Studies of Promoter Recognition and Start Site Selection by T7 RNA Polymerase Using a Comprehensive Collection of Promoter Variants†

Diane Imburgio, Minqing Rong,‡ Kaiyu Ma, and William T. McAllister*

Morse Institute of Molecular Genetics, Department of Microbiology and Immunology, State University of New York, Health Science Center at Brooklyn, 450 Clarkson Avenue, Brooklyn, New York 11203-2098

Received February 16, 2000; Revised Manuscript Received May 25, 2000

ABSTRACT: We have examined the behavior of T7 RNA polymerase (RNAP) at a set of promoter variants having all possible single base pair (bp) substitutions. The polymerase exhibits an absolute requirement for initiation with a purine and a strong preference for initiation with GTP vs ATP. Promoter variants that would require initiation at the normal start site (+1) with CTP or UTP result in a shift in initiation to +2 (with GTP). However, the choice of start site is little affected by base substitutions elsewhere in the initiation region. Furthermore, when the initiation region is shifted either one nucleotide (nt) closer or 1 nt further away from the binding region, transcription still begins the same distance downstream. These results indicate that the sequence around the start site is of little importance in start site selection and that initiation is directed a minimum distance of 5 nt downstream from the binding region. At promoters that initiate with +1 GGG, T7 RNAP synthesizes a ladder of poly(G) products as a result of slippage of the transcript on the three C residues in the template strand from +1 to +3. At promoter variants in which there is an opportunity to form a longer RNA−DNA hybrid, this G-ladder is enhanced and extended. This observation is not in agreement with recent suggestions that the RNA−DNA hybrid in the initiation complex cannot extend further than 3 bps upstream from the active site [Cheetham, G., Jeruzalmi, D., and Steitz, T. A. (1999) *Nature 399,* 80−83].

The T7 genome contains 17 promoters recognized by T7 RNA polymerase (RNAP), all of which are related to a 23 base pair (bp) consensus sequence that extends from −17 to +6 (Figure 1). These promoters are referred to as class II, class III, or replication promoters on the basis of their location, temporal utilization, and function (ref *1* and references therein). While the sequence of class III promoters is identical to the consensus sequence, the class II promoters differ from the conserved sequence at two or more positions and are generally weaker than the class III promoters. During infection, transcription from class II and class III promoters is regulated by T7 lysozyme, which preferentially inhibits transcription from class II promoters (*2*−*7*). Promoters for other phage RNAPs (e.g., T3, SP6, and K11) exhibit a similar consensus sequence and contain a core of bps from −7 to −3 that are identical to those of the T7 promoter, suggesting that this region may function in a similar manner for all phage RNAPs. On the other hand, the phage promoters differ from −12 to −8, consistent with studies showing that this region is responsible for specific promoter recognition and binding (*8*−*10*). Certain positions (e.g., −14, −7, −6, −4, and −3; shaded in Figure 1) are invariant among all phage promoters.

Earlier studies indicated that the T7 promoter is composed of two functional domains—an upstream binding region from −17 to −5 and an initiation region from −4 to +6 (*11*−*14*). In general, base substitutions in the upstream region were found to have large effects on RNAP binding but little effect on initiation, whereas substitutions in the downstream region have a greater effect on initiation but little effect on binding (*11*−*14*).

A variety of experimental approaches indicated that the polymerase recognizes one face of a closed DNA duplex in the binding region, and that the initiation region is melted open downstream from −5 (*13*−*18*). A collection of T7 promoter variants having single bp substitutions in the binding region was particularly useful in deducing the nature of specific promoter recognition and in identifying structural elements in the RNAP that interact with these determinants (*19*−*21*). The results of recent X-ray structural analysis of T7 RNAP complexed with its promoter are consistent with these earlier studies and reveal additional contacts not previously identified (*22*). To summarize, specific recognition of the T7 promoter involves interactions between bps in the major groove from −7 to −11 with amino acid residues in a specificity loop (residues 739−770) that projects into the DNA binding cleft of the RNAP. Additional contacts in the AT-rich region from −17 to −13 involve a flexible loop (the AT-rich recognition loop; residues 93−101) that extends into the minor groove. The transition from duplex DNA in the binding region to single-stranded DNA in the initiation region occurs between bps −5 and −4 and is stabilized by interactions with a β-hairpin that includes Val237, while the template strand is led down into the active site by numerous additional contacts.

† Supported by NIH Grant GM38147 to W.T.M.

* To whom correspondence should be addressed: Phone 718-270-1238; fax 718-270-2656.

‡ Present address: Department of Physiology and Biophysics, Mount Sinai School of Medicine, One Gustave Levy Place, New York, NY 10029.

10.1021/bi000365w CCC: $19.00
Published on Web 08/04/2000

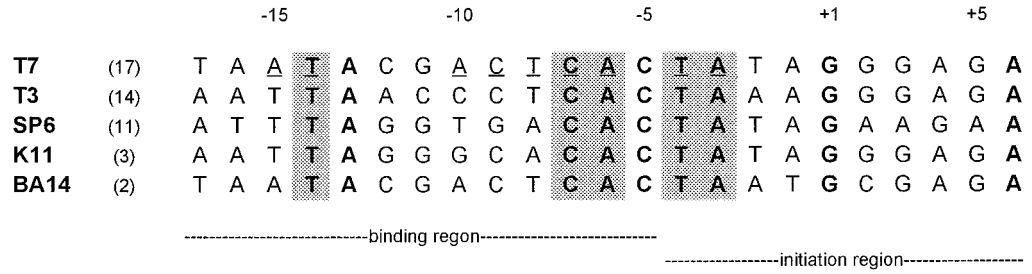

FIGURE 1: Comparison of phage promoter consensus sequences The sequence of the nontemplate strand is shown. The number of naturally occurring promoter sequences that have been determined for each set is given in parentheses (*19*, *61*). Positions that are shaded are invariant in all 47 natural phage promoter sequences. Positions in boldface type are common among all consensus promoters. Positions that are underlined are invariant in the 17 naturally occurring T7 promoters. The binding region (−17 to −5) and initiation region (−4 to +6) of the T7 promoter are indicated.

Events that occur after promoter binding and melting are less well understood. As with all RNA polymerases, T7 RNAP engages in a process of abortive initiation in which short nascent transcripts are continuously synthesized and released until the polymerase clears the promoter and forms a stable elongation complex (for review, see (ref *23*). The transition to a stable complex is accompanied by loss of upstream promoter contacts, enhanced retention of the RNA product, and isomerization of the ternary complex (as revealed by a more compact footprint and changes in protease sensitivity) (*24*−*26*). Promoters that deviate from the consensus sequence in the initiation region exhibit a lowered efficiency of promoter clearance and an altered pattern of abortive initiation (*27*−*30*).

Although a number of T7 promoter variants have been cloned and characterized (*11*, *31*, *32*) the sequence context and the conditions under which they were analyzed vary, making direct comparisons difficult. Furthermore, many of the variants contain changes at multiple positions. In this work, we describe the construction and characterization of a collection of T7 promoter variants having all possible single bp substitutions over the entire region from −17 to +6. These studies extend previous work that described single bp variants only in the region from −15 to −6 (*19*, *33*, *34*). We anticipate that this collection of promoters will prove useful in characterizing a number of aspects of promoter function.

MATERIALS AND METHODS

Construction of Promoter Templates. Desired mutations were introduced into the test promoter (P_x) in pGD13 by use of the polymerase chain reaction and mismatched primers as described in Diaz et al. (*19*). To introduce mutations at −17, −16, −5, and −4, the mutagenic primer was ATCGATCTGCAG**TA**ATACGACTCA**CT**ATAGGGAGAGG, which is degenerate at these positions (boldface type) and includes a *Pst*I site (underlined) upstream of the promoter. The second primer (ATAGCGCTAGCAGCACG) is complementary to a region that lies 228 bp downstream from P_X and contains a unique *Nhe*I site (underlined). The amplified products were digested with *Pst*I and *Nhe*I, purified from an agarose gel, and cloned back into the parental construct at the same restriction sites. Substitutions in the region from −3 to +6 utilized the mutagenic primer CGCATTGGATCC**TCTCCCTAT**AGTGAG and a second primer (GGCGTATCACGAGGCCC) that is complementary to a region 80 bp downstream from P_X. The amplified products were digested with *Bam*HI and *Eco*RI, purified from an acrylamide gel, and cloned back into the parental plasmid at the same restriction sites.

Plasmid DNA was purified by the Promega large-scale maxiprep system. Synthetic DNA oligomers were purchased from Macromolecular Resources (Fort Collins, CO).

Transcription Conditions. Histidine-tagged forms of wild-type T7 RNAP and the mutant enzyme Y639F (*35*) were purified as previously described (*36*). To determine relative promoter strength, reactions (10 μL) contained 1.0 μg of template DNA, 0.1 mM ATP, 0.5 mM CTP, GTP, and UTP (Pharmacia), 2.0 μCi of [α-^{32}P]ATP (New England Nuclear) and 10 ng of wild-type RNAP in GHT buffer [30 mM Hepes, pH 7.8, 0.25 mM EDTA, 1 mM dithiothreitol (DTT), 15 mM $Mg(OAc)_2$, 0.05% Tween 20, and 0.1 mM potassium glutamate] (*37*). Reactions were incubated at 37 °C for 15 min and terminated by chilling on ice and the addition of 10 μL 2× stop buffer (6 M urea, 0.01 M EDTA, 0.02% bromophenol blue, and 0.02% xylene cyanol). Following heat denaturation (100 °C, 2 min), aliquots (10 μL) of the reactions were resolved by electrophoresis in a 6% denaturing polyacrylamide gel (6 M urea and 0.1% SDS) in 1× TBE buffer (90 mM Tris−Borate and 2 mM EDTA) at 275 V for approximately 2 h (*38*). The gel was fixed with 10% acetic acid and 10% methanol, dried, exposed to a PhosphorImager screen, and analyzed by use of ImageQuant software.

To characterize abortive initiation, choice of start site, and poly(G) synthesis, reactions (10 μL) contained 1 μg of template, 20 ng of RNAP, and 2−3 μCi of [γ-^{32}P]ATP, [γ-^{32}P]GTP, or [α-^{32}P]GTP (as indicated) in a Tris buffer system (20 mM Tris-HCl, pH 7.9, 15 mM $MgCl_2$, 1 mM DTT, 0.1 mM EDTA, and 0.05% Tween 20). Abortive initiation assays included all four NTPs (0.5 mM). For start site experiments, NTPs and chain-terminating analogues of CTP and/or UTP were present at 0.5 mM. 3′-dUTP and 3′-dCTP were provided by Dr. Harry Osterman, P−L Pharmacia Laboratories; ddTTP and ddCTP were purchased from Amersham Pharmacia Biotech. For experiments involving poly(G) synthesis, reactions contained 0.5 mM GTP and 3.0 μCi of [γ-^{32}P]GTP. Reactions were incubated at 37 °C for 15 min and the products were resolved by electrophoresis in a highly cross-linked (12:1 acrylamide:bisacrylamide) denaturing 20% polyacrylamide gel (*38*).

RESULTS

Construction of Promoter Variants and Analysis of Promoter Strength and Specificity. T7 RNAP promoter variants were constructed by polymerase chain reaction with

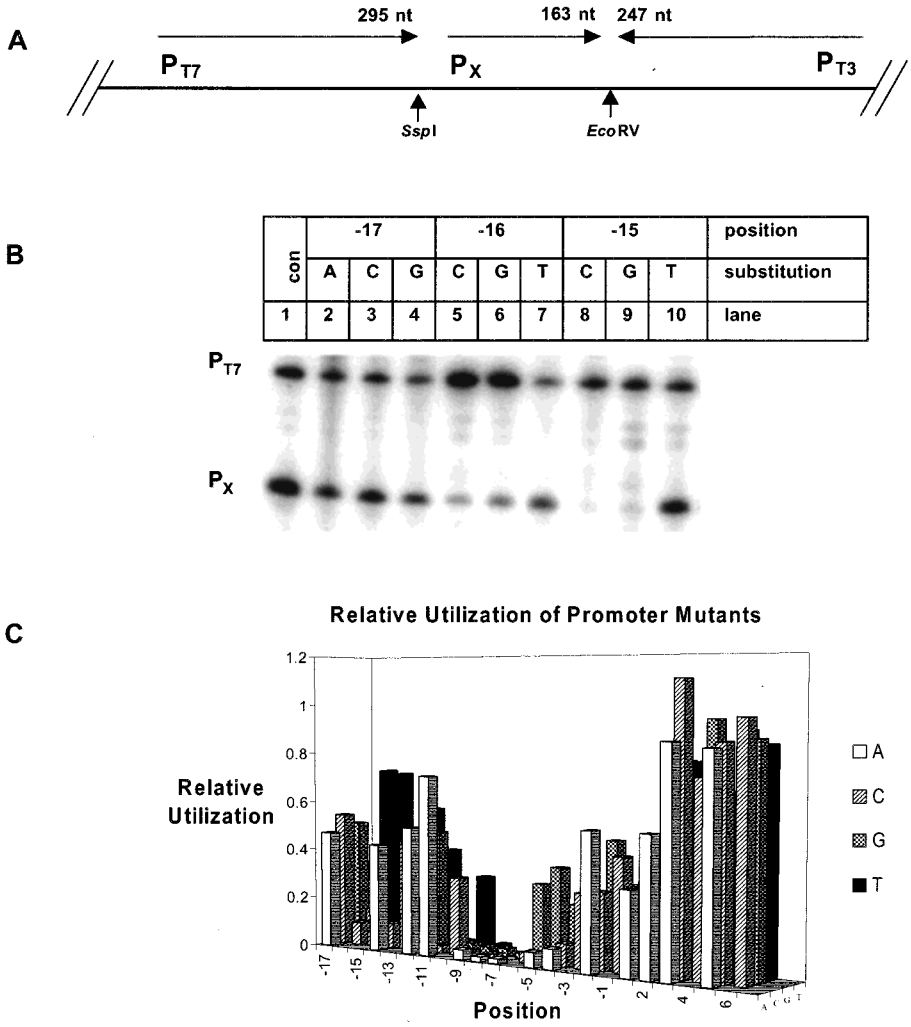

FIGURE 2: Characterization of T7 RNAP promoter variants Panel A. Template structure. Plasmid templates contain a test promoter (P_X) having the desired substitution, a reference promoter (P_{T7}), and a promoter for T3 RNAP (P_{T3}). Digestion with *Eco*RV and *Ssp*I results in the synthesis of 163, 295, and 247 nt runoff products, respectively, from these promoters (*19*). (B) Analysis of promoter activity. A representative promoter strength assay is shown for promoters having the indicated bp substitutions at −17, −16, or −15. Products from a control template in which P_X corresponds to the consensus promoter are shown in lane 1. Reaction products were resolved by electrophoresis in a denaturing 6% polyacrylamide gel and visualized by PhosphorImager analysis. In these experiments, the template was present in 5-fold molar excess of RNAP (50 nM template vs 10 nM RNAP). Assuming a K_d for the promoter−polymerase complex of 5 nM (*62*), the enzyme is expected to be >90% saturated with promoter under these conditions. The reduced production of total transcripts in lane 8 may be due to poor initiation (but good binding) by T7 RNAP on the −16T promoter variant, which would sequester the RNAP in poorly active initiation complex; further experiments will be needed to address this question. (C) Relative activity of promoter variants. The height of each bar indicates the activity of each promoter relative to that of the consensus promoter (data are from Table 1).

mismatched primers and the plasmid pGD13 as a template ((*19*); see Figure 2). The resulting plasmids contain two promoters recognized by T7 RNAP [a reference promoter (P_{T7}) and the test promoter (P_X)] and a promoter recognized by T3 RNAP (P_{T3}). The utilization of each of these promoters may be determined by comparing the production of a runoff product of characteristic size from each promoter (Figure 2). The results are summarized in Table 1. For convenience, we identify promoter variants by referring to the base in the nontemplate strand of the DNA (e.g., a −17A promoter).

Specific promoter recognition involves interactions of amino acid residues in the specificity loop with base pairs in the binding region. In earlier studies, we described the effects of a number of changes in the specificity loop on promoter specificity and RNAP activity (*19*−*21*; and see Table 2). However, we had not previously reported our results with regard to residue R756 and its interaction with the bp at −9 (*39*). In the crystal structure of a T7 RNAP−promoter complex this residue is observed to make direct hydrogen bonds with the G in the template strand at −9 (*22*). The effects of substitutions of R756 are complex and affect the preference of the mutant RNAP not only for the bp at −9, but also for the bps at −10 and −11 (Figure 3). As noted below, these alterations are likely to reflect interactions between R756 and the amino acid at position 748 (N748), which contacts the bps at −10 and −11.

Additional potential contacts had been identified in the region upstream from the recognition region. Specifically, UV-laser cross-linking studies revealed an intimate contact between the RNAP and the bp at −17, and fluorescent interference studies suggested close approaches of the polymerase to −17 and −15 (*40*, *41*). These interactions were confirmed in the crystal structure, where an A·T rich recognition loop (amino acid residues 93−101) was found to be inserted into the minor groove from −17 to −13. In the region from −16 to −14, A·T and T·A substitutions are generally well

Table 1: Utilization of T7 Promoter Variants

plasmid[a]	P_X[b]	consensus base[c]	naturally occurring promoters[d]	relativ promoter strength[e]	plasmid[a]	P_X[b]	consensus base[c]	naturally occurring promoters[d]	relativ promoter strength[e]
pGD13	consensus			1.00	pDP38	−5A	C		0.06
					pDP39	−5G	C	ϕ1.3	0.37
pDP2	−17A	T		0.47	pDP40	−5T	C		0.04
pDP3	−17C	T	ϕ1.1A, 4c, 4.7	0.54					
pDP4	−17G	T		0.50	pDP41	−4A	T		0.08
					pDP42	−4C	T		0.09
pDP5	−16C	A		0.09	pDP43	−4G	T		0.23
pDP6	−16G	A		0.14					
pDP7	−16T	A	ϕ 4.7	0.72	pDP44	−3C	A		0.29
					pDP45	−3G	A		0.21
pGD44	−15C	A		0.06*	pDP46	−3T	A		0.27
pGD45	−15G	A		0.11*					
pGD46	−15T	A		0.71*	pDP47	−2A	T	ϕ 1.5, 1.6, 3.8, 4c, 4.3	0.54
					pDP48	−2C	T		0.30
pGD47	−14A	T		0.43*	pDP49	−2G	T		0.49
pGD48	−14C	T		0.10*					
pGD49	−14G	T		0.10*	pDP50	−1C	A		0.44
					pDP51	−1G	A		0.32
pGD50	−13C	A	ϕ4c	0.44*	pDP52	−1T	A	ϕ 2.5	0.33
pGD51	−13G	A		0.47*					
pGD52	−13T	A	ϕ 3.8, 4.7	0.57*	pDP53	+1A	G	ϕ OL, 2.5	0.33
					pDP54	+1C	G		0.08
PDP17	−12A	C		0.51*	pDP55	+1T	G		0.18
PDP18	−12G	C	ϕ 3.8	0.48*					
PDP19	−12T	C		0.40*	pDP56	+2A	G	ϕ1.1A, 4c	0.54
					pDP57	+2C	G		0.50
pGD17	−11A	G	ϕ OL, 3.8	0.72*	pDP58	+2T	G		0.80
pGD18	−11C	G		<0.03*					
pGD19	−11T	G		<0.03*	pDP59	+3A	G	ϕ 1.1B,1.5,1.6,4.3,4.7	0.88
					pDP60	+3C	G		1.11
pGD14	−10C	A		0.31*	pDP61	+3T	G		0.80
pGD15	−10G	A		0.04*					
pGD16	−10T	A		0.30*	pDP62	+4C	G		0.75
					pDP63	+4G	A	ϕ1.1A, 1.1B, 1.5, 1.6, 2.5,4.3,4.7	0.96
pGD27	−9A	C		0.04*	pDP64	+4T	A		0.69
pGD28	−9G	C		<0.03*					
pGD29	−9T	C		<0.03*	pDP65	+5A	G	ϕ 1.6, 2.5, 4.3, 4.7	0.86
					pDP66	+5C	G	ϕ1.3	0.88
pGD30	−8A	T		<0.03*	pDP67	+5T	G		0.92
pGD31	−8C	T		<0.03*					
pGD32	−8G	T		<0.03*	pDP68	+6C	A	ϕ 1.6, 4.3	0.97
					pDP69	+6G	A		0.89
pGD33	−7A	C		<0.03*	pDP70	+6T	A	ϕ 1.5, 4.7	0.87
pGD34	−7G	C		<0.03*					
pGD35	−7T	C		<0.03*					
pGD41	−6C	A		0.04*					
pGD42	−6G	A		0.30*					
pGD43	−6T	A		0.06*					

[a] Plasmids in the pDP series were constructed in this work; those in the pGD series were described previously (*19*). [b] Promoter variants are designated as $P_{-n}X$, where *n* indicates the position in the promoter and X indicates the base in the nontemplate strand. [c] The base in the nontemplate strand of the consensus promoter at the indicated position. [d] Naturally occurring T7 RNAP class II promoters that contain the indicated substitution; all class II promoters have two or more substitutions. [e] Plasmid templates were transcribed as described in Figure 2, and the products were resolved by gel electrophoresis. After the number of adenosine residues (the labeled substrate) encoded in each transcript was taken into account, the data for each lane were normalized to the internal control (P_{T7}). The utilization of each mutant promoter was then expressed relative to that of the consensus promoter when it was present at the position of P_X (*19*). Data obtained for the previously described pGD series (*19*) are indicated with an asterisk, and are presented here for comparison.

tolerated, while C·G or G·C substitutions are not (Table 1). The preference for A·T or T·A bps in this region may reflect a need for flexibility during promoter recognition and initiation but would also be consistent with base-specific readout in the minor groove (*42*, *43*).

The transition from duplex DNA in the binding region to a melted form in the initiation region occurs between bps −5 and −4, and is stabilized by intercalation of a β-hairpin structure (residues 229−243) and stacking interactions between Val237 and the template strand base at −5 (*22*). There appears to be a moderate degree of specificity in this region of the promoter. Thus, while substitution of G is tolerated at −6, substitutions of C or T are not. Further downstream, substitution of G is tolerated at −5 while A or T substitutions are not, and at −4 there is tolerance for G but little tolerance for A or C substitutions. Beyond −4 (−3 to +6) the polymerase appears to tolerate nearly all substitutions, except at the start site for transcription (+1).

Table 2: Properties of T7 RNAP Mutants with Substitutions in the Specificity Loop

mutation[a]	plasmid[b]	activity at consensus promoter[c]	nonspecific catalytic activity[d]
L749M*	CAR52	NC	+++
M750I*	CAR29	NC	+++
M750A	BH218	NC	+++
M750V	BH219	NC	+++
F751L	BH215	NC	+++
G753A,Q754E	MR81	NC	+++
G753R,Q745T	MR82	NC	+++
Q754A	MR77	NC	+++
Q754N	MR78	NC	+++
Q754G	MR79	NC	+++
Q754D	MR80	NC	+++
R756E	MR85	none	+++
R756A	MR86	CH	+++
R756Q	MR87	CH	+++
R756S	MR88	CH	+++
R756K	MR89	CH	ND
R756D	MR90	none	ND
R756C	MR91	CH	ND
R756N,Q758K	MR96	CH	+++
T760A	MR108	low	+++
T760S	MR109	NC	+++
T760D	MR110	low	ND
N762A	BH217	NC	+++
N762E	MR84	NC	+++
T763R[e]	MR58	ND	ND
T763H	MR59	NC	+++
N764T	MR60	NC	+++
K765R	MR61	NC	+++
K765E	MR92	NC	ND
D766E	MR94	NC	ND
S767E	MR95	NC	ND

[a] Mutations are identified by position of the residue; the letter to the left indicates the wild-type residue, the letter to the right indicates the substitution. Mutants identified with asterisks have been previously described (*56*). [b] The plasmid that carries the indicated RNAP. [c] The activity of the mutant enzyme at the consensus promoter vs promoters with substitutions from −7 to −11 was determined as described in Figure 3. NC indicates no change in activity or specificity; low indicates decreased activity but no change in specificity; CH indicates change in specificity. [d] Nonspecific catalytic activity was determined as synthesis of poly(G) on a poly(dC) template; +++ indicates high activity (*39*). [e] T673R is highly sensitive to protease and could not be tested.

T7 polymerase prefers to initiate with GTP but has also been observed to initiate with ATP (*1*). In keeping with this observation, substitution of a T in the template stand (which directs initiation with ATP) is tolerated, while substitutions that direct initiation with CTP or UTP are not.

Previous work suggested that the base at −1 might be important in start site selection. Weston et al. (*44*) proposed that base stacking interactions between the −1 and +1 bases in the template strand contribute to the specificity of initiation, and Cheetham et al. (*22*) identified an interaction of Trp422 with the template base at −1. However, substitutions at −1, while reducing activity slightly, reveal little base specificity at this position. Furthermore, these substitutions have little effect on the choice of start site (see below).

Although the bps at certain positions are conserved among all phage promoters (i.e., −14, −7, −6, −4, and −3; see Figure 1), this does not necessarily indicate that the polymerase will not tolerate substitutions at these positions. While no substitutions are acceptable at −7, the polymerase will tolerate substitution of A at −14, G at −6, G at −4, and nearly all substitutions at −3 (Table 1).

Effects of Substitutions in the Initiation Region on Start Site Selection. The use of plasmid templates that contain both a reference promoter and a test promoter was convenient in the experiments described above because these templates provide an internal standard for activity (P_{T7}). However, to examine the effects of bp substitutions on other aspects of promoter function such as choice of start site or synthesis of abortive initiation products, the presence of two promoters in these templates makes it difficult to identify products that arise solely from the test promoter. In subsequent studies it was therefore necessary to subclone individual mutant promoters into a background that contained no other promoters (Table 3).

To characterize the choice of start site, we carried out transcription under conditions of limiting substrate in order to limit synthesis to the production of short products whose lengths could be readily determined (Figure 4). The consensus promoter in pDP71 initiates with the sequence GGGAGAGGATC... Transcription of this template is expected to give rise to a 10 nt product in the presence of GTP, ATP, and UTP and to a runoff product of 21 nt in the presence of all four NTPs. However, an additional product 1 nt larger ($n + 1$) is observed under both conditions (lanes 3 and 4). This phenomenon has been attributed to the addition of a non-template encoded base at the 3′ end of the transcript (*28*) and/or to slippage of the RNA on the template during the early stages of initiation (*45*). To prevent incorporation of additional nucleotides at the 3′ end of the transcript, we employed 3′-deoxy chain terminators. As shown in Figure 4, when UTP is replaced by 3′-deoxy-UTP (dU; lane 5) synthesis of the extraneous (11 nt) product is eliminated. The extra band observed in our experiments is therefore due to the nontemplated addition of a single nt to the 3′ end of the nascent RNA and does not arise by slippage at the initiation site. Using dUTP to prevent the addition of the extra nt, we examined the effects of changes around the start site in various promoter variants, labeling the resulting transcripts either with [γ-^{32}P]GTP or [γ-^{32}P]ATP (lanes 6−16).

Transcription from the +1A promoter in the presence of [γ-^{32}P]ATP results in the synthesis of a discrete product of 10 nt, indicating initiation at +1 (Figure 4, lane 13). However, some initiation also occurs at +2 at this promoter, as evidenced by the synthesis of a 9 nt product labeled with [γ-^{32}P]GTP (lane 7). Weaker bands above 9 and 10 nt are observed under both conditions, indicating some slippage during initiation at this promoter. Although T7 RNAP can initiate with either GTP (at +2) or with ATP (at +1) at this promoter, it prefers to start at +1, as demonstrated by the predominant synthesis of the 10 nt product, when the transcripts are labeled with [α-^{32}P]GTP (see Figure 5B, lane 5).

At the +4G promoter, initiation occurs at +1 (as evidenced by production of a 10 nt product; Figure 4, lane 10), but enhanced synthesis of products >10 nt is also observed. The greater synthesis of the latter products probably reflects an increased opportunity for slippage of the nascent transcript at this promoter due to the longer uninterrupted run of C residues in the template strand from +1 to +5.

At the −1G promoter, strong initiation at +1 was observed (Figure 4, lane 6, 10 nt product) but enhanced synthesis of products >10 nt was again observed. At this promoter, these products could arise either by slippage of the nascent RNA or by initiation at −1 (see below).

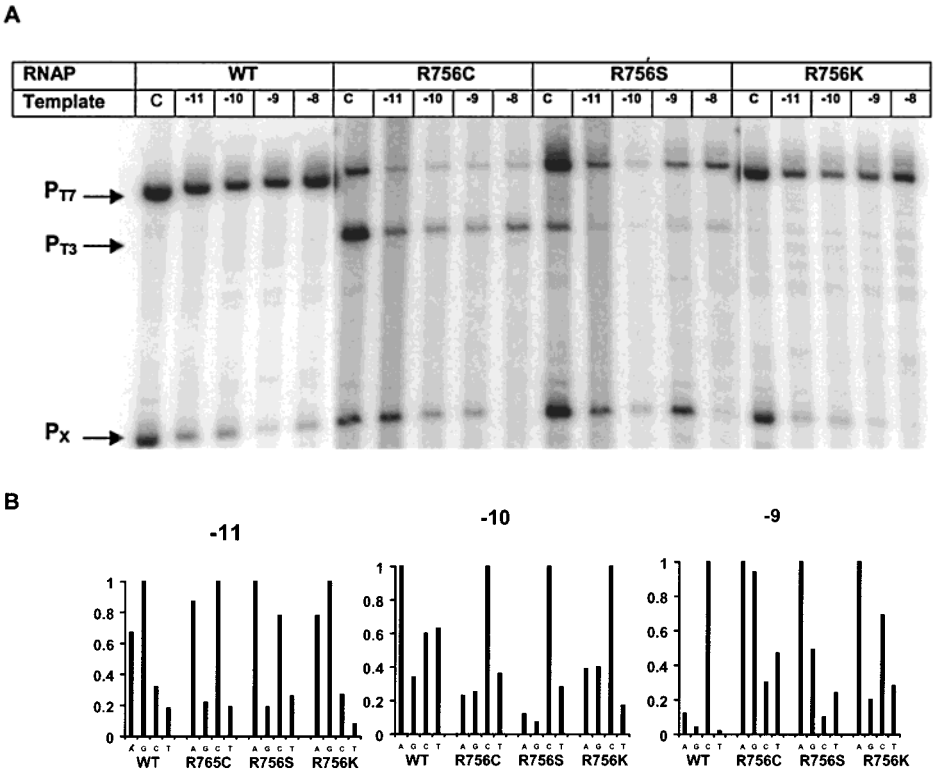

FIGURE 3: T7 RNAP mutants with altered promoter specificities. (A). Plasmid templates were prepared as described as in Figure 2 and transcribed with the RNAP indicated. Reactions in lanes marked C contained a control plasmid in which the consensus T7 promoter was present at P_X. Reactions in lanes marked −11 through −8 contained a mixture of three plasmid templates, each having a promoter with one of the nonconsensus bps at the indicated position. Note that the wild-type (WT) enzyme preferred the reference promoter (P_{T7}) over the nonconsensus mixtures. In contrast R756C preferred a T3 promoter (which differs from the T7 promoter at −10 and −11) to the consensus T7 promoter (lane C), and furthermore preferred templates having a nonconsensus bp at −11, −10, and −9, but not at −8. Similar results were obtained with R756S, but in this case a stronger preference for promoters with substitutions at −9 was observed. (B). The preference of the wild-type, R756K, R756C, and R756S RNAPs for promoter variants with individual bp substitutions at −11, −10, and −9 was determined as described in Figure 2. Each panel represents the activity of the RNAP at a T7 promoter having the substitution indicated. The activities in each set have been normalized to the most active promoter in the set. Thus, at −11 the WT RNAP prefers −11G whereas R756C prefers −11C and R756S prefers −11A. At −10 the WT RNAP prefers −10A while the mutant enzymes prefer −10C. At −9, the WT RNAP prefers −9C while the mutant enzymes prefer −9A.

Table 3: Single-Promoter Constructs

plasmid[a]	promoter	plasmid[a]	promoter
pDP71	consensus	pDP79	+1A
pDP72	−5G	pDP80	+1C
pDP73	−5T	pDP81	+1T
pDP74	−4C	pDP91	+2A
pDP75	−4G	pDP92	+2C
pDP76	−3G	pDP93	+2T
pDP77	−2C	pBH199[b]	+3T
pDP89	−1C	pDP82	+4G
pDP78	−1G	pBH197[b]	+4T
pDP90	−1T		

[a] The promoter indicated was excised from the plasmid described in Table 1 by digestion with *Bam*HI and *Kpn*I and subcloned into the corresponding sites of pUC19. [b] Plasmids pBH197 and pBH199 were incorrectly identified in He et al. (*60*); the correct assignments are given here.

Substitution of C or T at +1 results in initiation at +2 (as evidenced by synthesis of a 9 nt product labeled with [γ-^{32}P]-GTP; Figure 4, lanes 8 and 9). However, we cannot exclude initiation at +1 on the basis of this experiment as we did not employ [γ-^{32}P]CTP or [γ-^{32}P]UTP to label the products. To examine this in more detail, we carried out similar experiments with [α-^{32}P]GTP as the label, employing ddCTP or ddTTP as chain terminators and taking advantage of a mutant RNAP (Y639F) that is able to incorporate dideoxynucleotides (*35*) (Figure 5). In the presence of GTP, ATP, and UTP, the wild type (WT) polymerase again gave rise to *n* (10 nt) and *n* + 1 (11 nt) products, which was suppressed in the presence of dU (Figure 5A, lanes 1 and 2). When presented with ddTTP, the WT enzyme mostly terminated before the incorporation of the analogue, resulting in production of a 9 nt product and small amounts of the 10 nt product (Figure 5A, lane 3). The mutant enzyme gave rise to *n* and *n* + 1 products in the presence of GTP, ATP, and UTP (Figure 5A, lane 4), but this was completely suppressed by the addition of either dU or ddTTP (Figure 5A, lanes 5 and 6). Similar results were observed with GTP, ATP, UTP, and ddCTP (Figure 5A, lanes 8, 9, 11, and 12), however; due to the sequence of the template, transcription is extended 1 nt further under the latter conditions, resulting in the synthesis of an 11 nt product. In the presence of all four rNTPs (Figure 5A, lanes 7 and 10) both enzymes produced runoff products of the expected lengths.

Using this procedure, we observed initiation at +1 at the −1C, −1G, +1A, +2A, +2C, +4G, −1T, and +2T promoters, as evidenced by production of a discrete 10 nt product (Figure 5B, lanes 3, 4, 5, 7, 8, 9, 11, and 13, respectively). However, transcription from the +1C and +1T promoters resulted in synthesis only of a 9 nt product, indicating initiation at +2 (Figure 5B, lanes 6 and 12).

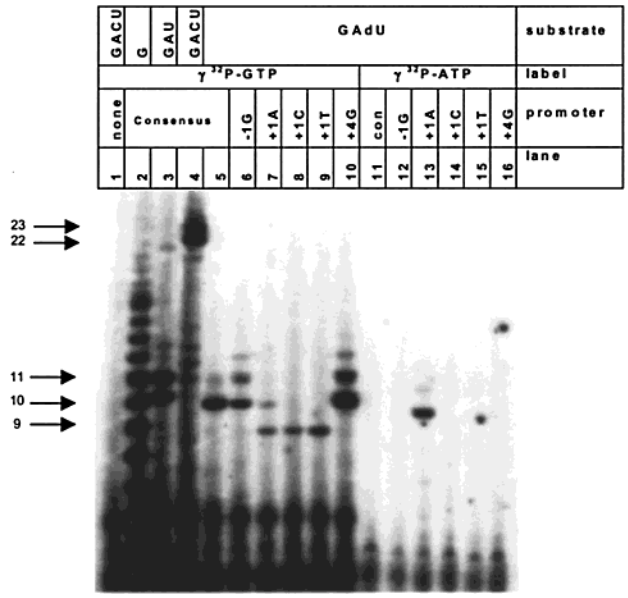

FIGURE 4: Effects of substitutions in the initiation region on choice of start site. Plasmids that contain the indicated promoter (see Table 3) were linearized by digestion with *Hin*cII, which results in the synthesis of runoff products of 21 nt. Transcription reactions contained substrates as indicated (0.4 mM each); the reactions in lanes 5−16 contained 3′-deoxy-UTP (dU) in place of UTP. Products were resolved by electrophoresis in a denaturing 20% polyacrylamide gel; the lengths of the products (in nucleotides) are indicated at the left.

Increased synthesis of slippage products was again observed at the −1G and +4G promoters (Figure 5B, lanes 4 and 9; also see below). We also examined the effects of single bp substitutions in the region from −5 to −2, and in all instances observed initiation exclusively at +1 (*46*).

Effects of Repositioning the Initiation Region. The results above demonstrate that the choice of start site by T7 RNAP is quite stringent. However, with the exception of substitutions at +1, most single bp changes in the initiation region had little effect on start site selection. To explore further the question as to whether sequence-specific information in the initiation region contributes to start site selection, we constructed partially single-stranded (pss) templates in which the consensus initiation sequence (−4 to +6) was positioned either 1 nt closer or 1 nt further away from the upstream binding region by inserting or deleting a base at −4 (Figure 6). It had previously been shown that removal of the nontemplate strand downstream of −5 has little effect on the kinetics of initiation or the choice of a start site (*14*, *44*, *47*); however, the stability of T7 RNAP elongation complexes on pss templates is not as great as on completely double-stranded (ds) templates, and increased premature release of products in the 11−14 nt range is observed on the pss templates (*48*). Nevertheless, the largest, and predominant, product from the consensus promoter in the presence of GTP, ATP, CTP, and dUTP is the expected transcript of 16 nt (Figure 6, lane 3). When the initiation sequence is moved one position closer to the upstream region, initiation occurs at the same distance downstream from the binding region, resulting in the synthesis of products that are 1 nt shorter (lane 4). When the initiation region is moved 1 nt away from the upstream binding region, initiation still occurs the same distance downstream, resulting in products that are 1 nt larger (lane 5). Since the base in the template strand at the new start site in this promoter is a T, the latter transcript is expected to start with an A. This was confirmed by labeling of this transcript with [γ-^{32}P]ATP but not [γ-^{32}P]-GTP (Figure 6, lanes 13 and 9, respectively).

In summary, T7 RNAP prefers to start at +1 if it can initiate with a purine (and preferably with GTP), but when a pyrimidine is encoded at +1 and a purine start is available at +2, it will initiate at the latter position. When given the opportunity to start at −1, even in the context of an optimal initiation sequence, the polymerase continues to initiate 5 nt away from the downstream boundary of the binding region.

Effects of Substitutions on Abortive Initiation. The abundance and distribution of abortive products made by T7 RNAP depends on the sequence of the initially transcribed sequence (ITS) that lies downstream of the start site, and it has been proposed that these differences reflect sequence and position effects on substrate K_m (*27*, *28*, *30*). We have also observed changes in the pattern of abortive products as a result of substitutions in the ITS. The most dramatic effects are observed at promoters that direct the early incorporation of UMP. (In all natural T7 promoters UMP is not incorporated prior to +6.) For example, a promoter that directs the incorporation of UMP at +3 (+3T) results in increased production of both 3 and 4 nt products (*46*). This may be the result of slower substrate incorporation by T7 RNAP due to weak rU·dA interactions and poor orientation of the incoming or accepting base in the active site (*49*−*51*).

An Extended G:C Hybrid Gives Rise to a Longer Ladder of Poly(G) Products. When GTP is provided as the sole substrate for T7 RNAP at a promoter that directs initiation with GGG... a ladder of poly(G) transcripts that extends to 14 nt is observed (Figure 7). It is thought that these products arise as a result of repeated cycles of slippage of the nascent RNA on the template strand and subsequent addition of GMP (*49*). The synthesis of poly(G) products is greatly reduced on promoters that initiate with GGA..., suggesting that a minimum of three rG·dC bps is required to maintain the stability of the complex (*49*, *52*). The synthesis of poly(G) product is enhanced on supercoiled templates vs linear templates (Figure 7, A), which reflects the greater stability of the initiation complex on supercoiled DNA due to a reduced tendency of the duplex DNA to renature and displace the RNA (*53*, *54*). Enhanced synthesis of poly(G) products is also observed on pss templates (*46*).

A more robust poly(G) ladder is also observed at the −1G and +4G promoters (Figure 7A, lanes 4 and 5; Figure 7B, lanes 3 and 4). However, as compared to the poly(G) ladder made at the consensus promoter, the distribution of products is skewed toward a greater length on these templates, particularly for the −1 G promoter (Figure 7C). Presumably this is due to the enhanced retention of the nascent RNA as a result of the opportunity to form a longer G·C hybrid.

During initiation, the affinity of the RNAP for the upstream binding region hinders the free extension of the active site downstream, resulting in competition between slippage of the nascent RNA vs translocation of the active site (*55*). The different behavior of the −1G promoter vs +4G promoter may reflect this competition. At the −1G promoter there is an opportunity to form a longer hybrid (by slippage of the nascent product back to −1) while the active site need extend only to +3 (as at the consensus promoter). At the +4G

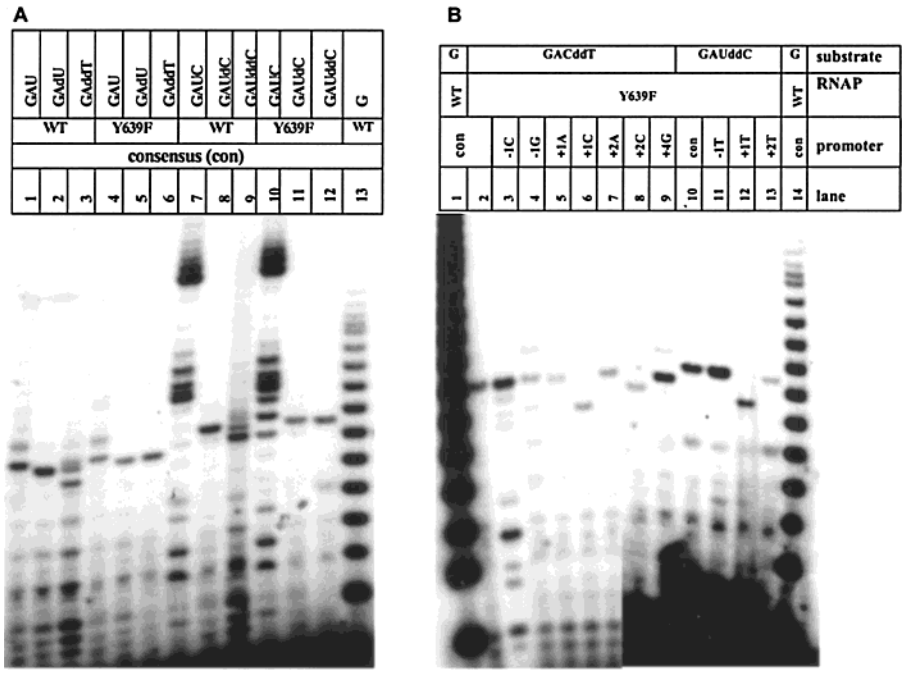

FIGURE 5: Choice of start site in promoter variants. Experiments were carried out as in Figure 4 with either WT T7 RNAP or a mutant enzyme (Y639F) (*35*) that is able to incorporate dNTPs. The chain-terminating analogues 3′-deoxy-UTP (dU), 3′-deoxy CTP (dC), dideoxy-TTP (ddTTP), or dideoxy-CTP (ddCTP) were present as indicated.

promoter (which initiates with the sequence +1 GGGGGA•••), the formation of an even longer G•C hybrid is possible, but the active site must extend to +5.

To examine this further, we utilized synthetic templates that alter the position of the consensus ITS relative to the upstream binding site (Figure 8). A promoter in which the G3 tract is 1 bp closer to the binding region (template 2) gives rise to a weaker poly(G) ladder that extends to ~5 nt. At this promoter, initiation occurs at the central G in the G3 tract (see Figure 6**)** and even though a 3 bp hybrid could be formed by slipping back to −1, this would occur when the active site is at +2, not at +3 as at the consensus promoter. At a promoter in which the G3 tract is moved one position further from the upstream binding region (template 3), initiation occurs with the A just ahead of the G3 tract. Under these conditions, more G-ladder production is observed than with template 2 but less than with the consensus promoter (template 1). Here the opportunity to form a 3 bp hybrid is possible (as at the consensus promoter), but the active site must extend to +4. From these experiments, we conclude that the formation of a poly(G) ladder reflects both the length of the G•C hybrid and the position of the active site relative to the binding region.

DISCUSSION

Studies of T7 RNA polymerase have provided significant insights into RNAP−promoter interactions. These studies have employed both structural and biochemical approaches, as well as the use of promoter variants. A collection of single base pair variants in the binding region from −15 to −6 was particularly useful in probing the basis of specific promoter recognition. The collection of promoter variants reported here (which covers the entire promoter region from −17 to +6) should prove useful in characterizing other aspects of promoter function such as abortive initiation, isomerization, and inhibition by T7 lysozyme.

Promoter Structure. In keeping with prior observations that suggested a two-domain model for promoter function (*11*−*13*), substitutions in the upstream binding region have a stronger effect on promoter strength than substitutions in the initiation region. Surprisingly, however, alterations of certain positions that are conserved in all phage promoters (not just T7 promoters), while weakening the promoter, did not prevent its function (for example, at −14, −6, −4, and especially at −3). Why are these positions so highly conserved? Perhaps these bps are critical for promoter function in vivo but not in vitro, or perhaps the moderate decrease in the level of utilization seen in vitro is not tolerated in the infected cell. In previous work, Chapman and Burgess (*11*) noted dramatic differences in transcription from certain promoter mutants in response to changes in template topology or reaction conditions. Similar effects may be important to promoter function in vivo.

A characteristic feature of all phage promoters is the presence of an AT-rich region that extends from −13 to −17. While A•T vs T•A substitutions are well tolerated in this region, G•C or C•G substitutions are not. This could reflect a direct base readout in the minor groove (*42*, *43*) or could reflect a need for low helix stability or flexibility of the promoter in this region. Two of the amino acid residues in the AT-rich recognition loop are in positions where they could form hydrogen bonds with the nontemplate strand (R96 with N3 of A at −16 and K96 with N3 of A at −13; see ref *22*); however, these bonds do not discriminate among different base pairs and would not account for the observed specificity at these positions. In the crystal structure of the T7 RNAP−promoter complex, the insertion of the AT-rich recognition loop is accompanied by a widening of the minor groove and a slight bend of the DNA, suggesting that the inherent flexibility of this region may be important in recognition (*22*). All class III promoters have an uninterrupted run of A•T bps that extends upstream beyond −17

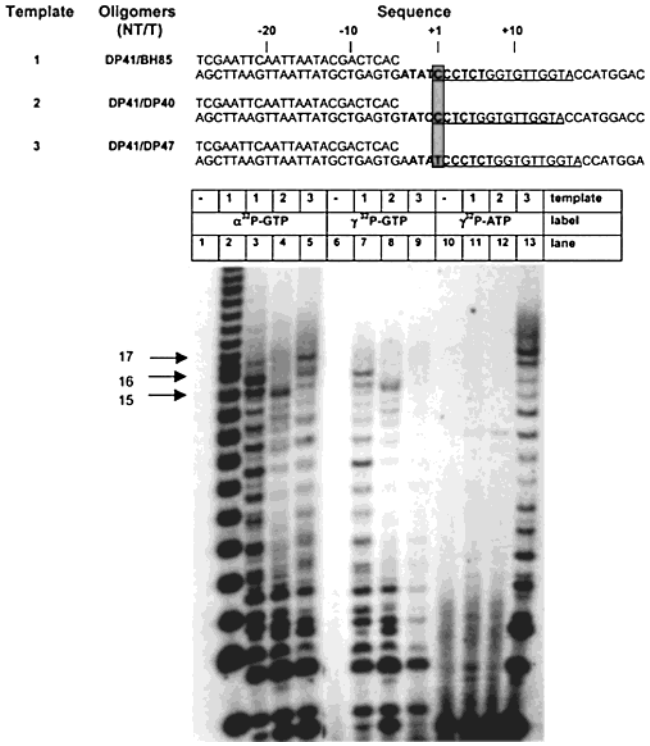

FIGURE 6: Effects of moving the initiation region on choice of start site. Synthetic templates were prepared by annealing together template (T) and nontemplate (NT) strand oligomers as indicated. Template 1 contains the consensus T7 promoter; positions in this template are numbered relative to the start site of transcription (+1), the initiation region from −4 to +6 is shown in bold-face type. In templates 2 and 3 the position of the initiation region relative to the upstream binding region has been shifted by the deletion of the base at −4 (template 2) or insertion of a base between −5 and −4 (template 3). The templates were transcribed by T7 RNAP in the presence of GTP, ATP, CTP, and dUTP with the indicated substrate labeled, and the products were resolved by electrophoresis in a denaturing 20% polyacrylamide gel (− indicates no template was present). Transcription from the consensus promoter is expected to terminate at 16 nt following incorporation of dU (underlined). Shorter products due to premature release are visible in each lane. The start site for transcription in each template (as determined by these experiments) is boxed and shaded.

(up to −22 in some promoters) while nearly all class II promoters have a G·C interruption that limits the AT run to 4−6 bps. Interestingly, experimental class III promoter constructs in which the AT run was truncated upstream of −17, or interrupted in the region from −17 to −13 by a single G·C pair, behave as class II promoters in vivo (*31*).

Promoter Specificity. Previous efforts to characterize RNAP−promoter interactions involved the construction and characterization of RNAP mutants having alterations in the specificity loop. In such a manner, we found that N748 makes specific contacts with the bps at −10 and −11, R756 interacts with the bp at −9, and Q758 interacts with the bp at −8 (*20*, *21*, *39*, *56*). The crystal structure of a T7:RNAP promoter complex confirmed these interactions and revealed an additional contact between the template strand base at −7 and R746 (*22*). While no bp substitutions are tolerated at positions −7, −8, or −9, some changes are accepted at −10 and −11 (Table 1 and Figure 2). The interactions at −7, −8, and −9 each involve two direct hydrogen bonds, and the bidentate nature of these contacts probably accounts for the high specificity at these positions. In contrast, the interaction at −11 involves one direct and one water-mediated hydrogen bond and the interaction at −10 involves only a single water-mediated bond; this may account for the lower stringencies at the latter positions.

As noted in Figure 3, substitution of R756 with other amino acids altered the preference of the mutant RNAP not only for the bp at −9, but also for the bps at −10 and −11. The observation that the change of a single amino acid affected promoter preference over a 3 bp range was initially puzzling but is understandable in light of the crystal structure, where it is observed that, in addition to the two hydrogen bonds R756 makes with the 6-keto and 7-imino groups of G at −9 in the template strand, it also forms a hydrogen bond with N748 via its guanidinium nitrogen. As noted above, the latter residue (N748) is involved in contacts with the bps at −10 and −11. All of the substitutions tested here are expected to change the interaction with N748. While some changes (such as R756K) resulted in altered promoter preference only at −9 and −10, others, such R756C and R756S, resulted in more dramatic changes over the 3 bp interval, suggesting that they have a greater potential to alter the orientation of N748. A similarly broad set of effects on promoter recognition was observed when N748 was replaced with the corresponding residue found in T3 RNAP (T7-N748D; *20*), suggesting that a reorientation of side chain interactions might also be involved in this case (*57*).

Choice of Start Site. The question as to how T7 RNAP chooses the correct site for initiation has been the subject of a number of recent investigations. Two (nonexclusive) models for start site selection have been discussed. In the first, the polymerase initiates at a position on the template strand that is a preferred distance from the upstream binding region, while in the second model, specificity is dependent upon information in the template strand in the initiation region (*44*). In the crystal structure of the T7 RNAP−promoter complex, the template strand is led down into the active site by contacts along the surface of the RNAP (*22*). It is clear from the studies reported here that, with the exception of substitutions at +1, changes in the initiation region are well tolerated and have little effect on start site selection. These observations are consistent with the notion that there is little sequence-specific information in the template strand in the initiation region that is required for start site selection and that initiation occurs at a particular distance downstream from the binding region. The effects of deleting or inserting base pairs between the binding region and the initiation region are also consistent with the polymerase counting off the distance from the binding region to the start site. For example, repositioning the entire initiation region (−4 to +6) either 1 nt closer or 1 nt further away from the upstream binding site still resulted in initiation the same distance downstream (Figure 6).

In earlier work, Weston et al. found that the insertion of flexible nonnucleosidic linker segments between the binding region and the initiation region had little effect on the choice of start site (*44*). While this observation would seemingly argue against a model for start site selection that involves counting off a fixed distance from the upstream binding region, the linkers used in these studies did not allow initiation within the inserted region (as there were no bases to direct incorporation). Instead, the work by Weston et al., together with the results presented here, are consistent with a model in which the RNAP initiates a minimum distance

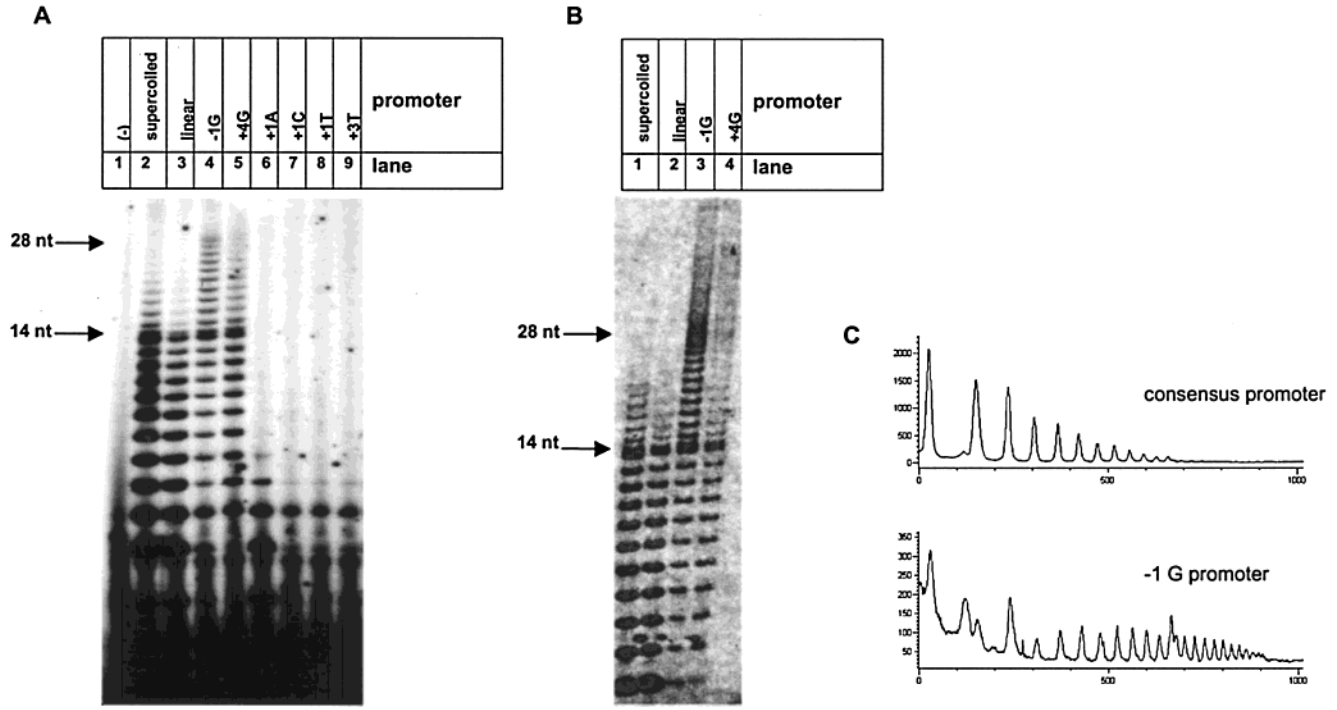

FIGURE 7: Effects of base-pair substitutions in the initiation region on poly(G) synthesis. (A). Plasmid templates carrying the indicated promoter (see Table 3) were digested with *Hin*cII. Reactions (10 μL) contained 500 ng of template, 0.5 mM GTP, 20 ng of WT RNAP, and [γ^{32}-P]GTP and were incubated at 37 °C for 15 min. The products were resolved by electrophoresis in a highly cross-linked 20% polyacrylamide gel (acrylamide:bisacrylamide 12:1). The length (in nucleotides) of each transcript is indicated at the side. (B) Same as panel A except that the labeled substrate was [α-^{32}P]GTP and the gel was less cross-linked (acrylamide:bisacrylamide) (19:1) to improve the resolution of longer products. (C) Densitometric scans of lanes 3 and 4 of panel A. Since the transcripts were labeled at their 5′ end with [γ^{32}-P]GTP, the peak heights provide an indication of the relative molar production of each species.

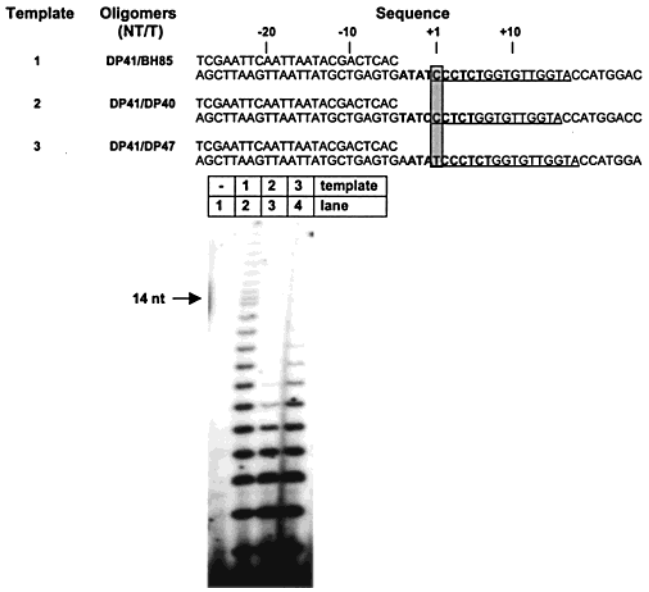

FIGURE 8: Effects of repositioning the initially transcribed sequence on poly(G) synthesis. Synthetic templates were prepared as described in Figure 6. The conserved initiation region is shown in boldface type, and the start site of transcription at each promoter (as determined in Figure 6) is shaded and boxed. Templates were transcribed by T7 RNAP in the presence of [γ-^{32}P]GTP as the sole substrate, and the products were resolved by electrophoresis in a 20% polyacrylamide gel.

(i.e., not less than 5 nt) downstream from the binding region. Superimposed upon this restraint is the absolute preference for initiation with purines (and especially GTP). Thus, while we never detected initiation closer than this interval (even when the consensus initiation region was shifted 1 nt closer to the binding region), we detected initiation at +2 when a pyrimidine start was required at +1. Recent work by Huang et al. (*58*) indicates that the strong preference for initiation with GTP may involve residue H784.

Weston et al. (*44*) found that removal of the base at −1 or substitution with an abasic linker at this position resulted in changes in start site selection, suggesting that stacking interactions between the −1 and +1 bases might contribute to start site specificity. However, we have found that initiation occurs accurately at +1 regardless of which base is present at −1. In the crystal structure of a T7 RNAP−promoter complex, Cheetham et al. (*22*) observed stacking interactions between the template strand base at −1 and the aromatic side chain of Trp422, and suggested that these interactions might be important in positioning the base at +1 in the active site. While this finding could account for the effects of an abasic linker at −1 on start site selection observed by Weston et al., our results indicate that this stacking interaction is not sequence-specific, as base substitutions at this position do not affect start site selection.

Poly(G) Synthesis and the Length of the RNA:DNA Hybrid. The synthesis of poly(G) products by T7 RNAP is thought to result from repeated cycles of slippage of the nascent RNA on the template strand and subsequent addition of GMP (*49*). We have found that this process is enhanced on promoter variants in which there is an opportunity to form a more extended G·C hybrid (such as the −1G or +4G promoters). Presumably this is due to a decreased tendency of the RNA to dissociate before it can again be extended.

In solving the crystal structure of a T7 RNAP initiation complex at a consensus promoter, Cheetham and Steitz (*59*) utilized the nonhydrolyzable ATP analogue α,β methylene-

ATP in an apparent attempt to limit RNA synthesis to the trinucleotide pppGpGpG. They noted that the base at the 5′ terminus of the trinucleotide was involved in non-Watson−Crick base pairing, suggesting that it might be peeling off the template strand at this point, and further observed that extension of the RNA·DNA hybrid beyond 3 bp would result in steric clashes with the N-terminal domain of the RNAP. On the basis of these findings, they suggested that the RNA·DNA hybrid may not extend further than 3 bp upstream from the active site in the IC. However, we and others have found that under standard conditions of transcription α,β-methylene-ATP does not inhibit poly(G) synthesis, even at very high ratios of the analogue to GTP (W.T.M., unpublished observations; Craig Martin, personal communication); furthermore, the analogue is not visible in the structure presented (*59*). The likely heterogeneity of the RNA in the crystal complicates the interpretation of the diffraction data. In addition, the conclusion that the RNA·DNA hybrid cannot extend further than 3 bp upstream from the active site (*59*) is not consistent with our observation that the opportunity to form a more extended hybrid results in enhanced synthesis of poly(G) products. Additional studies involving complexes in which the transcript is longer than 3 nt will be needed to resolve this issue.

ACKNOWLEDGMENT

This work was supported by NIH grant GM381477 and was submitted to the State University of New York in partial fulfillment of the requirements for the doctoral degrees of D.I and M.R. We thank Manli Jiang and Ray Castagna for assistance in carrying out these experiments.

REFERENCES

1. Dunn, J. J., and Studier, F. W. (1983) *J. Mol. Biol. 166*, 477−535.
2. McAllister, W. T., and Wu, H.-L. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 804−808.
3. Ikeda, R. A., and Bailey, P. A. (1992) *J. Biol. Chem. 267*, 20153−20158.
4. Zhang, X., and Studier, F. W. (1997) *J. Mol. Biol. 269*, 10−27.
5. Moffatt, B. A., and Studier, F. W. (1987) *Cell 49*, 221−227.
6. Kumar, A., and Patel, S. S. (1998) *Biochemistry 36*, 13954−13962.
7. Villemain, J., and Sousa, R. (1998) *J. Mol. Biol. 281*, 793−802.
8. Klement, J. F., Moorefield, M. B., Jorgensen, E. D., Brown, J. E., Risman, S., and McAllister, W. T. (1990) *J. Mol. Biol. 215*, 21−29.
9. Bailey, J. N., Klement, J. F., and McAllister, W. T. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 2814−2818.
10. Lee, S. S., and Kang, C. (1993) *J. Biol. Chem. 268*, 19299−19304.
11. Chapman, K. A., and Burgess, R. R. (1987) *Nucleic Acids Res. 15*, 5413−5432.
12. Chapman, K. A., Gunderson, S. I., Anello, M., Wells, R., and Burgess, R. R. (1988) *Nucleic Acids Res. 16*, 4511−4524.
13. Li, T., Ho, H. H., Maslak, M., Schick, C., and Martin, C. T. (1996) *Biochemistry 35*, 3722−3727.
14. Ujvari, A., and Martin, C. T. (1997) *J. Mol. Biol. 273*, 775−781.
15. Muller, D. K., Martin, C. T., and Coleman, J. E. (1989) *Biochemistry 28*, 3306−3313.
16. Jorgensen, E. D., Durbin, R. K., Risman, S. S., and McAllister, W. T. (1991) *J. Biol. Chem. 266*, 645−651.
17. Schick, C., and Martin, C. T. (1995) *Biochemistry 34*, 666−672.
18. Maslak, M., Jaworski, M. D., and Martin, C. T. (1993) *Biochemistry 32*, 4270−4274.
19. Diaz, G. A., Raskin, C. A., and McAllister, W. T. (1993) *J. Mol. Biol. 229*, 805−811.
20. Raskin, C. A., Diaz, G. A., Joho, K., and McAllister, W. T. (1992) *J. Mol. Biol. 228*, 506−515.
21. Rong, M., He, B., McAllister, W. T., and Durbin, R. K. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 515−519.
22. Cheetham, G., Jeruzalmi, D., and Steitz, T. A. (1999) *Nature 399*, 80−83.
23. McAllister, W. T. (1997) Transcription by T7 RNA polymerase, in *Nucleic Acids and Molecular Biology* (Eckstein, F., Lilley, D. M. J., Eds.) pp 15−25, Springer, Berlin.
24. Ikeda, R. A., and Richardson, C. C. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 3614−3618.
25. Gunderson, S. I., Chapman, K. A., and Burgess, R. R. (1987) *Biochemistry 26*, 1539−1546.
26. Sousa, R., Patra, D., and Lafer, E. M. (1992) *J. Mol. Biol. 224*, 319−334.
27. Ikeda, R. A. (1992) *J. Biol. Chem. 267*, 11322−11328.
28. Milligan, J. F., Groebe, D. R., Witherell, G. W., and Uhlenbeck, O. C. (1987) *Nucleic Acids Res. 15*, 8783−8798.
29. Lopez, P., Guillerez, J., Sousa, R., and Dreyfus, M. (1997) *J. Mol. Biol. 269*, 49−51.
30. Guajardo, R., Gopal, V., Lopez, P., and Sousa, R. (1998) *J. Mol. Biol. 281*, 777−792.
31. Jolliffe, L. K., Carter, A. D., and McAllister, W. T. (1982) *Nature 299*, 653−656.
32. Schneider, T. D., and Stormo, G. D. (1989) *Nucleic Acids Res. 17*, 659−674.
33. Ikeda, R. A., Warshamana, G. S., and Chang, L. L. (1992) *Biochemistry 31*, 9073−9080.
34. Ikeda, R. A., Ligman, C. M., and Warshamana, S. (1992) *Nucleic Acids Res. 20*, 2517−2524.
35. Sousa, R., and Padilla, R. (1995) *EMBO J. 14*, 4609−4621.
36. He, B., Rong, M., Lyakhov, D. L., Gartenstein, H., Diaz, G. A., Castagna, R. C., McAllister, W. T., and Durbin, R. K. (1997) *Protein Expression Purif. 9*, 142−151.
37. Maslak, M., and Martin, C. T. (1994) *Biochemistry 33*, 6918−6924.
38. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A laboratory manual (2nd ed.)*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
39. Rong, M. (1998) Characterization of promoter specificity determinants and template requirements of T7 RNA polymerase, Ph.D. Thesis, State University of New York Health Science Center at Brooklyn.
40. Place, C., Oddos, J., Buc, H., McAllister, W. T., and Buckle, M. (1999) *Biochemistry 38*, 4948−4957.
41. Rastinejad, F., and Lu, P. (1993) *J. Mol. Biol. 232*, 105−122.
42. Seeman, N., Rosenberg, J., and Rich, A. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 804−808.
43. Steitz, T. A. (1990) *Q. Rev. Biophys. 23*, 205−280.
44. Weston, B. F., Kuzmine, I., and Martin, C. T. (1997) *J. Mol. Biol. 272*, 21−30.
45. Pleiss, J. A., Derrick, M. L., and Uhlenbeck, O. C. (1998) *RNA 4*, 1313−1317.
46. Parrotta-Imburgio, D. (2000) Effects of changes in promoter sequence on initiation by T7 RNA polymerase, and the role of a conserved structural motif (DX_2GR) in polymerase function, Ph.D. Thesis, State University of New York Health Science Center at Brooklyn.
47. Maslak, M., and Martin, C. T. (1993) *Biochemistry 32*, 4281−4285.
48. Mentesana, P. E., Chin-Bow, S. C., Sousa, R., and McAllister, W. T. (2000) *J. Mol. Biol.*, in press.
49. Martin, C. T., Muller, D. K., and Coleman, J. E. (1988) *Biochemistry 27*, 3966−3974.
50. Ling, M. L., Risman, S. S., Klement, J. F., McGraw, N. J., and McAllister, W. T. (1989) *Nucleic Acids Res. 17*, 1605−1618.
51. Rong, M., Durbin, R. K., and McAllister, W. T. (1998) *J. Biol. Chem. 273*, 10253−10260.

52. Gross, L. B., Chen, W. J., and McAllister, W. T. (1992) *J. Mol. Biol. 228*, 488−505.
53. Diaz, G. A., Rong, M., McAllister, W. T., and Durbin, R. K. (1996) *Biochemistry 35*, 10837−10843.
54. Villemain, J., Guajardo, R., and Sousa, R. (1997) *J. Mol. Biol. 273*, 958−977.
55. Guajardo, R., and Sousa, R. (1997) *J. Mol. Biol. 265*, 8−19.
56. Raskin, C. A., Diaz, G. A., and McAllister, W. T. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 3147−3151.
57. Raskin, C. A. (1992) Characterizing a promoter specificity determinant of bacteriophage T7 and T3 RNA polymerases, Ph.D. Thesis, State University of New York Health Science Center at Brooklyn.
58. Huang, J., Brieba, L. G., and Sousa, R. (2000) *Biochemistry*, in press.
59. Cheetham, G., and Steitz, T. A. (1999) *Science 286*, 2305−2309.
60. He, B., Rong, M., Durbin, R. K., and McAllister, W. T. (1997) *J. Mol. Biol. 265*, 275−288.
61. McAllister, W. T., and Raskin, C. A. (1993) *Mol. Microbiol. 10*, 1−6.
62. Ujvari, A., and Martin, C. T. (1996) *Biochemistry 35*, 14574−14582.

BI000365W