Volume 15 Number 13 1987 Nucleic Acids Research

# Construction of bacteriophage T7 late promoters with point mutations and characterization by *in vitro* transcription properties

Kenneth A.Chapman* and Richard R.Burgess

McArdle Laboratory for Cancer Research, University of Wisconsin, Madison, WI 53706, USA

Received January 22, 1987; Revised and Accepted June 4, 1987

## ABSTRACT

This paper describes the construction of 18 cloned bacteriophage T7 late promoters with single point mutations. In vitro transcription experiments were used to characterize the properties of these promoters. Since the mutated promoters are cloned into identical backgrounds, differences seen in the transcription assays are directly attributable to the point mutations.

All of the mutated promoters are less active than wildtype, but they can be divided into two types. Type A mutations map from -4 to +1 and reduce promoter activity when the template is linearized or when 60mM NaCl is added to the reaction buffer. Type B mutations map from -9 to -7 and reduce promoter activity under all conditions tested. At several sites all three possible point mutations are available. At these sites we observed hierarchies of base pair preference, as determined by promoter activity, that may indicate that T7 RNA polymerase interacts with groups in the major groove.

## INTRODUCTION

The viral encoded RNA polymerases of T7-like bacteriophages show a high degree of template specificity because they initiate transcription only at the late promoters, which are found on the phage DNA but are absent on the host DNA (1). The 17 late promoters of T7 are considerably different from bacterial promoters, in that they consist of a single uninterrupted sequence that is 23 base pairs long (2, 3). There are two classes of late promoters. The class II late promoters are responsible for the transcription of the middle genes, and the class III late promoters are responsible for the expression of the late genes. The late promoter sequence is highly conserved. In class III promoters the 23 basepair region from -17 to +6 is absolutely conserved (4). The class II promoters tend to show some deviations, averaging 4 non-consensus base pairs each.

This paper describes the construction and characterization by in vitro transcription assays of 18 cloned T7 late promoters that each contain a single point mutation in the 11 bp region from -9 to +2 region. This 11 bp region is highly conserved in the naturally occurring T7 late promoters, and also in the

related phages T3 and SP6. In the region from -10 to -3, there is only one example of a T7 promoter with a deviation from the consensus sequence. This is a change in the class II promoter at 16.0%. This promoter has a G, instead of a C, at position -5. It was shown by Panayatatos and Wells that this promoter was less active than nearby class II promoters when tested in an *in vitro* transcription system using a linearized template (5).

The point mutations were made using deoxyoligonucleotide mediated site-directed mutagenesis methods (6). A piggyback mutagenesis procedure, which is described in the materials and methods section, made possible a biological screen for mutations in the late promoter. This screening procedure was independent of promoter function and thus facilitated the isolation of mutations without regard for their effect on promoter activity.

In the class II late promoters there are some examples of deviations from the consensus sequence which occur in the 11 bp region. As mentioned above, the 16.0% promoter has a C to G change at -5 (5). Other naturally occurring deviations in the 11 bp region of interest are found from -2 to +2. The substitution of an A for the T at -2 is fairly common, it occurs in 4 class II promoters. However, all of these promoters also contain additional deviations from the consensus. While these changes are interesting, it is better to study such a deviation from the consensus sequence in a cloned promoter where the change being studied is the only difference from the consensus. In this type of system the effects on the promoter activity can be definitely correlated to the specific change that has been made.

## MATERIALS AND METHODS

### Enzymes, Strains, and Media

The restriction enzymes Eco RI, Bam HI, Hinf I and Hind III and also T4 DNA ligase were purchased from Promega Biotec. Polynucleotide kinase and Sph I were purchased from New England Biolabs. Klenow fragment of *E. coli* DNA polymerase was purchased from Boehringer Mannheim. Enzyme reactions were done using the conditions recommended by the manufacturers.

T7 RNA polymerase was purified from an overproducing *E. coli* strain containing the plasmid pAR1219 (7). Cells were grown, induced, lysed, and a cleared lysate prepared as described by Studier (7). The lysate was then diluted with an equal volume of buffer with 50 mM $NH_4Cl$ and loaded onto a DE52 column. The polymerase was eluted with a 50 to 500 mM $NH_4Cl$ gradient. Fractions containing the peak of activity were pooled, diluted with 2 volumes of buffer without salt and the DE52 chromatography was repeated. The second

Table 1 SYNTHETIC OLIGONUCLEOTIDES USED FOR IN VITRO MUTAGENESIS

| Name | Sequence | Site | Change * |
|---|---|---|---|
| KC-1D | 5'AATTCGGTTAATACGACTCACTA*AGGGAGATAGGGG | -2 | T to V |
| KC-1C | 3' GCCAATTATGCTGAGTGAT*TCCCTCTATCCCCCTAG | -2 | A to B |
| KC-2 | 3' GCTGAGT*ATATCCCT | -5 | G to H |
| KC-3 | 3' GCTGAG*GATATCCCT | -6 | T to V |
| KC-4 | 3' GAGTGA*ATCCCTC | -3 | T to V |
| KC-5 | 3' GAGTGATA*CCCTC | -1 | T to V |
| KC-6 | 3' GAGTGATAT*CCTCT | +1 | C to D |
| KC-7 | 3' GAGTGATATC*CTCT | +2 | C to D |
| KC-8 | 3' TGAGTG*TATCCCTC | -4 | A to B |
| KC-9 | 3' GCTGA*TGATATCCCTC | -7 | G to H |
| KC-10 | 3' TGCTG*GTGATATCCCTC | -8 | A to B |
| KC-11 | 3' ATGCT*AGTGATATCCCTC | -9 | G to H |

* B = not A, D = not C, H = not G, V = not T

DE52 column helped remove lysozyme, a major contaminant of the lysate. Peak fractions were pooled and diluted with 1 volume of buffer P plus 0.2M NaCl (8) then further purified by chromatography on a Whatman P11 column as described previously (8). Peak fractions were concentrated by dialysis against storage buffer containing 50% glycerol (8). The final yield was 134,000 units of enzyme at a specific activity of 240,000 units/mg from 1.3g of wet cells.

The principle strains of E. coli used were: 7118, for the growth of M13 (9), and C600 galK$^-$,recA$^-$, which is used in conjunction with the plasmid pKO1 (10). The vectors for the mutagenesis and recloning were respectively the single stranded DNA phage M13 mp8 (11), and the plasmid pKO1 (12). The standard plate media was LB (13), with ampicillin added to a concentration of 0.1 mg/ml. Colored plaque assays used plates containing IPTG and X-gal (14).

Synthetic Oligonucleotides

The synthetic oligonucleotides that were used for in vitro mutagenesis procedures are shown in table 1. All oligonucleotides were synthesized using the solid support phosphoramidite procedure, most using an Applied Biosystems automated DNA synthesizer in the UW Biotechnology Center (15).

For the in vitro mutagenesis procedure to work the oligonucleotide must have a 5' phosphate. Kinase reactions were done using about 20 umol. of each oligonucleotide. Most of the ATP in the reactions was non-radioactive, however, a small amount of gamma $^{32}$P-ATP was added to monitor the efficiency of the phosphorylation. The 5' phosphorylated oligonucleotides were purified by gel electrophoresis. In each case the top band was excised from the gel and the oligonucleotide was recovered by the crush and soak method (16).

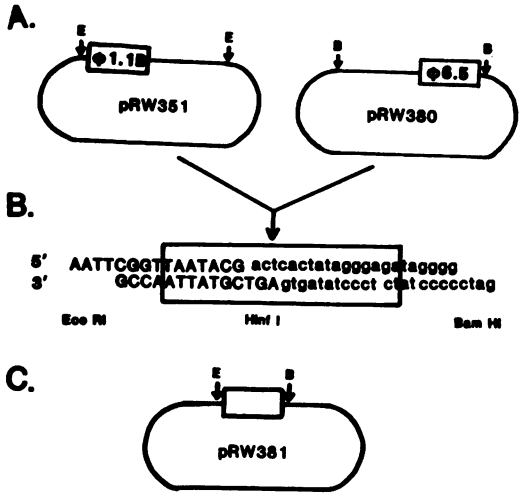


Figure 1
A. The plasmids pRW351 and pRW380 were described previously (8). The class II late promoter in pRW351 has an Eco RI site at -25. pRW380 has a class III promoter with a Bam HI site at +12.

B. The Hinf I site, located at -10, was used to construct a 37 bp fragment containing the left part of the 14.8% class II promoter and the right part of the 46.0% class III late promoter.

C. The 37 bp fragment was cloned into pBR322 to make pRW381.

The oligonucleotides were concentrated and desalted using SepPak C18 disposable reverse phase columns. New columns were first washed with 30% acetonitrile then with 10mM Tris-HCl pH 7.9, 1mM EDTA. The oligonucleotide in gel elution buffer was loaded and the column washed with water. Oligonucleotides were eluted with 30% acetonitrile and dried in a Savant Speed Vac. Sequences were confirmed by using a modified Maxam-Gilbert procedure (16, 17).

Mutagenesis Methodology

The techniques and procedures of oligonucleotide mediated site directed mutagenesis have been described by Smith and Zoller (6). To obtain a single stranded template containing the site to be mutagenized, a wildtype T7 late promoter was cloned into a derivative of the single stranded DNA phage M13. The promoter used as the wildtype is in fact not a naturally occurring T7 late promoter, but a hybrid consisting of the left half of a class II promoter and the right half of a class III promoter (see figure 1). A similar promoter constructed by Jollife, Carter and McAllister was shown to act like a class II late promoter in *in vitro* transcription experiments (18). The hybrid promoter was chosen as the wildtype because: 1) it is an active promoter for T7 RNA polymerase utilization *in vitro*, 2) it has the consensus sequence of T7 late promoters from -19 to +6, 3) it could be easily isolated on a 37 bp Eco RI - Bam HI fragment, 4) the 37 bp fragment can be cloned in a specific orientation in M13 mp8, and 5) this construction permits a biological screen to be used for finding mutations in the promoter sequence.

Figure 2 shows the procedure used to make the collection of mutated

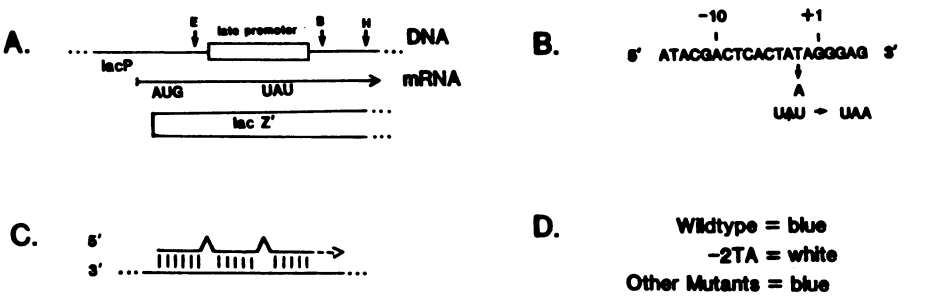


Figure 2
A. Map of the N-terminal portion of the lac Z' gene of M13 mp8 with the 37 bp Eco RI - Bam HI fragment cloned into the polylinker site. Cloning the promoter introduced neither a frameshift mutation nor a premature termination codon in the lac Z' gene. The location of the translation initiation codon is shown as AUG. A UAU codon is located in the late promoter sequence. E, B and H stand for the Eco RI, Bam HI, and Hind III restriction sites.
B. Changing the T at -2 to an A changes the UAU to a termination codon.
C. The piggyback mutagenesis scheme uses a synthetic deoxyoligonucleotide with mismatches at 2 sites. The template for the piggyback mutagenesis reactions was the + strand of M13-KCT7PM -2TA. One of the mutations is the reversion of -2TA back to a T, the other is silent.
D. The selection scheme for mutations in the late promoter.

late promoters. When the 37bp fragment was cloned into M13 mp8 it became a part of the N-terminal portion of the still active lac Z' gene. Expression of the lac Z' gene is easily monitored on a plate containing IPTG and X-gal. Mutations in the late promoter sequence could be screened for if they also affected the expression of lac Z' gene. A piggyback mutagenesis procedure was used to obtain most of the promoters with point mutations. The oligonucleotide was designed to produce two point mutations, one that could be detected easily, and a second that was silent but closely linked with the first. Since the selectable mutation involved the reversion of a previously obtained promoter mutation, the silent point mutation ended up being the only change from the wildtype.

The clone of the wildtype promoter in pKO1 is designated pKCT7P. The system used to identify point mutants is as follows: -2TA represents a point mutation at position -2 relative to the transcription initiation site. The wildtype promoter has a T in the antisense strand at this position, -2TA is the replacement of this T by an A. Similarily -4TG would indicate a mutant with the T at -4 being replaced by a G.

Successful application of this mutagenesis scheme requires obtaining the -2TA mutation. This mutant produces a white plaque. However, spontaneous

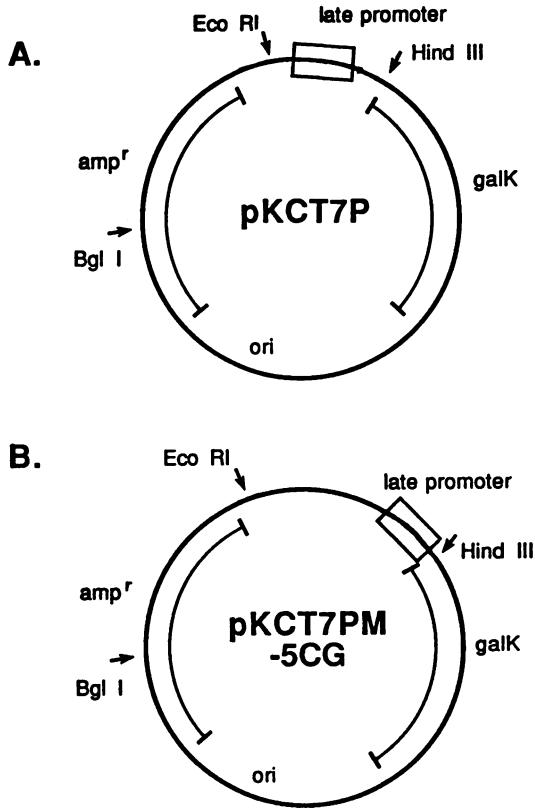


Figure 3
A. The wildtype or mutated promoters were cloned into pKO1 on 57 bp Eco RI - Hind III fragments, placing the galK gene under the transcriptional control of the late promoter.

B. A 438 bp Taq I fragment containing the 16.0% late promoter was isolated from the plasmid pRW307 (5), and cloned into the Cla I site of pBR322. From this clone a 469 bp Eco RI - Hind III fragment containing the late promoter was isolated, and cloned into the plasmid pKO1.

mutations of the lac promoter or lac Z' gene can also produce a white plaque. Oligonucleotides 1C and 1D were used to make the -2TA mutation. Together they produce the 37 bp Eco RI - Bam HI fragment with a mixture of nonwildtype base-pairs at -2. This 37bp fragment was cloned into M13 mp8, yielding white and blue plaques at a ratio of 4:5. Two white plaques were sequenced, both had the -2TA point mutation. Of 13 blue plaques sequenced, all were -2TC.

Single stranded DNA from M13-KCT7PM -2TA was used as the template for the piggyback mutagenesis procedure. Picking out blue plaques from a background of white plaques gives fewer artifacts because a functional lac Z' gene is required to get a blue plaque.

Recloning the Mutated Promoters into the Plasmid pKO1

The wildtype and mutated promoters were excised from the M13 vector on a 57 bp Eco RI - Hind III fragment and cloned into pKO1. pKO1 carries a gene for resistance to ampicillin and has an unexpressed, promoterless,

Table 2 CONDITIONS FOR TYPICAL IN VITRO TRANSCRIPTION ASSAYS

| Components | Methodology |
|---|---|
| Water<br>Buffer<br>T7 RNA polymerase<br>Template DNA<br>ATP,GTP,CTP,UTP ($^3$H-UTP) | mix components<br>incubate<br>spot 3MM filter<br>wash filter with TCA<br>count |
| **Typical Reactions** | **Data analysis** |
| volume = 0.100 ml<br><br>take 2 aliquots of 0.040 ml each to spot filters<br><br>two timepoints at 6 and 12 min. | reactions done in triplicate<br>two filters spotted per rxn.<br>to determine rates,<br>assume linearity from 0<br>extrapolate to 1 hr.<br>determine mean and std. dev.<br>of 6 data points |
| **Constants** | **Variables** |
| Buffer: 50 mM Tris-HCl pH 7.6<br>0.1 mM dithiothreitol<br>10 mM $MgCl_2$<br>2.5 mM spermidine<br><br>[NTPs]: 0.3 mM each<br>polymerase: 50 U/ml (2 nM)<br>temperature: 37$^o$ | Template: promoter type<br>concentration<br>linear or supercoiled<br>Salt: concentration<br>anion<br>(chloride or glutamate)<br>(NaCl or NaGlu) |

galactokinase gene (10). Cloning the 57 bp fragment in this manner placed the galactokinase gene under the transcriptional control of the late promoter (see figure 3A). The DNA sequences of the recloned promoters were confirmed by either the Maxam-Gilbert or dideoxy sequencing methodologies (16, 19).

Transcription Reactions

Table 2 describes the conditions and methodology of the in vitro transcription reactions. For most experiments the [DNA] was 0.06mM in terms of nucleotide monomer. Since the plasmids containing the various promoters are 3743 bp long, 0.06mM of nucleotide monomer is about 8nM of plasmid. The enzyme concentration in most reactions was about 2nM, or 50 units/ml. A unit was defined by Chamberlin as the amount of enzyme that can incorporate 1 nmol of UMP into RNA in 1 hr. (20). Many reactions were done using the buffer described in table 2 with no additional salt, these are refered to as low salt conditions. The incorporation of $^3$H-UMP into RNA was assayed by taking 40ul aliquots at various timepoints and spotting Whatman 3MM filters then washing the filters with trichloroacetic acid (21). The filters were counted in a liquid scintillation counter using a toluene based fluor.


Figure 4
Hypothetical mechanism for the appearance of the unexpected but frequent G to A mutation at +12.

## RESULTS

### Point Mutants Obtained

Of the 33 possible single point substitutions that can occur in the 11 bp region, 18 were obtained. They are: -9CT, -8TG, -8TC, -8TA, -7CT, -7CA, -7CG, -6AT, -5CT, -4TG, -4TC, -4TA, -3AT, -3AC, -3AG, -2TA, -2TC and +1GT. In addition the promoter from 16.0% of the T7 genome, that contains a G instead of a C at -5, was cloned into pKO1 in a manner similar to that used for the mutated promoters (see figure 3b).

### A Mutagenesis Artifact

6 of the 18 point mutants proved to contain an unexpected additional mutation outside the promoter sequence. The mutants -4TG, -4TC, -7CT, -7CG, -8TG and -9CT all have an A rather than a G at +12. This mutation probably has no effect on promoter function because: a) it is outside the conserved region of T7 late promoters, b) it does not affect the length of the transcript, c) there is no correlation between the presence of +12GA and the hierarchies of base preference observed, and d) it is outside the protected region when the promoter is tested in a footprint assay (22, 23, 24). A hypothetical mechanism for the appearance of the +12GA mutation is shown in figure 4. This mechanism requires the hybridization of 2 oligonucleotides to the template and the removal of the excess portion of the second oligonucleotide by the 3' exonuclease activity of Klenow fragment. T4 DNA ligase could then join the two oligonucleotides. Circumstantial evidence favors this hypothetical mechanism. The oligonucleotides used to make these mutants all have the same 5' end. Oligonucleotides used in other mutagenesis reactions have different 5' ends or would not form such a structure. These did not produce the +12GA mutation.

It is a good practice to check the sequence of the template for sites

Table 3 IN VITRO TRANSCRIPTION ACTIVITIES OF MUTATED PROMOTERS ON LINEAR TEMPLATES

| Promoter Type | Activity on Supercoiled Template | Activity on Linearized Template | linear / supercoiled |
|---|---|---|---|
| wildtype | 100 | 53 ± 11 | 0.5 |
| +1GT | 80 ± 5 | 20 ± 10 | 0.25 |
| -2TA | 94 ± 7 | 27 ± 6 | 0.3 |
| -2TC | 89 ± 10 | 20 ± 4 | 0.2 |
| -3AC | 95 ± 6 | 22 ± 7 | 0.2 |
| -3AT | 94 ± 6 | 11 ± 5 | 0.1 |
| -3AG | 90 ± 14 | 8 ± 4 | 0.1 |
| -4TA | 77 ± 5 | 1.3 ± 0.7 | 0.02 |
| -4TC | 75 ± 3 | nd | nd |
| -4TG | 95 ± 5 | 7 ± 2 | 0.07 |
| -5CT | 31 ± 9 | 2 ± 1 | 0.06 |
| -5CG | 88 ± 14 | 11 ± 6 | 0.13 |
| -6AT | 84 ± 9 | 6 ± 3 | 0.07 |

Reactions were run at 37°C using the standard buffer with no additional salt.

where the mutagenic oligonucleotide could hybridize and produce undesired mutations. A typical computerized search often ignores sites that have only a short segment homologous to the oligonucleotide because these sites are too numerous. However, the frequent occurrence of the unexpected +12GA mutation suggests that hybridization of an oligonucleotide can require only a short segment of homologous sequence if the secondary binding site is adjacent to the primary one. The sequences flanking the primary binding site should be carefully examined.

Transcription Assays

*In vitro* transcription assays were used to test the plasmids that contained the wildtype or mutated late promoters for the ability to function as templates for T7 RNA polymerase. None of the mutated promoters were more active than the wildtype promoter under any of the conditions tested. Figure 5 shows the results of typical *in vitro* transcription experiments, using supercoiled templates and low salt. These experiments indicated that there were two types of mutated promoters. Mutants that had activities similar to the wildtype were called type A. Mutants that were clearly less active than wildtype were called type B. Under the conditions used, type A mutants had activities that were at least 75% of the wildtype rate, but type B mutants less than 50%. The type A mutations occur at positions -4 to +1 and the type B mutations are found at -9 to -7. It may be significant that the two types of mutations occur in two distinct clusters. At positions -5 and -6 the

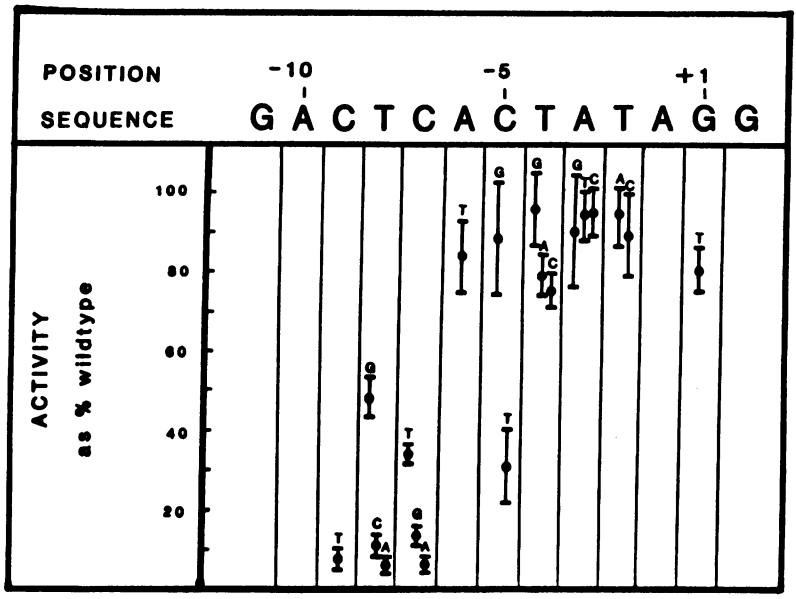


Figure 5
The in vitro transcription activities of the mutated promoters are shown relative to the wildtype promoter. Supercoiled templates and low salt conditions were used. Type A mutants are clustered in the upper right part of the figure, whereas type B mutants are clustered in the lower left.

mutations can be either type A or B. This suggests that there are two parts of the promoter that may have distinctly different functions.

From these experiments it was possible to determine the hierarchies of base preference at some positions. At positions -7 and -8 the complete sets of substitution mutations are available, and they are all type B. At positions -7 and -8 any substitution mutation reduces promoter activity but some mutations are worse than others. At position -7 the wildtype basepair, which is a C in the antisense strand, produces the most active promoter. If the C is changed to a T, the promoter activity is reduced. If the C is changed to a G or A, the promoter activity is reduced even further. Based on the activities of the promoters, the hierarchies of base preference at these two positions are: -7; C > T > G > A, and -8; T > G > C > A.

Determining hierarchies of base preference is a more difficult task for the type A mutations because they are nearly as active as the wildtype promoter when supercoiled templates and low salt conditions are used. This problem was approached by altering conditions of the transcription assays.

DNA Dependence Assays

The activities observed with the wildtype or mutated promoters are dependent on the concentration of template DNA. DNA dependence experiments using supercoiled and linearized templates were run over a range of 5 - 200 uM in terms of nucleotide monomer at a polymerase concentration of 2nM (data not shown). Over the range of DNA concentrations tested, the hierarchies of base preference seen at -7 and -8 remain unchanged and the type A mutants have activities that are indistinguishable from wildtype. (Except for -4TA and -4TC that are consistently less active than wildtype.) The activities observed at 60 uM are rather typical, so most of the data presented was obtained at this concentration.

Supercoiled vs. Linearized Templates

Under some conditions the type A mutated late promoters have lower activities than the wildtype promoter. The pKO1 derived plasmids can be cut once with Bgl I. These linear templates can be expected to produce a runoff transcript of 2844 bases in length. All type A mutated promoters are less active than wildtype when the template is linearized. The activities of the various mutants are affected to different degrees by loss of supercoiling, making it possible to see hierarchies of base preference for type A mutants. The activities of some mutants, such as -4TG and -4TA, are drastically reduced when the template is linearized. Table 3 shows the results of *in vitro* transcription experiments using linearized templates. The activity of a mutated promoter is given as a percentage of that seen for the wildtype promoter on supercoiled DNA. The conditions for the experiments shown in table 3 are the same as those used in figure 5. The degree to which the activity of a promoter is affected by the loss of supercoiling is expressed by the ratio shown in the righthand column of the table:

activity on linear / activity on supercoiled.

Promoters with lower ratios are more drastically affected by linearization.

Salt Dependence Experiments

Figure 6 shows the effects of added salt on the *in vitro* transcription activities of the wildtype and -4TA promoters. The addition of either NaCl or monosodium glutamate (NaGlu) reduces the activity of both promoters. For the wildtype promoter the effects of added NaCl and NaGlu are similar up to 140 mM, at higher concentrations NaGlu may be less inhibiting than NaCl. The NaCl dependence for -4TA drops more steeply than for the wildtype. Under low

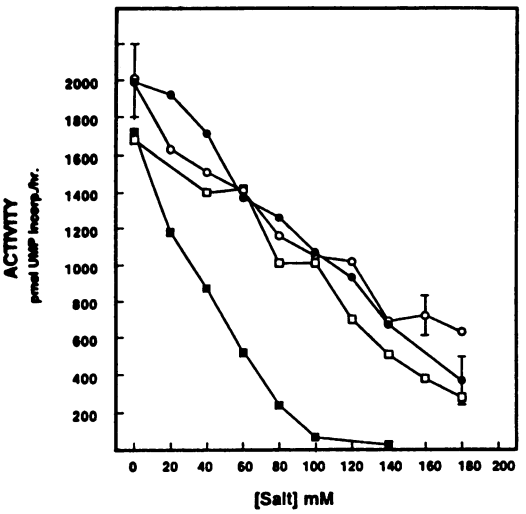


**Figure 6**
Salt dependence using either NaCl (closed symbols) or NaGlu (open symbols), with the wildtype promoter (circles) and -4TA (squares).

salt conditions -4TA is about 80% as active as wildtype, but at 60 mM NaCl, -4TA is about 40% as active as wildtype. The NaGlu dependence curve for -4TA does not drop as steeply as the NaCl curve. At lower concentrations of NaGlu the salt dependence of -4TA is similar to wildtype, however at higher concentrations NaGlu affect -4TA more than wildtype.

Table 4 shows the effects of added salt on the activities of the mutated promoters. All of the activities are given as a percent of the activity of the wildtype promoter under low salt conditions. The wildtype promoter activity drops to about 60% of the standard activity when 60mM NaCl is added. The effect of added salt on the activity of a given promoter is expressed as a ratio found in the righthand columns. Most of the mutated promoters are inhibited more than the wildtype by the addition of 60mM NaCl, and thus have lower ratios in the righthand columns. Type B mutants are more sensitive to added salt than the type A mutants. The effect of adding 120mM NaGlu is about the same as adding 60mM NaCl.

The addition of either 60mM NaCl or 120 mM NaGlu allows one to see hierarchies of base preference at some of the type A sites. The differences between one mutant and the next are very slight, but the order of base preference generally agrees with the results obtained with linearized templates. At -3 the wildtype A is the best, and C is marginally better than T which may be better than G. At -4; T > G > A,C. At -2 an A appears to be slightly better than C. Hierarchies at type B positions are unchanged.

Table 4 SALT EFFECTS ON THE ACTIVITY OF MUTATED PROMOTERS

| Promoter | ACTIVITY | | | 60 mM NaCl / no added salt | 120 mM NaGlu / no added salt |
|---|---|---|---|---|---|
| | no added salt | 60 mM NaCl | 120 mM NaGlu | | |
| wildtype | 100 | 60 ± 11 | 58 ± 10 | 0.6 | 0.6 |
| +1GT | 80 ± 5 | 41 ± 7 | 44 ± 4 | 0.5 | 0.55 |
| -2TA | 94 ± 7 | nd | 53 ± 4 | nd | 0.6 |
| -2TC | 89 ± 10 | nd | 46 ± 5 | nd | 0.5 |
| -3AT | 94 ± 6 | 45 ± 8 | 37 ± 7 | 0.5 | 0.4 |
| -3AC | 95 ± 6 | 53 ± 8 | 44 ± 6 | 0.6 | 0.5 |
| -3AG | 90 ± 14 | 36 ± 6 | 32 ± 7 | 0.4 | 0.4 |
| -4TA | 77 ± 5 | 29 ± 4 | 26 ± 4 | 0.4 | 0.3 |
| -4TC | 75 ± 3 | 29 ± 4 | 28 ± 8 | 0.4 | 0.4 |
| -4TG | 95 ± 9 | 40 ± 4 | 40 ± 12 | 0.4 | 0.4 |
| -5CG | 88 ± 14 | 54 ± 7 | 56 ± 8 | 0.6 | 0.6 |
| -6AT | 84 ± 9 | 23 ± 3 | 15 ± 2 | 0.3 | 0.2 |
| -5CT | 31 ± 9 | 3.5 ± 2 | 4 ± 2 | 0.1 | 0.1 |
| -7CA | 6 ± 1 | 1.4 ± .3 | 0.6 ± .1 | 0.2 | 0.1 |
| -7CG | 13 ± 2 | 2.1 ± .5 | 1.7 ± .3 | 0.2 | 0.1 |
| -7CT | 33 ± 2 | 3.7 ± .3 | 6 ± 2 | 0.1 | 0.2 |
| -8TA | 6 ± 1 | 1.8 ± .6 | 1.5 ± .5 | 0.3 | 0.25 |
| -8TC | 11 ± 2 | 1.9 ± .4 | 2 ± .5 | 0.2 | 0.2 |
| -8TG | 48 ± 5 | 10 ± 2 | 11 ± 3 | 0.2 | 0.2 |

Reactions were done with supercoiled templates, at 37° C.

Temperature Effects on Type I Mutated Promoters

Figure 7 shows a temperature dependence experiment with the wildtype promoter and 3 type A mutants using supercoiled templates. Temperature changes affect the activities of the wildtype and mutated promoters to the same degree. Two of the mutants tested in this experiment have composition changes in the region that melts upon the formation of the open complex. However the mutants do not show reduced activity relative to wildtype at lower temperatures.

Glycerol Effects on Type II Mutated Promoters

It has been previously shown that the addition of glycerol could enhance the relative activity of promoters with deletion mutations (8). This was due to both an increase in the activity of the mutated promoter and a decrease of the wildtype activity. Table 5 shows the effect of increased glycerol concentration on the relative activities of type B mutants. The effect of the added glycerol is expressed as the ratio:

activity at 12.6% glycerol / activity at 2.6 % glycerol.

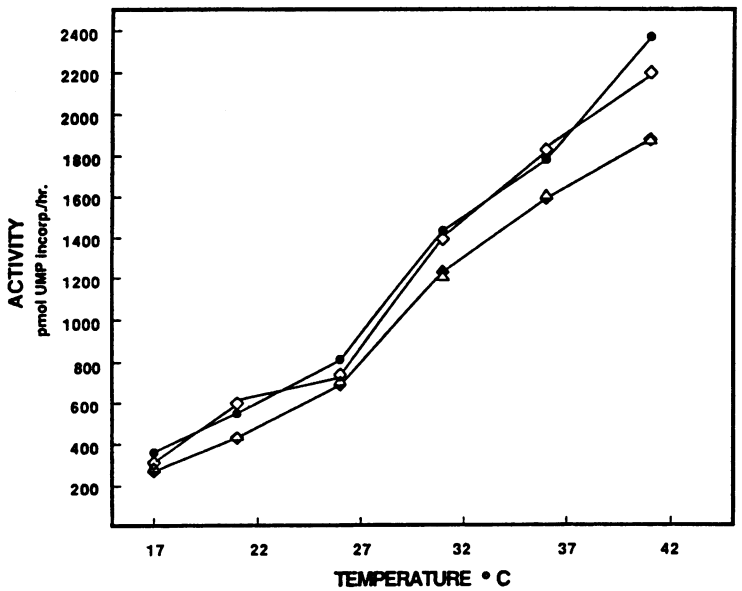


Figure 7
Temperature dependence of the transcription reaction for the wildtype and several type A mutated promoters. Supercoiled templates and low salt conditions were used. Symbols: closed circles are wildtype, closed diamonds are -4TG, open triangles are -4TA, and open diamonds are -3AC.

The activity of the wildtype promoter decreases with added glycerol, but the activities of the type II mutants stay the same or increase slightly. This is consistant with earlier observations. The glycerol experiment was run at higher DNA and RNA polymerase concentrations than the supercoiling and salt dependence experiments, so the activities of the type B promoters at 2.6% glycerol are higher than under the standard conditions given in table 3 or 4.

## DISCUSSION

### Transcription Experiments

The extreme specificity of the RNA polymerases from T7 and related bacteriophages insures that transcription will initiate only at a T7 late promoter, or a site that is very similar to a late promoter. This makes it possible to obtain meaningful results using the simple *in vitro* transcription assay described in this paper. One must keep in mind that the assay is for the rate of transcription, not initiation. But since the relationship between the rates of transcription and initiation is caused by the sequence

Table 5 GLYCEROL EFFECTS ON THE ACTIVITY OF MUTATED PROMOTERS

| Promoter | 2.6 % glycerol | 12.6 % glycerol | 12.6% glycerol / 2.6% glycerol |
|---|---|---|---|
| wildtype | 100 | 79 ± 3 | 0.8 |
| -5CT | 56 ± 6 | 58 ± 3 | 1.0 |
| -7CA | 14 ± 2 | 19 ± 1 | 1.3 |
| -7CG | 22 ± 2 | 31 ± 3 | 1.4 |
| -7CT | 60 ± 11 | 57 ± 6 | 0.9 |
| -8TA | 14 ± 4 | 22 ± 4 | 1.6 |
| -8TC | 24 ± 5 | 29 ± 4 | 1.2 |
| -8TG | 67 ± 10 | 65 ± 7 | 1.0 |
| -9CT | 7 ± 2 | 9 ± 2 | 1.3 |

Templates were supercoiled, no added salt, 37° C.

of the transcribed and surrounding region, the use of identical or very similar templates permits the quantitative comparison of one transcription assay with another. Using plasmids that are identical, except for point mutations in the late promoter, insures that the differences seen in transcription activities are due to differences in the interactions of the promoters with the RNA polymerase.

All of the single point changes examined by the use of this collection of mutated promoters affect one or more of the steps of the initiation process. The type A mutated promoters are less active than the wildtype promoter but only when the template is linearized or higher concentrations of salt are used. A change of one base pair can be sufficient to significantly alter the response of a promoter to fluctuations of either the degree of DNA supercoiling or the ionic conditions.

The addition of salt, either NaCl or NaGlu, inhibits the mutated promoters more than the wildtype, especially the type B mutants. Salt acts as a competitive inhibitor for site specific DNA binding by proteins (25). It also acts to increase helix stability (26) and thus could be acting on either the binding of the polymerase to the promoter or the isomerization of the polymerase-promoter from the open to the closed forms. Added salt could also affect the binding of NTPs and promoter clearance steps.

The addition of glycerol increases expression from type B promoters, perhaps by increasing the binding affinity of the polymerase to the promoter. Glycerol has been shown to increase the binding affinity of T7 RNA polymerase for the wildtype promoter (22). An increase in binding affinity could

conceivably reduce the expression from the wildtype promoter by reducing the promoter clearance rate. Alternatively, glycerol is known to destabilize the DNA double helix, so it might reduce the isomerization rate (27).

The Two Domain Promoter Hypothesis

The type B mutations at -9, -8, and -7 tend to be more deleterious than the type A mutations at -4, -3, -2 and +1. Mutations at -5 and possibly -6, may be either type A or type B. This grouping of mutations suggests they may be affecting different parts of the initiation process. We propose that the T7 promoter is divided into two functional domains. A binding domain from about -16 to about -5, and an initiation domain -4 to about +5.

Several lines of evidence support the two domain promoter hypothesis.

1) It has been shown that in DNA protection studies you can get two different sizes of footprint, depending on whether or not GTP is present. For the wildtype promoter the region protected by T7 RNA polymerase extends from -16 to -5 without GTP, but expands in the presence of 0.4mM GTP to cover the bases from -16 to +5 (22, 23). This indicates that there are two types of promoter complex. The protection observed in the absence of GTP, from -16 to -5, may be caused by polymerase binding to the promoter in a closed complex. The addition of GTP may induce or stabilize either an open complex or an intermediate form of closed complex from the isomerization pathway. The boundary between the type A and type B mutants falls at the right hand end of the footprint obtained without GTP.

2) A comparison of the late promoters from T7, T3 and SP6 shows some species-dependent differences and also some regions of identity. The differences tend to be to the left of -7. The region between -7 and +2 tends to be highly conserved. This could be caused by different rates of evolution in the two different domains or parts of the promoter. The initiation domain is highly conserved between the phage species, whereas the differences largely occur in the binding domain.

3) Strothkamp, Oakley and Coleman tested the effects of a single strand specific endonuclease on the polymerase-promoter complex (28). They determined that the binding of the polymerase melts the base pairs from -6 to +4, at least transiently. The lefthand boundary of the melted region is in about the same location as the right hand boundary of the footprint without GTP present, and the boundary between the type A and type B mutations.

4) Footprinting experiments indicate that the type B mutants, but not the type A mutants, reduce the binding affinity of the polymerase to the promoter (K. Chapman, S. Gunderson, and R. Burgess, unpublished results).

In 1980 McAllister and Carter proposed a three domain model for the late promoter consisting of, a "recognition" region from -17 to -1, an "initiation" region from +1 to +6, and a region from -13 to about -23 which they call the "melting" region (29). The major differences between the two models are: the left hand boundary of the "initiation" domain, and the role of the AT rich region from -13 to -22. Recent experiments indicate that the left hand boundary of the "initiation" domain is at about -5 (22, 28). The AT-rich region could conceivably destabilize the initiation site, but it might be involved in sequence specific contacts with the polymerase. The "melting" region hypothesis is based on the observations that increased salt selectively inhibits transcription from promoters which have shorter AT-rich regions, and that this inhibition can be reversed by DMSO. Increased salt could decrease the rate of initiation by increasing the stability of the DNA double helix, while the DMSO counteracts this effect. An alternative hypothesis is that the increased salt may preferentially reduce the binding of the polymerase to class II promoters, and that DMSO reverses this effect by increasing the binding affinity.

Hierarchies of Base Preference

It seems reasonable to hypothesize that there are important interactions between T7 RNA polymerase and late promoters that occur in the major groove.

Since there are several mechanisms by which a given base pair can affect the binding or interaction of a protein, it is difficult to know why a given mutation influences activity. But it is possible to make predictions about the hierarchy of base preference at a given site, if you make assumptions about the role of a base pair in the DNA-protein interaction. Base pair hierarchies cannot prove the importance of a given functional group, but they can be used to eliminate some of the possibilities.

There are several mechanisms that must be considered when guessing why a given site on the DNA looks unique to a binding protein. Some mutations are less deleterious than others because one new base pair may have some features similar to the wildtype base pair, while other base pairs may be less similar or even have unfavorable features. Base pairs have functional groups that protrude into the major and minor grooves and that can form hydrogen bonds and/or hydrophobic contacts with the amino acid residues of a binding protein. This aspect of DNA structure is an important part of models that have been developed for the specific binding of several proteins, such as the lac repressor, and the CI and cro repressors of bacteriophage lambda (30). The four base pairs have similarities as well as differences in regards to

the types and positioning of these functional groups (31, 32). However a given base pair may be important for other reasons. The AT content of a given DNA segment can affect the activity of nearby promoters (33). Also, point mutations could produce backbone conformational changes that significantly rearrange the charged phosphates. It is well documented that ionic attractions provide a substantial portion of the favorable free energy for both nonspecific and specific DNA-protein binding (25), and that even in B-DNA there are sequence specific variations in the propeller twist and the helix turn or helix rise per base pair (34).

Consider a T=A base pair as an example. If the N7 of the A is a key group, one could predict a hierarchy of: T > C > G > A, because C≡G preserves this group while A=T introduces a methyl group in it's position. If the key group is the 6-$NH_2$ of the A, or the keto O4 of the T, then then one might see a hierarchy of: T > G > A,C. Or, if the key group is the hydrophobic 5-methyl of the T, then the hierarchy may be T > C > A,G because the 5-H of C is more likely to be partially acceptable than the N7 of a purine. If a key group is in the minor groove, the hierarchy observed might be: T > A > C > G.

For a G≡C base pair, if the key group is the N7 of the G, the hierarchy G > A > C > T, might be observed. If the key feature is in the major groove on either the keto or amino group one might see: G > T > A, C. If the key contacts are in the minor groove then perhaps: G > C > A,T.

If the important feature of a base pair is its contribution to local helical thermal stability, composition is the relevant factor (35).

If an A=T base pair has an important effect on the conformation of the phosphate backbone, the observed hierarchy might be A,G > C,T. This is based on a theoretical method proposed by Dickerson to estimate helical-twist angle and base roll angle in B-DNA (36). The theoretical conformation of the helix backbone is determined by whether a base of interest is purine or pyrimidine.

In the T7 late promoter the hierarchy observed at -8 is: T > G > C > A. This is consistant with an interaction between the polymerase and the keto and/or amino groups located in the middle of the major groove. It is not consistant with phosphate backbone conformation theory (T,C > G,A), the base pair composition theory (T,A > C,G), minor groove contacts (T > A > C > G), sole contact at N7 of A (T > C > G > A), or contact with the methyl of the T (T > C > G,A). Similarily the hierarchy at -7 (C > T > G > A) is consistant with either an interaction at N7 of the G, or the phosphate backbone theory, but is inconsistant with the key interaction by the major groove keto or amino group (C > A > T > G), and with minor groove interactions (C > G > T >

A). At -4 the hierarchy (T > G > A,C,) is similar to that -8. And at -3 the hierarchy (A > C > T,G) is the same as for -4 and -8, reversing the strands.

Stahl and Chamberlin showed there are contacts between the polymerase and promoters in the minor groove, but they do not rule out the possibility of major groove contacts, except at the 5 position of purines (37).

ACKNOWLEDGEMENTS

We thank Dr. R.L. Niece of the UW Biotechnology Center for assistance in the synthesis of mutagenic oligonucleotides. This work was supported by the National Institutes of Health USA, Institutional Training Grants GM 07215 and CA 09135, additional support came from CA 07175 and CA 23076.

*Present address: MSU-DOE Plant Research Laboratory, Michigan State University, East Lansing, MI 48824, USA

References

1) M. Chamberlin, and J. Ring (1973) J. Biol. Chem., 248: 2235-2244
2) J. J. Dunn and F. W. Studier (1983) J. Mol. Biol., 166: 477-535
3) J. Oakley and J. Coleman (1977) Proc. Natl. Acad. Sci. USA, 74: 4266-4270
4) M. D. Rosa (1979) Cell, 16: 815-825
5) N. Panayotatos and R. D. Wells (1979) J. Mol. Biol., 135:91-109
6) M. Zoller and M. Smith (1983) Methods in Enzymol., 100: 468-500
7) P. Davenloo, A. Rosenberg, J. J. Dunn and F. W. Studier (1984) Proc. Nat. Acad. Sci. USA, 81: 2035-2039
8) K. Chapman and R. D. Wells Nuc. Acids Res., 10: 6331-6340 (1982)
9) J. Messing, B. Groenenborn, P. Muller-Hill and P. Hobschneider (1977) Proc. Natl. Acad. Sci. USA, 74: 3642-3646
10) C. Gross, A. Grossman, H. Liebke, W. Walter and R.R. Burgess (1984) J. Mol. Biol., 172: 283-300
11) J. Messing and J. Vieira (1982) Gene, 19: 269-276 (1982)
12) K. McKenney, H. Shimatake, D. Court, U. Schmeissner, C. Brady and M. Rosenberg in Gene Amplification and Analysis, vol. 2 Elsevier - North Holland, eds. Chirikjian and Papas (1981)
13) T. Maniatis, E. Fritsch and J. Sambrook (1982) in Molecular Cloning A Laboratory Manual page 68 (LB) Cold Spring Harbor
14) Davies and Jacob (1973) Histochemie, 37: 375
15) M. Matteucci and M. Caruthers (1981) J. Am. Chem. Soc., 103: 3185-3191
16) A. Maxam and W. Gilbert Methods in Enzymology 65, 499-549 (1977)
17) A. Banuszuk, K. Deugau and B. Glick (unpublished) Biologicals Incorporated, Toronto, Canada
18) L. Jolliffe, A. Carter and W. McAllister Nature, 299: 653-656
19) F. Sanger, S. Nicklen, and A. Coulson (1977) Proc. Natl. Acad. Sci. USA, 74: 5463-5467
20) M. Chamberlin, J. McGrath and L. Waskell (1970) Nature, 288: 227
21) M. Ryan and R.D. Wells (1976) Biochem., 15: 3778-3782
22) S. Gunderson, K. Chapman and R. Burgess (1987) Biochem., 26: 1539-1546
23) R. Ikeda and C. Richardson (1986) Proc. Natl. Acad. Sci. USA, 83: 3614-3618

24) S. Basu and U. Maitra (1986) J. Mol. Biol., 190: 425-438
25) T. Lohman, P. deHaseth and M.T. Record Biochem., 19: 3522-3530
26) S. Nakanishi, S. Adhya, M. Gottesman and I. Pasten (1975) J. Biol. Chem., 250: 8202-8208
27) S. Nakanishi, S. Adhya, M. Gottesman and I. Pasten (1974) J. Biol. Chem., 249: 4050-4056
28) R. Strothkamp, S. Oakley and J. Coleman (1980) Biochem., 19: 1074-1078
29) W.T. McAllister and A.D. Carter (1980) Nuc. Acids Res., 8: 4821-4837
30) C. Pabo and R. Sauer (1984) Ann. Rev. Biochem., 53: 293-321
31) N. Seeman, J. Rosenberg and A. Rich (1976) Proc. Natl. Acad. Sci. USA, 73: 804-808
32) J. Rosenberg and P. Greene (1982) DNA, 1: 117-124
33) H. Vollenweider, M. Fiandt and W. Szybalski (1979) Science, 205, 508-570
34) R. Dickerson (1983) J. Mol. Biol., 166: 419-441
35) R. Blake and P. Haydock (1979) Biopolymers, 18: 3089-3109
36) R. Dickerson (1984) Scientific American, Dec. 94-111
37) S. Stahl and M. Chamberlin (1979) J. Biol. Chem., 253: 4951-4959