plasmids (pCM-X#'s) that in combination with plasmid pKGP1-1 endowed *E. coli* JM101 with growth characteristics identical to pCAT10-1 were considered to contain strong or wild-type T7 promoters and were categorized as 'strong' selection plasmids.

If a promoter selection plasmid does not contain a viable T7 promoter, the *E. coli* containing the promoter selection plasmid and pKGP1-1 grows equally well on plates containing either kan and amp or kan, amp, and IPTG; induction of T7 RNA polymerase is not lethal. However, the lack of a viable T7 promoter to transcribe the chloramphenicol acetyl transferase gene results in sensitivity to chloramphenicol and no bacterial growth on plates containing kan, amp, and cam. Promoter selection plasmids (pCM-X#'s) that exhibited these growth characteristics were considered to contain inactive T7 promoters and were categorized as 'inactive' selection plasmids.

Lastly, a third phenotype was observed in the screening assay. Many of the promoter selection plasmids allowed growth on all three of the antibiotic plates. The promoter selection plasmids conferred chloramphenicol resistance to the host *E. coli* in the presence of basal amounts of T7 RNA polymerase, but were not lethal to the host in the presence of high levels of T7 RNA polymerase (IPTG induction of the *tac* promoter). In our opinion, this behavior is indicative of a promoter that can be recognized by T7 RNA polymerase, but is not as strong a promoter as $\phi 10$. These promoter selection plasmids were categorized as 'intermediate' selection plasmids.

To confirm the phenotypic characterizations, representative strong, intermediate, and inactive clones were assayed for T7 promoter activity *in vitro*, and these *in vitro* results support the *in vivo* observations.

Sequencing the promoter regions of the intermediate clones reveals mutations that reduce promoter strength but are not absolutely essential for T7 promoter activity. Compilation of these mutations (Table 3) and the nonconsensus bases found in the ϕOL and the class II promoters (Figure 3) shows that few bases are absolutely essential for T7 promoter activity; however, many of the possible mutations that do not appear in this compilation congregate between positions -7 and -11 . While Schneider and Stormo previously identified mutations in the T7 promoter that do not reduce promoter activity (20), their results are similar to the results obtained here with the intermediate promoters. All of this data appears to support Schneider and Stormo's hypothesis that positions -7 to -9 are important for T7 promoter activity (20).

Intermediate Mutations

Inactive Mutations

Figure 4: The phenotypes of the point mutations analyzed. Each lower case letter found either above or below the capitalized wild type T7 promoter sequence represents the single base change found in a point mutant. The mutations shown above the wild type sequence were classified as intermediate clones and the mutations shown below the wild type sequence were classified as inactive clones.

The importance of specific bases within the T7 promoter can be directly verified by sequencing the inactive mutants. The sequences of the inactive mutants that carry a single base substitution are listed in Table 5 and reveal that a number of the bases between -7 and -11 are essential for T7 promoter activity. Characterization of 20 of the 21 possible point mutants from -5 to -11 show that only seven mutations inactivate the T7 promoter. Mutations that change C to A or G at -7; T to A at -8; C to A, T, or G at -9, or G to T at -11 all inactivate the T7 promoter, while the 13 remaining single base substitutions between -5 and -11 generate intermediate promoters (Figure 4). Since much of the -12 to -17 region of the T7 promoter can be deleted or changed without abolishing promoter activity (17,18,19), and since the remaining point mutations between -5 and -2 do not seriously reduce T7 transcription *in vivo* (16), it is likely that the seven point mutations that produce inactive T7 promoters are the only point mutations that inactivate the promoter.

In previous work, McAllister and coworkers (12) found that methylation of the G residues at positions -7, -9, and -11 of the T7 promoter interferes with binding of the promoter by T7 RNA polymerase. This suggests that T7 RNA polymerase contacts the major groove of the T7 promoter between positions -7 and -11. When a G is replaced with a T, the 5-methyl group of the T occupies the same position as the methyl group of a N7-methylated G; consequently, inactivation of the T7 promoter by a -11G to T substitution, a -9C to A substitution, and a -7C to A substitution may be due to the same steric effects that prevent binding of T7 RNA polymerase to a promoter carrying N7-methylated G residues between positions -7 and -11. This seems to further emphasize that T7 RNA polymerase and its promoter must be very close to one another at the N7 positions of the G residues at -7, -9, and -11 of the double helical promoter. Furthermore, specificity sites analysis (as described by Seeman *et al.* (30) and McClarin *et al.* (31)) of major groove functionality of the wild type, intermediate, and inactive bases at positions -7, -8, -9, and -11 of the double helical T7 promoter seems to suggest that a hydrogen bond acceptor at the N7 position of the G at -7, close contact with the N7 position of the A at -8, a hydrogen bond donor at the 4-amino group of the C at -9, and a hydrogen bond acceptor at the 6-keto group at the G at -9 are important for T7 promoter activity.

In conclusion, it appears that mutations at many of the bases within the conserved consensus sequence of the T7 RNA polymerase promoter can reduce the efficiency of the promoter, but that only bases at positions -7, -8, -9, and -11 are essential for T7 promoter function.

ACKNOWLEDGEMENTS

This work was supported by a FIRST grant (AI-24905) from the National Institute of Allergies and Infectious Diseases. We thank Dr. Martha Briggs for helpful comments and discussion.

REFERENCES

- Dunn, J. J. and Studier, F. W. (1983) *J. Mol. Biol.*, **166**, 477-535.
- Studier, F. W. and Dunn, J. J. (1982) *Cold Spring Harbor Symp. Quant. Biol.*, **47**, 999-1007.
- Moffatt, B. A., Dunn, J. J. and Studier, F. W. (1984) *J. Mol. Biol.*, **173**, 265-269.
- Oakley, J. L., Strothkamp, R. E., Sarris, A. H. and Coleman, J. E. (1979) *Biochemistry*, **18**, 528-537.
- Klement, J. F., Moorefield, M. B., Jorgensen, E., Brown, J. E., Risman, S. and McAllister, W. T. (1990) *J. Mol. Biol.*, **215**, 21-29.
- Joho, K. E., Gross, L. B., McGraw, N. J., Raskin, C. and McAllister, W. T. (1990) *J. Mol. Biol.*, **215**, 31-39.
- Oakley, J. L., Pascale, J. A. and Coleman, J. E. (1975) *Biochemistry*, **14**, 4684-4691.
- Ikeda, R. A. and Richardson, C. C. (1986) *Proc. Natl. Acad. Sci. U. S. A.*, **83**, 3614-3618.
- Gunderson, S. I., Chapman, K. A. and Burgess, R. R. (1987) *Biochemistry*, **26**, 1539-1546.
- Basu, S. and Maitra, U. (1986) *J. Mol. Biol.*, **190**, 425-437.
- Muller, D. K., Martin, C. T. and Coleman, J. E. (1989) *Biochemistry*, **28**, 3306-3313.
- Jorgensen, E. D., Durbin, R. K., Risman, S. S. and McAllister, W. T. (1991) *J. Biol. Chem.*, **266**, 645-651.
- Stahl, S. J. and Chamberlin, M. J. (1978) *J. Biol. Chem.*, **253**, 4951-4959.
- Stahl, S. J. and Chamberlin, M. J. (1976) in Chamberlin, M. and Losick, R. (eds.), *RNA Polymerase*. Cold Spring Harbor Press, Cold Spring Harbor, New York, pp. 429-440.
- Chapman, K. A. and Burgess, R. R. (1987) *Nucleic Acids Res.*, **15**, 5413-5432.
- Chapman, K. A., Gunderson, S. I., Anello, M., Wells, R. D. and Burgess, R. R. (1988) *Nucleic Acids Res.*, **16**, 4511-4524.
- Chapman, K. A. and Wells, R. D. (1982) *Nucl. Acids Res.*, **10**, 6331-6340.
- Osterman, H. L. and Coleman, J. E. (1981) *Biochemistry*, **20**, 4884-4892.
- Martin, C. T. and Coleman, J. E. (1987) *Biochemistry*, **26**, 2690-2696.
- Schneider, T. D. and Stormo, G. D. (1989) *Nucl. Acids Res.*, **17**, 659-674.
- Tabor, S. and Richardson, C. C. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 1074-1078.
- Ikeda, R. A. and Richardson, C. C. (1987) *J. Biol. Chem.*, **262**, 3790-3799.
- Chamberlin, M., McGrath, J. and Waskell, L. (1970) *Nature*, **228**, 227-231.
- Brosius, J. and Lupski, J. R. (1987) *Methods in Enzymology*, **153**, 54-68.
- Brosius, J. and Holy, A. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 6929-6933.
- Chang, A. C. Y. and Cohen, S. N. (1978) *J. Bacteriol.*, **134**, 1141-1156.
- Rose, R. E. (1988) *Nucl. Acids Res.*, **16**, 356.
- Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) in *Molecular Cloning: A Laboratory Manual*, 2 ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
- Brutlag, D. and Kornberg, A. (1972) *J. Biol. Chem.*, **247**, 241-248.
- Seeman, N. C., Rosenberg, J. M. and Rich, A. (1976) *Proc. Natl. Acad. Sci. USA*, **73**, 804-808.
- McClarin, J. A., Frederick, C. A., Wang, B.-C., Greene, P., Boyer, H. W., Grable, J. and Rosenberg, J. M. (1986) *Science*, **234**, 1526-1541.