J. Mol. Biol. (1993) 231, 569–580

Transcriptional Slippage During the Transcription Initiation Process at a Mutant *lac* Promoter *in Vivo*

Xiao Feng Xiong† and William S. Reznikoff‡

Department of Biochemistry, College of Agricultural and Life Science
University of Wisconsin-Madison
Madison, WI 53706, U.S.A.

(Received 22 September 1992; accepted 28 January 1993)

A C·G to A·T transversion at position +10 of the *lac* promoter activates a nascent sigma 70-dependent promoter (the +10A promoter). The *lac* +10A promoter has two unusual properties; it programs a large family of transcripts with multiple 5′ ends, and its sequence bears little resemblance to other sigma 70-dependent promoters. The 5′ end of the +10A *in vivo* mRNA was determined to contain oligo(U) sequences of varying lengths suggesting that the true start site was at a run of three T·A base-pairs located 20 to 22 bp downstream of the *lac* wild-type promoter start site, and that the transcription initiation process involved a transcriptional slippage event (which resulted in multiple rU incorporation). Only mutations at or near the start site and those deletions that changed the location of the start site abolished this transcriptional slippage property of the transcription initiation process. This transcriptional slippage was also found to be promoter independent because changing the *lac* UV5 start site to a run of five T·A base-pairs (−1 to +4) resulted in similar transcriptional slippage. Saturated mutagenesis of the +10A promoter identified a potential −10-like region and indicated that sequences immediately upstream of the −10 region contributed to the promoter's activity. Decreasing the weak −35 region homology did not change promoter strength; however, introduction of the consensus −35 hexamer TTGACA increased expression tenfold. RNA polymerase bound to the +10A promoter partially protects a 20 base-pair sequence from DNase I digestion upstream of the start site. These results suggest that RNA polymerase interacts with the +10A promoter in a different manner from that for the majority of sigma 70 promoters.

Keywords: reiterative; transcription; initiation; polymerase; promoter

1. Introduction

The transcription initiation process is a key step in gene expression. It has been the object of extensive analysis for over 30 years. Three important lines of inquiry into this process include: (1) the determination of the DNA sequence determinants that are recognized by RNA polymerase (what sequences make up a promoter?); (2) the analysis of biochemical processes occurring during the formation of the first few phosphodiester bonds (how do RNA polymerase–promoter complexes initiate polymerization?); and (3) the examination of how promoter structure influences the nature of the polymerization initiation process. The mutant *lac* promoter, +10A, has two unusual properties whose elucidation may contribute to these lines of inquiry. The +10A promoter programs the synthesis of mRNAs with multiple apparent 5′ ends and the +10A promoter is not composed of an obvious promoter sequence.

The conversion of RNA polymerase from an open complex into a transcribing core complex is a multistep process that ends with the release of the sigma subunit and the release of the favorable sequence-specific RNA polymerase–DNA contacts found in the open complex. During this process many RNA polymerase–promoter complexes will, *in vitro*, undergo an abortive initiation process; the synthesis and release of short oligonucleotides that have been properly initiated but not productively elongated into full-length RNA chains (Carpousis & Gralla, 1980; Munson & Reznikoff, 1981). It is not known whether or not this process occurs *in vivo*. Another feature of the RNA polymerase initiation process that may be linked to RNA chain initiation and promoter clearance, is the transcriptional slippage that has been observed in a few cases *in vitro*.

† Current address: New England Biolabs Inc., 32 Tozer Road, Beverly, MA 01915, U.S.A.
‡ Author for correspondence.

0022-2836/93/110569-12 $08.00/0

At some promoters in which the start site is part of a stretch of three to five A·T or T·A base-pairs, the polymerase incorporates more copies of the corresponding nucleotide than is programmed by the template sequences (Machida *et al.*, 1984; Harley *et al.*, 1990; Guo & Roberts, 1990; Jacques & Susskind, 1990). Guo & Roberts (1990) proposed that a slippage model could account for the *in vitro* synthesis of such unusual RNAs programmed by the late gene promoter of phage 82. In this case, RNA polymerase starts transcription at the first of the three consecutive dT bases in the template strand (+1). After the first three rA·dT base-pairs are formed in the transcription bubble, the initiated transcript can slip one base-pair toward the 5′ end, so that an additional AMP can be templated by the third of three consecutive dT bases in the template strand. The initiation complexes can go through numerous slippage cycles before they switch into an elongation complex, resulting in oligo(A) chains of varying lengths at the 5′ end of transcripts.

In previous studies, it was found that single point mutations in the consecutive base-pairs at start sites will abolish the transcription slippage (Guo & Roberts, 1990; Jacques & Susskind, 1990). However, it is not known if consecutive base-pairs of any of the four nucleotides will give rise to transcriptional slippage during the initiation process and it is not known whether or not other aspects of promoter structure are required for this process to occur.

The +10A mutation of the *lac* promoter was isolated both as a pseudorevertant of the −11G *lac* mutation and, in another study, as a CAP-cAMP independent *lac*P mutation (Maquat & Reznikoff, 1978; Karls *et al.*, 1989; Rothmel & LeClerk, 1989). It has a unique property of programming the synthesis of a family of transcripts with multiple apparent start points from around +20 to +1 or further upstream relative to the wild-type *lac* promoter start site both *in vivo* and *in vitro* (Karls *et al.*, 1989). Here, we demonstrate that the +10A promoter programs transcriptional slippage during the initiation process from around +20. This is the first demonstration that transcriptional slippage occurs during the transcription initiation process *in vivo* in *Escherichia coli*. The features of promoter structure that give rise to transcriptional slippage were studied by examining the effects of over 50 single point mutations of the *lac* +10A promoter and start site mutations of the *lac* UV5 promoter.

The DNA sequences of several hundred sigma 70-dependent *E. coli* promoters have been determined. In general most promoters resemble the so-called consensus sequence with a −35 region hexamer (TTGACA) displaced by approximately 17 bp from a −10 region hexamer (TATAAT). Moreover the level of expression is, in most cases, related to the promoter's similarity to this consensus sequence (the similarity is quantified by the promoter's "homology score", Mulligan *et al.*, 1984). Some exceptions to this promoter sequence–activity relationship have been found. For instance, the *gal* P1 promoter has no apparent −35 region but instead there are a few base-pairs just upstream of the −10 region that appear to be crucial for its functioning (Busby & Chan, 1989). It is of interest to examine other promoters that do not follow the consensus relationship in part because they may reveal additional or alternative sequence patterns that can be recognized by RNA polymerase.

The +10A promoter has been shown to be sigma 70-dependent (Karls *et al.*, 1989). However, it has an extremely poor homology score (37·3), which is equivalent to the bottom seven out of 112 promoters examined by Mulligan *et al.* (1984). The +10A promoter's level of activity is substantially greater than that predicted by its homology score. An analysis of +10A promoter point mutations revealed that, similar to *gal*P1, the +10A promoter has a non-functional −35 region and sequences immediately upstream from the −10 region play an important role in its activity.

2. Materials and Methods

(a) *Enzymes and reagents*

AMV reverse transcriptase was purchased from Molecular Genetic Resources, Inc. Phage T4 DNA polymerase, phage T4 DNA ligase and the 1212 primer (located at +95 of the *lac* operon, dGTTTTCCCAGTCACGAC) were purchased from New England Biolabs. Phage T4 nucleotide kinase, terminal transferase and 5× terminal transferase buffer were purchased from Promega. *E. coli* RNA polymerase was a gift from Dr Richard Burgess. The degenerate oligonucleotides that were used for mutagenesis were synthesized by the University of Wisconsin Biotechnology Center. The −105 primer (dCTGGCACGACAGGT) was synthesized in the Department of Biochemistry of the University of Wisconsin-Madison. Diethyl pyrocarbonate, ampicillin and *o*-nitrophenyl-β-D-galactopyranoside were purchased from Sigma. Actinomycin D was purchased from Boehringer-Mannheim Biochemicals.

(γ-^{32}P)-labeled adenosine triphosphate (>5000 Ci/mmol) and (α-^{32}P)-labeled deoxy adenosine triphosphate (>3000 Ci/mmol) were purchased from the Amersham Corporation.

(b) *Bacterial strains and plasmids*

DH5α [F⁻ *endA1*, *hsdR17*, *supE44*, *thi*, *recA1*, *gyrA96*, *relA1*, Δ(*lac*)*U169*, *φ80dlacZM15*, *endA1*] was used as the host for the screening of mutants. CSH26 [F⁻ Δ(*lac-proAB*)*thi*] and CSH26 Δ*cya* [F⁻ Δ(*lac-proAB*)*thi* Δ*cya*] were used for β-galactosidase and primer extension assays. *λplac5-T743* and its derivatives were used to make single copy λ lysogens.

Starting plasmids pRZ3207 and pRZ6524 were constructed in this laboratory by replacing the 322 bp *Pvu*II fragment of pUC119 with a *lac Pvu*II fragment (268 bp) that contains a wild-type *lac* sequence (pRZ3207) or an A to G mutation at −11 (pRZ6524).

pRZ6522 is identical with pRZ3207, except that it contains the *lac* UV5 promoter rather than the wild-type *lac* promoter in pRZ3207. pRZ6505 is a pRZ6524-derived plasmid containing −11G, +10A as described by Karls *et al.* (1989).

(c) *Oligonucleotide-directed mutagenesis*

The oligonucleotide-directed mutagenesis followed the general protocols described by Sambrook *et al.* (1989). We obtained a mixed population of oligonucleotides spanning the region from −9 to +34 of the *lac* promoter, each of which contains +10A, plus on the average one other change between −3 and +28. The degenerate oligonucleotides were hybridized to single-stranded DNA derived from pRZ6524, and the double-stranded plasmids were synthesized with T4 DNA polymerase. Mutants were screened amongst DH5α transformants. All mutants obtained in this way were confirmed to contain −11G, +10A and one other change between −3 and +26 by DNA sequencing. The −11G mutation was incorporated in these mutants to eliminate the transcript from wild-type *lac* P1 promoter (Karls *et al.*, 1989), therefore these mutants were assayed in strain CSH26. Some mutants were originally obtained by Russell Karls from the pRZ3207 vector. Those mutants, which include +1T, +1G, +1C, +2T, +3A, +3G, +3C, +6C, +7T, +11T, +11C and +12A, were assayed in strain CSH26 Δ*cya* to minimize the effect of *lac* P1 transcripts.

Start site mutations of the *lac* UV5 promoter (see Fig. 4(ε)), UV5+1As; UV5+1Ts; UV5+1Gs and UV5+1Cs, were obtained through site-directed mutagenesis of pRZ6522.

Start site mutations (+20As, +20Gs and +20Cs, see Fig. 5(a)) and +35 hexamer mutations (−35up (TTGACA) and −35dn (ACAGGC)) of the +10A promoter were obtained through site-directed mutagenesis of plasmid pRZ6505.

(d) *β-Galactosidase assays*

The mutations were crossed onto *λplac5-T743*, which contains *lacZ* but is phenotypically Lac⁻ because it has the −11G mutation in the *lac* regulatory region (Karls *et al.*, 1989). CSH26 or CSH26 Δ*cya* were lysogenized with the recombinant phages as described by Karls *et al.* (1989). β-Galactosidase assays were performed by the methods of Miller (1972) from lysogens grown in M9 plus glucose. Values for β-galactosidase activities given in Fig. 7 are an average of 4 independent isolates. Standard deviations are less than 10% of sample averages.

(e) *Primer extension*

The preparation of *in vivo* mRNA and subsequent primer extension analysis were performed as described by Xiong *et al.* (1991).

(f) *RNA sequencing*

The RNA sequencing reactions were performed according to procedures described by Air (1979) and DeBorde *et al.* (1986). Approximately 10 μg of RNA from pRZ6505 prepared as described (Xiong *et al.*, 1991) was dried in a speed vacuum and redissolved in 10 μl of hybridization buffer (60 mM NaCl, 50 mM Tris·HCl (pH 8·3), 10 mM DTT) containing 5 ng of (γ-^{32}P)-labeled 1212 primer. The hybridization step was done at 37 °C for 15 min. In tube A, 2·5 μl of RNA-primer mix was mixed with 0·125 μl of AMV reverse transcriptase (16 units/μl), 0·25 μl of actinomycin D (2·5 mg/ml), 0·5 μl of $MgCl_2$ hybridization buffer (36 mM $MgCl_2$, 60 mM NaCl, 50 mM Tris·HCl (pH 8·3), 10 mM DTT) and 3 μl of RNA sequencing mixture A (36 mM $MgCl_2$, 60 mM NaCl, 50 mM Tris·HCl (pH 8·3), 10 mM DTT, 375 μM dATP, 375 μM dTTP, 375 μM dGTP, 375 μM dCTP and 375 μM ddATP); tube T was the same as tube A except it contained 375 μM ddTTP and no ddATP. Tube G and tube C contain ddGTP and ddCTP correspondingly. The completed mixes were incubated at 45 °C for 30 min. Stop solution (6 μl of 95% (v/v) formamide, 20 mM EDTA, 0·05% (w/v) bromophenol blue and 0·05% (v/v) xylene cyanol FF) was added. The contents of tubes A, T, G and C were loaded onto a 4% (w/v) polyacrylamide gel and electrophoresed for 40 min at 20 W to eliminate the excess primer. The gel slices containing the cDNAs identified by the ^{32}P image were cut out. The cDNAs in the gel slices were eluted overnight in 0·5 ml of 0·5 M ammonium acetate and 1 mM EDTA. After precipitation of cDNAs with ethanol, the cDNAs were resuspended in 10 μl of water, boiled for 2 min and quick-chilled on ice. To resolve the 5′-terminal mRNA sequence indicator (termination product due to incorporation of ddNTP at the correct site) from the termination product of cDNA due to the reverse transcriptase merely reaching the 5′ end of a subset of the mRNAs, 10 μl of terminal transferase mix (16 μl of 5× terminal transferase buffer, 16 μl of water and 8 μl of terminal transferase, 20 units/μl) was then added. The samples were incubated at 37 °C for 1 h. This allowed terminal transferase extension of the 3′ OH ends that were products of the reverse transcriptase termination at mRNA 5′ ends. The cDNAs were then precipitated with ethanol, resuspended in 95% formamide and loaded on an 8% polyacrylamide gel with 8 M urea.

(g) *DNase I footprint*

pRZ6505 plasmid DNA, (γ-^{32}P)-labeled 1212 primer and −105 primer were used for PCR amplification of a 218 bp fragment which served as the substrate for the DNase I footprint experiment. A control experiment was also performed on the *lac* UV5 promoter where the template DNA was obtained in the same way except using pRZ6522 plasmid DNA for PCR† amplification.

About 160 ng of a labeled DNA fragment in 20 μl of DNase I footprint buffer (10 mM Tris·HCl (pH 7·9), 30 mM potassium glutamate, 10 mM $MgCl_2$, 10 mM $CaCl_2$, 0·1 mM DTT, 0·2 mM UTP (or not UTP for the UV5 promoter) and 25% glycerol) was incubated for 5 min at 37 °C. Then 8 μl of 0·6 μg *E. coli* RNA polymerase/ml was added. After 15 min further incubation at 37 °C, 1 μl of 1·5 mg heparin/ml was added, followed by 1 μl of 8 mM rifampicin. At 5 min later, 2 μl of 0·4 μg DNase I/ml was added. The DNase I digestion was stopped after 1 min by adding 5 μl of stop solution containing 2·6 M ammonium acetate, 87 mM EDTA and 130 μg calf thymus DNA/ml. The reaction mixture was mixed with 5 μl of 60% (w/v) sucrose and then loaded on a 4% polyacrylamide gel (acrylamide to bis-acrylamide, 39:1, w/w). The gel slices containing DNA-RNA polymerase complex and free DNA were cut out. The DNA was eluted from the gel slice into 0·5 M ammonium acetate and 1 mM EDTA, precipitated with ethanol and loaded onto an 8% polyacrylamide gel.

(h) *Nomenclature*

Here, the nucleotide pairs are numbered according to their positions in the wild-type and UV5 *lac* promoter sequence. Thus the +10A mutation is 10 bp downstream from the wild-type *lac*P start site and the +10A promoter

† Abbreviation used: PCR, polymerase chain reaction.

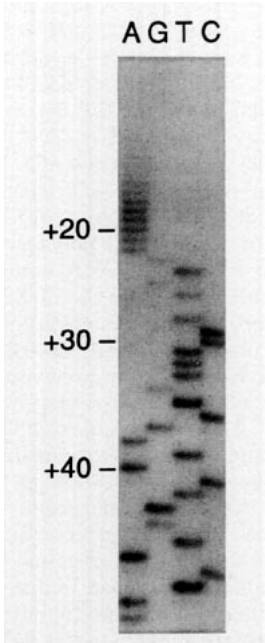

Figure 1. Sequence of +10A mRNA. Primer 1212 was extended using AMV reverse transcriptase in the chain-terminator procedure followed by terminal transferase chasing as described in Materials and Methods.

start site is at or near +20, 20 bp downstream from the wild-type *lac*P start site.

3. Results

(a) *The 5′ end of the +10A mRNA consists of oligo(U) tails*

To identify the nature of the apparent multiple start sites for the +10A mutant of the *lac* promoter, we directly sequenced the 5′ ends of +10A *in vivo* mRNAs using AMV reverse transcriptase. Since the 5′ ends of the +10A mRNAs differ in length, sequencing this family of mRNA molecules led to ambiguities beginning just 5′ to the shortest mRNA present (data not shown). This is because reverse transcriptase will terminate cDNA synthesis corresponding to the 5′ nucleotide of the mRNA even though it incorporates a deoxynucleotide, giving an apparent termination in all four ddNTP reactions, as well as a result of incorporating dideoxy nucleotides giving the sequence-specific stops. Therefore, terminal transferase extension was used to resolve the true dideoxy-terminated primer extension products by extending the length of products with 3′ OH groups (cDNAs that have terminated at the various mRNA 5′ ends). As shown in Figure 1, the 3′ end of +10A cDNA consists of oligo(A), which indicates that the 5′-terminal nucleotides of the +10A mRNA consists of oligo(U) of different lengths. The cDNA sequence from +20 and downstream is identical with the DNA sequence of the template strand. These results suggest that RNA polymerase is reiteratively transcribing the three consecutive A bases of the template strand (+20 to +22), which generates oligo(U) chains of varying lengths at the 5′ terminus of the transcripts. We term this process "transcriptional slippage".

(b) *Effect of +10A promoter point mutations on transcriptional slippage*

To study what sequences in the +10A promoter are responsible for the transcriptional slippage nature of the transcription initiation process, a large collection of point mutations of the +10A promoter was obtained through oligonucleotide directed mutagenesis (shown in Fig. 2). This collection also includes four single base-pair deletion mutations at −1, +15, +18 and +20. Primer extension analysis of all *in vivo* mutant mRNAs showed that most mutants gave the same transcriptional slippage pattern as the parent +10A promoter (data not shown). However, +19C, +20A, +21C, +22C, Δ+15, Δ+18 and Δ+20 mutations appeared to abolish the transcriptional slippage pattern (Fig. 3). In addition, six of these mutations changed the transcription start site sequence. Of these six mutations, four (+20A, +21C, +22C and Δ+20T) directly changed the start site sequence, two single base-pair deletion mutants (Δ+15A and Δ+18A) effectively changed the start site sequence by shifting the start site downstream. One mutation, +19C, changed the base-pair immediately upstream of the start site consecutive T·A base-pairs.

One mutation created an additional start. The +1T mutation showed a combination pattern of transcriptional slippage and a strong start at +13 (data not shown). The +13 start is due to the generation of another promoter (P′115 or P+1T), which has been reported by Maquat & Reznikoff (1980), Karls *et al.* (1989) and Rothmel & LeClerk (1989).

```
          -30        -20        -10         +1         +10        +20        +30
5'  CTTTATGCTT CCGGCTCGTG TGTTGTGTGG AATTGTGAG AGGATAACAA TTTCACACAG GAAACA  3'
                                    ↓ ↓ ↓↓↓↓↓↓↓↓ ↓↓↓↓↓↓↓↓↓ ↓↓↓↓↓↓↓
                -35up  TTGAC A      C Δ TTAAAAATA  AATGTTATC ACCATAC
                -35dn  ACAGG C        GCGGTCTGC  T GCGGTC  Δ TGT
                                      C C C  C   C C CCGΔ     C
                                                      Δ
```

Figure 2. DNA sequences of +10A promoter and promoter mutations. The DNA sequence is numbered relative to the transcription start site of the wild-type *lac* promoter (+1). Transcription of +10A promoter initiated from +20. The locations of the −10 and −35 hexamers of the +10A promoter are underlined. Mutations used in this study are shown below the +10A promoter sequence. (Δ) The locations of single base-pair deletion mutations.

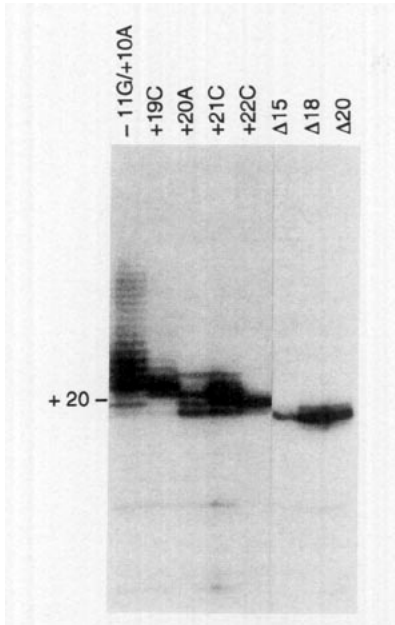

Figure 3. Primer extension analysis of +10A promoter mutations. The *in vivo* transcripts of +10A promoter mutations were examined by primer extension analysis in strain CSH26 as described in Materials and Methods. The −11G/+10A lane shows a typical pattern for +10A promoter transcription products. Lanes +19C, +20A, +21C and +22C show analyses of transcripts of start site mutations. Lanes Δ15, Δ18 and Δ20 show the transcript analyses of single base-pair deletion mutants at positions +15, +18 and +20, respectively.

(c) *Effect of a homonucleotide stretch at the transcription initiation site*

On the basis of the above mutant analysis, it is clear that the transcription start site sequence is crucial for transcriptional slippage. The transcription start site of the +10A promoter contains a run of three T·A base-pairs. Since homonucleotide stretches at start sites have been observed in all previously studied cases of transcription initiation involving transcriptional slippage, we wanted to test if a homonucleotide stretch at a start site is sufficient to cause this process. In order to accomplish this, the transcription initiation site (−1 to +3) of the *lac* UV5 promoter has been changed to a tandem sequence of four A·T base-pairs (UV5+1As), or T·A base-pairs (UV5+1Ts), or G·C base-pairs (UV5+1Gs) or C·G base-pairs (UV5+1Cs) as shown in Figure 4(a). The UV5+1Gs and UV5+1Ts constructs actually had five G·C and five T·A base-pairs at the start site due to the pre-existing structures. The mRNA primer extension analysis of the *lac* transcripts showed that changing the *lac* UV5 promoter start site to a run of five T·A base-pairs resulted in a pattern diagnostic of transcriptional slippage during the transcription initiation process *in vivo* (see Fig. 4(b)). The UV5+1As, UV5+1Gs and UV5+1Cs did not demonstrate transcriptional slippage patterns *in vivo*. We note, however, that UV5+1Cs appears to program a collection of 5′ ends. Because of their location (limited to a region near +1) we assume that these represent true ambiguity in the initiation reaction or 5′ end degradation.

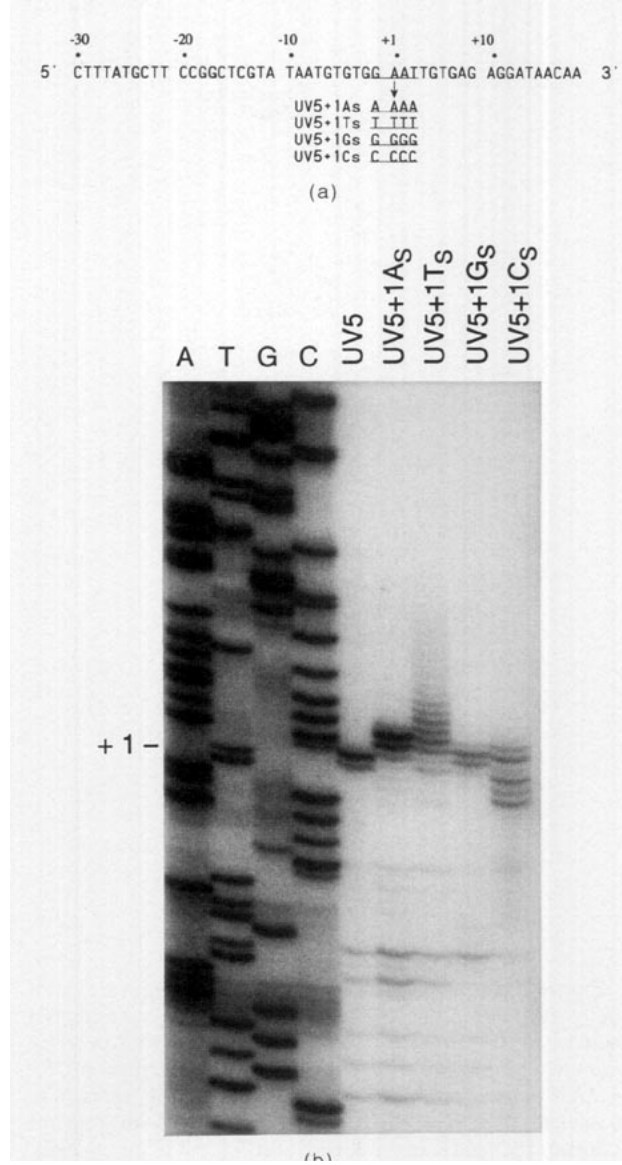

Figure 4 (a) DNA sequences of start site mutations in the *lac* UV5 promoter. The DNA sequence is numbered relative to the transcription start site of the *lac* UV5 promoter (+1). The transcription start site sequence GAAT of the *lac* UV5 promoter is underlined. The start site sequence GAAT has been replaced by AAAA in mutant UV5+1As; by TTTT in mutant UV5+1Ts; by GGGG in UV5+1Gs; and by CCCC in UV5+1Cs. (b) Primer extension analysis of start site mutations of the *lac* UV5 promoter. The *in vivo* transcripts of start site mutations of the *lac* UV5 promoter were examined by primer extension analysis in strain CSH26 as described in Materials and Methods. Lanes A, T, G and C represent a DNA sequencing ladder of the wild-type *lac* promoter region from plasmid pRZ3207. The *lac* UV5 promoter transcripts from pRZ6522 are shown in lane UV5. The transcripts of start site mutations are shown in each corresponding lane.

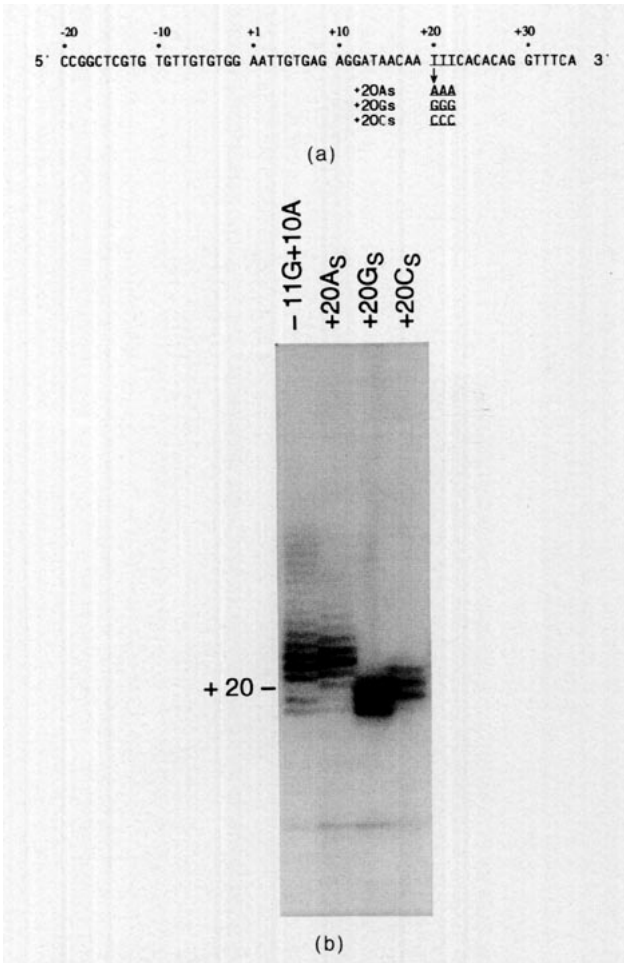

Figure 5. (a) DNA sequences of start site mutations of the +10A promoter. The DNA sequence is numbered relative to the transcription start site of the wild-type *lac* promoter (+1). The transcription start site sequence TTT of the +10A promoter is underlined. The start site sequence TTT has been replaced by AAA in mutant +20As; by GGG in +20Gs; and by CCC in +20Cs. (b) Primer extension analysis of start site mutations of the +10A promoter. The *in vivo* transcripts of start site mutations of the +10A promoter were examined by primer extension analysis in strain CSF26 as described in Materials and Methods. The +10A promoter transcripts are shown in lane −11G/+10A. The transcripts of start site mutations are shown in the corresponding lanes.

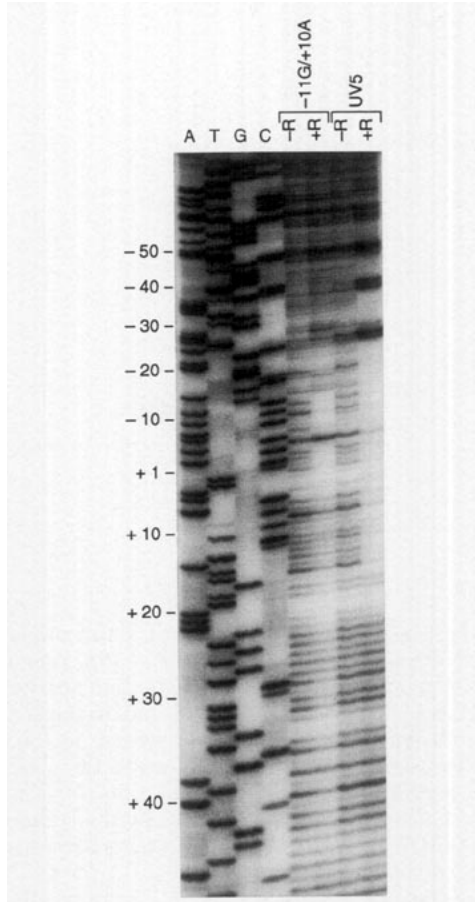

Figure 6. DNase I footprint protection studies of the +10A promoter. A 218 bp fragment (−11G/10A) containing +10A promoter (and lacking an active *lac* P1 promoter) was used to study the DNase I protection pattern by *E. coli* RNA polymerase. A 218 bp fragment containing the *lac* UV5 promoter was used in the control experiment. Lanes A, T, G and C represent a DNA sequencing ladder of the +10A promoter region from pRZ6505. The +1 site indicates the location of the wild-type *lac* promoter transcription initiation site. For each promoter the −R lane shows the DNase I cleavage pattern that results in the absence of RNA polymerase; the +R lane shows the cleavage pattern that results in the presence of RNA polymerase.

Transcriptional slippage preferably occurred for a run of T·A base-pairs at the start site in the *lac* UV5 promoter. Previously reported examples of transcriptional slippage *in vitro* also occurred during transcription initiation at a run of T·A or A·T base-pairs. To study if transcriptional slippage shows a similar preference at the +10A promoter, the start site (+20 to +22) of the *lac* +10A promoter has been changed to five consecutive A·T (+20As), or three G·C (+20Gs) or three C·G base-pairs (+20Cs) through site-directed mutagenesis as shown in Figure 5(a). The +20As mutation shows a similar but less extensive transcriptional slippage pattern as compared with the +10A promoter. The +20Gs and +20Cs mutations did not show a pattern of transcriptional slippage (see Fig. 5(b)).

(d) *DNase I footprint localization of the RNA polymerase binding site*

The +10A promoter sequence bears little resemblance to other sigma 70-dependent promoters. To test the idea that the apparent transcription initiation from +20 is driven by a nascent promoter created by the +10A mutation of the wild-type *lac* promoter, DNase I protection experiments were performed with sigma 70 *E. coli* RNA polymerase and a DNA fragment containing the *lac* promoter region of pRZ6505. Preliminary experiments indicated that +10A promoter fragments formed

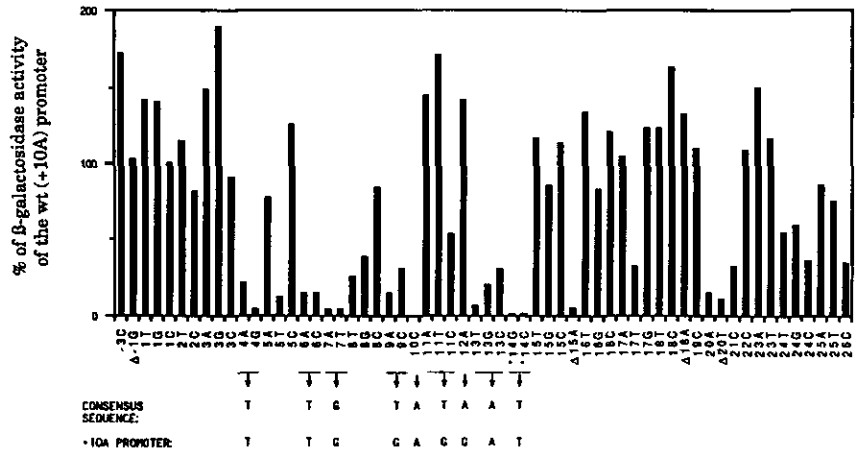

Figure 7. The effect of single point mutations on +10A promoter activity. Promoter sequence effect on +10A promoter activity was examined by measuring β-galactosidase activities of the indicated mutants as monocopy lysogens in strain CSH26; or in CSH26 Δ*cya* for those mutants derived from pRZ3207 (see Materials and Methods). The histograms are the β-galactosidase activities expressed as a percentage of activity found with the −11G/+10A construct (100%). Each value represents the average of 4 independent isolates. Standard deviations are less than 10% of sample averages. The −11G/+10A construct programmed 377 units of β-galactosidase expression, which was 30% of fully activated wild-type *lac* promoter (1238 units), but 6-fold greater than wild-type *lac* in the absence of CAP-cAMP (65 units). No β-galactosidase activity data for the +14G and +14C mutants are available because the phenotypes of the mutants are so weak that selection of the recombinant mutants from plasmid onto phage λ was unsuccessful. However, the phenotypes of the +14G and +14C DH5α transformants on Xgal indicator plates suggest that the β-galactosidase activities of these 2 mutants are less than that of the +5T mutant by at least several-fold.

stable complexes with RNA polymerase inefficiently. In order to maximize the recovery of RNA polymerase–+10A promoter complexes we formed ternary complexes by including UTP in the reaction mixture. In addition, in order to remove unbound DNAs from the footprinting analysis, we electrophoresed the samples after the limited nuclease digestion and analyzed the complexes that were retarded in mobility. The RNA polymerase–UV5 experiment was done in a similar fashion except that UTP was excluded in the preincubation. It has been reported that the RNA polymerase–UV5 binary complex footprint is identical with that for the ternary complex formed in the presence of ATP (Carpousis & Gralla, 1985). As shown in Figure 6, *E. coli* RNA polymerase bound to the +10A promoter yielded a well-protected region in the middle (from −15 to +5, relative to the wild-type *lac* promoter start site), and an enhancement at −5. A DNase I footprint pattern of the sigma 70-dependent UV5 promoter is also presented. The *lac* UV5 footprint pattern contains three enhanced bands in the upstream region at −25, −38 and −48, and a very well-protected region in the core and downstream regions (−23 to +20) of the *lac* UV5 promoter. The pattern of the resulting footprint of the +10A promoter is quite different from that found for UV5. In particular the downstream region of the +10A promoter is only partially protected if at all.

(e) *Identification of +10A promoter elements by mutational analysis*

The +10A promoter sequence has little resemblance to the consensus sigma 70 promoter sequence. To examine what sequences are important for +10A promoter activity, the promoter strength of all +10A promoter mutations has been quantified using β-galactosidase assays. As shown in Figure 7, promoter down mutations are clustered from +4 to +14 or around +20. Based on consensus promoter structure and the distance from the transcription initiation site, the −10 region of the +10A promoter should be located from +9 to +14 (see Fig. 7). The +10A mutation of the −11G *lac* promoter gives 377 units of β-galactosidase activity, while the wild-type −11G *lac* promoter gives no detectable β-galactosidase expression (data not shown). These results indicate that the +10A mutation is required for promoter function. All changes at positions +9, +10, +13 and +14 significantly decreased promoter activities. Changes from G to T or A at position +11 and from G to A at position +12 increased promoter activity less than twofold. However, all mutations at positions +4, +6 and +7 dramatically decreased promoter activities. These results are consistent with the −10 region being located at +9 to +14 and also indicate that sequences immediately upstream from the −10 region are crucial for +10A promoter activity.

The fact that the −35 region of +10A promoter has little homology with the consensus −35 sequence (2 out of 6 match to the −35 hexamer) and that the +10A promoter requires the sequences immediately upstream from the −10 hexamer suggested that this promoter may have some similarity to the *gal* P1 promoter, which contains a nonfunctional −35 region (Busby & Chan, 1989). To further test the importance of −35 region sequences in +10A promoter function, we constructed two additional promoter mutations through site-

directed mutagensis as shown in Figure 2. The −35dn mutation, which decreases further the weak −35 region homology (the −35dn mutation introduced the least favorable sequence (ACAGGC) according to Moyle *et al.*, 1991), did not affect +10A promoter activity suggesting that the −35 region was indeed non-functional. Consistent with that conclusion, we found that introduction of the consensus −35 hexamer TTGACA increased expression 10·6-fold (data not shown). The −35 up mutation programmed the same pattern of transcriptional slippage as other +10A promoter constructs (data not shown).

4. Discussion

Transcription initiation at the +10A mutant *lac* promoter has two unusual properties. First, transcripts synthesized from this promoter have multiple apparent start sites both *in vivo* and *in vitro*. Second, transcription initiation is not driven by an obvious promoter sequence (Karls *et al.*, 1989; Rothmel & LeClerk, 1989; this communication and R. Karls, personal communication). Here we report experiments that study these two properties of the +10A promoter.

(a) *The +10A promoter programs transcription slippage during the* in vivo *initiation process*

Our sequencing results demonstrate that the 5′ ends of +10A transcripts are composed of oligo(U) chains of varying lengths and the mRNA sequence downstream of +20 is homologous with its DNA template. This observation suggests that the multiple apparent start sites are due to transcriptional slippage that occurred during the transcription initiation process at +20. We propose that the sequence AAA (+20 to +22) in the DNA template strand is reiteratively copied by RNA polymerase through slippage of the initial oligo(U) RNAs across the surface of the template strand as illustrated by Figure 8 (see Guo & Roberts (1990) for a description of this model).

Previous investigators have reported the *in vitro* occurrence of transcriptional slippage (Chamberlin & Berg, 1964; Machida *et al.*, 1984; Harley *et al.*, 1990; Guo & Roberts, 1990). This is the first report of transcriptional slippage *in vivo* in *E. coli*. Transcription slippage *in vivo* has been reported previously for vaccinia virus late gene promoters (Wright & Moss, 1987).

(b) *Transcriptional slippage associated with initiation depends on the presence of a (T·A)-(T·A)-(T·A), or to a lesser extent (A·T)-(A·T)-(A·T), sequence at start sites*

At several promoters in which the transcription initiation sites consist of a stretch of three to five identical base-pairs, RNA polymerase incorporates *in vitro* more copies of the corresponding nucleotide than the template sequence codes for (Machida *et al.*, 1984; Harley *et al.*, 1990; Guo & Roberts, 1990; Jacques & Susskind, 1990). Guo & Roberts (1990) and Jacques & Susskind (1990) reported that any mutation that destroyed the start site homo-oligomeric sequence totally abolished transcriptional slippage *in vitro*. Here we provide additional evidence that a homo-oligomeric sequence at the start site is crucial for *in vivo* transcriptional slippage during the initiation process. Through the study of 59 point mutations in the +10A promoter region (Fig. 2) we found that with one exception (a change immediately preceding the start site sequence) only the mutations that changed the start site sequence abolished the transcriptional slippage. The fact that changing the start site of another promoter, the *lac* UV5 promoter, to five consecutive T·A base-pairs (from −1 to +4), caused transcriptional slippage is further evidence for the requirement of a homo-oligomeric sequence at the start site.

In previously reported cases, such transcriptional slippage during initiation occurred when the DNA template strands consisted of tandem T or tandem A bases at the start sites. We changed the *lac* UV5 promoter start site to four consecutive A·T, five T·A, five G·C and four C·G base-pairs. Only the change to five consecutive T·A base-pairs showed transcriptional slippage initiation. The +10A promoter contains three consecutive T·A base-pairs at the start site. We also changed the +10A start site to three A·T (resulting in 5 A·T base-pairs in a

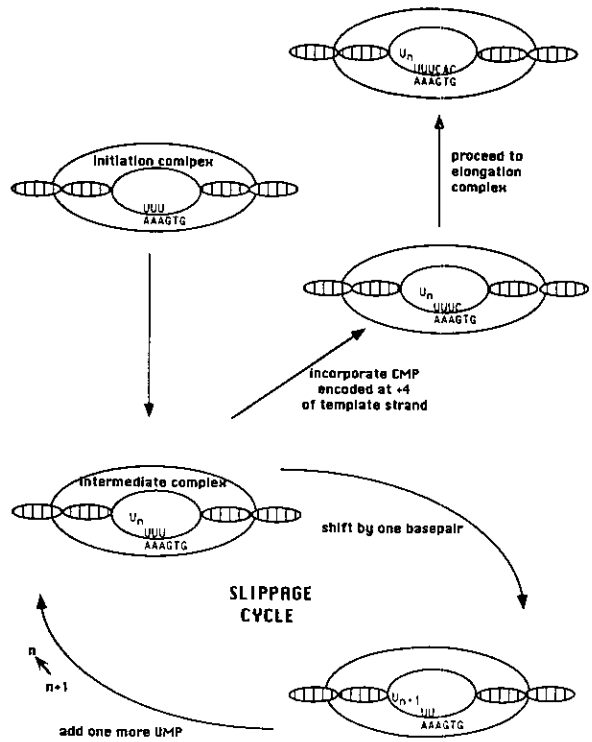

Figure 8. The slippage event of transcription initiation of the *lac* +10A promoter. This cartoon of transcriptional slippage is similar to that presented by Guo & Roberts (1990).

run), G·C or C·G base-pairs to see which homo-oligomeric run favors transcriptional slippage during initiation at the +10A promoter. The change to five A·T base-pairs also showed transcriptional slippage, though to a lesser extent than found for T·A base-pairs. However, the G·C and C·G constructs did not demonstrate transcriptional slippage.

Based on the slippage model illustrated in Figure 8, transcriptional slippage during the initiation process requires that the initiated transcript slip one base-pair toward the 5′ end across the surface of the template DNA strand. Such slippage will result in the breakage of hydrogen bonds between RNA·DNA base-pairs. It would be reasonable to assume that breaking A·U and T·A base-pairs will be easier than breaking G·C base-pairs. This may provide one explanation for why transcriptional slippage preferentially happens when the DNA template strand consists of consecutive A or T bases at the start site.

However, we also observed an apparent preference for T·A *versus* A·T runs at the start site. This observation may be explained by the lower stability of dA·rU base-pairs *versus* rA·dT base-pairs. Martin & Tinoco (1980) have found that a DNA–RNA hybrid oligonucleotide duplex that contains a $(dA \cdot rU)_5$ sequence is at least 200 times less stable at room temperature than the corresponding duplex containing an $(rA \cdot dT)_5$ sequence.

(c) *Transcriptional slippage during the initiation process is not promoter specific*

Transcriptional slippage was observed by Guo & Roberts (1990) for the late gene promoter of bacteriophage 82 *in vitro*. When the upstream sequences of the λP_R promoter was fused with the sequence downstream from +1 of the bacteriophage 82 late gene promoter, it was found that this fusion promoter can also show transcriptional slippage initiation *in vitro* (Guo & Roberts, 1990). These results demonstrated that the upstream and core regions of a promoter are not specifically required for transcriptional slippage initiation. However, these experimental results did not rule out a dependence on the downstream region of a promoter. The fact that changing the start site of the *lac* UV5 promoter to five consecutive T·A base-pairs is sufficient to cause transcriptional slippage initiation supports the idea that this process is not promoter specific and that it primarily depends on a run of T·A (or sometimes A·T) base-pairs at the start site. However, although the start site structure is obviously the major factor determining whether transcriptional slippage initiation occurs or not, the rest of the promoter sequence may influence the process. We note that the A·T construct of the +10A promoter gives rise to transcriptional slippage initiation while the A·T construct of UV5 does not. This difference may be a consequence of the sequence immediately around the start site as suggested by the effect of the +19C mutation on the +10A promoter. Alternatively it may suggest influences of the overall promoter structure on this process.

Does transcriptional slippage initiation occur at other *E. coli* promoters *in vivo*? From the review by Harley & Reynolds (1987), of all 263 *E. coli* promoters with known transcriptional start points, we identified seven promoters that contain three or more consecutive T·A base-pairs at the transcription initiation site. They are: Fplas-traY/Z IS1ins PL, MuPe-1, Tn2661bla-Pa, fumA, ilvIH-P1 and metBL. Transcriptional slippage initiation had been reported for IS1ins PL *in vitro* (Machida *et al.*, 1984). Fplas-traY/Z showed multiple apparent start sites in an *in vitro* transcription experiment. However further characterization of Fplas-traY/Z *in vitro* transcripts demonstrating products of transcriptional slippage has not been reported (Fowler *et al.*, 1983). Unfortunately the 5′ ends of the mRNAs from the other five promoters were studied mainly by the S_1 nuclease mapping technique. Since the pseudo-templated product of transcriptional slippage initiation is not homologous to the template DNA at the 5′ end, S_1 mapping would not allow one to observe these products.

(d) *Transcriptional slippage initiation and abortive initiation*

A number of investigators have reported that during *in vitro* transcription initiation RNA polymerase undergoes an abortive initiation event in which oligoribonucleotides complementary to the start region template are synthesized, released and RNA polymerase reinitiates. There are some superficial similarities between the abortive initiation phenomenon and transcriptional slippage initiation: namely, that a short oligonucleotide is made, the oligonucleotide is released from the template and RNA polymerase reinitiates RNA synthesis. There is, however, one obvious difference between the two processes. During transcriptional slippage initiation the product RNA is not released from the enzyme and in fact serves as a primer. It could be that the two processes are mechanistically related with the critical differences being that oligo(U) (or oligo(A)) might bind more tightly to the active site of RNA polymerase holoenzyme than do mixed oligonucleotides, and that they are capable of base-pairing to the start site sequence.

A mechanistic relationship between transcriptional slippage initiation and abortive initiation leads to two predictions. Promoters that demonstrate transcriptional slippage initiation should program the synthesis of fewer recycled products. This in turn suggests that, for these promoters, a higher proportion of initiation events would lead to elongating transcripts.

(e) *A possible physiological role for 5′ oligo(U) sequences*

It has recently been reported that *E. coli* mRNAs can carry poly(A) tracts at their 3′ ends (Cao &

Sarkar, 1992). This observation suggests that oligo(U) tracts at the 5′ ends could facilitate a base-pairing interaction between the 5′ end and 3′ end perhaps affecting mRNA stability. We are currently designing experiments to test if such a 5′–3′ interaction occurs and, if so, whether it modulates the half-life and/or translation yield of the mRNA.

(f) *The +10A promoter requires extended sequences upstream from the −10 region to compensate for its poor −10 region and the non-functional −35 region*

A single point mutation at the +10 position in the wild-type *lac* promoter resulted in new transcription initiation from +20. Initially, we were not sure how this single change generated a new promoter. By examining the DNA sequences in the region that could program transcription starting at +20, two hexamers were identified that might function as −10 and −35 regions (Fig. 2). Other possible −10 and −35 regions have been discussed by Karls *et al.* (1989). However, even these two hexamers have very poor homologies to the consensus −10 and −35 regions (Fig. 2). In the proposed −10 region, only three out of six positions are identical with the consensus sequence; in the proposed −35 region only two out of six positions are identical with the consensus sequence. Based on homology scores calculated by the method of Mulligan *et al.* (1984), we obtained a homology score of only 37·3 for the +10A promoter. This score ranks the promoter in the bottom 10% of all 112 promoters examined by Mulligan *et al.* (1984). However, the +10A promoter (containing the −11G mutation to eliminate wild-type *lac*P1) programs sixfold higher levels of *lac*Z expression than the wild-type *lac*P1 promoter (in the absence of CAP-cAMP, Karls *et al.*, 1989). The homology score of wild-type *lac*P1 promoter is 49·7. This comparison suggests that the +10A promoter's activity is not accurately predicted by its homology score and, therefore, that it is not a typical promoter.

We studied the +10A promoter sequence by analyzing the effects of mutations on β-galactosidase expression. A general principle for predicting promoter activity is that any change toward consensus increases promoter activity while any change away from consensus decreases promoter activity. This principle is basically true for the proposed −10 region (+9 to +14) of this new promoter. Changes toward consensus at +11 (from G to T), at +12 (from G to A) did increase promoter activity. Changes away from consensus at +13 (from A to T, G or C) did decrease promoter activity.

When examining the sequences upstream from the proposed −10 region (+4 to +8), we noticed that many mutations in this region dramatically decreased promoter activity. The effects are most obvious for positions +4, +6 and +7. However, the +4 T and +6, +7 TG are only weakly conserved in *E. coli* promoters. The dependence of the +10A promoter on these three nucleotides reveals some similarities to the *gal* P1 promoter. The *gal* P1 promoter is also of low homology to the consensus *E. coli* promoter sequence and also requires a T and TG at positions equivalent to +4 and +6, +7 of the +10A promoter (Busby & Chan, 1989; Ponnambalam *et al.*, 1986).

Also like the *gal* P1 promoter, the +10A promoter contains a non-functional −35 region. However, the change to a consensus −35 hexamer increases the +10A promoter activity more than tenfold. It will be interesting to test if the consensus −35 region can eliminate the requirement for the sequences immediately upstream from the −10 region.

Besides *gal* P1 and *lac* +10A promoters, several other promoters, such as *gal* P2, λ Pre, *omp*F and *uvr*A, have been reported to function without a specific −35 region (Ponnambalam *et al.*, 1986; Keilty & Rosenberg, 1987; Inokuchi *et al.*, 1984; Backendorf *et al.*, 1983). A comparison of these promoter sequences reveals sequence homologies different from that of the conventional sigma 70-dependent promoter sequence (Fig. 9). The consensus −10 hexamer of a conventional promoter is TAtaaT (capital letters indicate the most highly conserved bases). However, the −10 region of this promoter subclass is -A-gaT. Note, however, that, for the +10A promoter, the mutant data suggests the sequence -A-AAT is somewhat preferred. These promoters also exhibit sequence conservation in the immediate upstream region of the −10 region (T at −17 and TG at −15, −14). These sequences have been demonstrated to be crucial for promoter function of *lac* +10A, *gal* P1 and λ Pre (this work; Ponnambalam *et al.*, 1986; Keilty & Rosenberg, 1987). As expected, this subclass of sigma 70-dependent promoters does not show obvious homology in the −35 region. However, in regions upstream of the −35 region, T is highly conserved at three positions (−42, −52 and −72). It is not clear whether these three T bases play any role in promoter function. For the *gal* P1 promoter, deletion of the region upstream of −49 has been shown to decrease promoter activity twofold (Busby *et al.*, 1987). The spacing of these three T bases (10 and 20 bp) suggests that RNA polymerase could directly contact these nucleotides because they are located on the same side of the DNA helix.

5. Conclusion

Our analyses of the +10A promoter have shown that: (1) transcriptional slippage initiation producing RNAs with oligo(U) (or sometimes oligo(A)) at the 5′ end occurs *in vivo*. (2) The occurrence of transcriptional slippage initiation is primarily a function of the start site sequence, occurring when the sequence is (T·A)−(T·A)−(T·A) (or (A·T)−(A·T)−(A·T)). (3) A subset of sigma 70-dependent promoters exist that partially compensates for a non-functional −35 region with the sequence TXTGX immediately upstream of the −10 region.

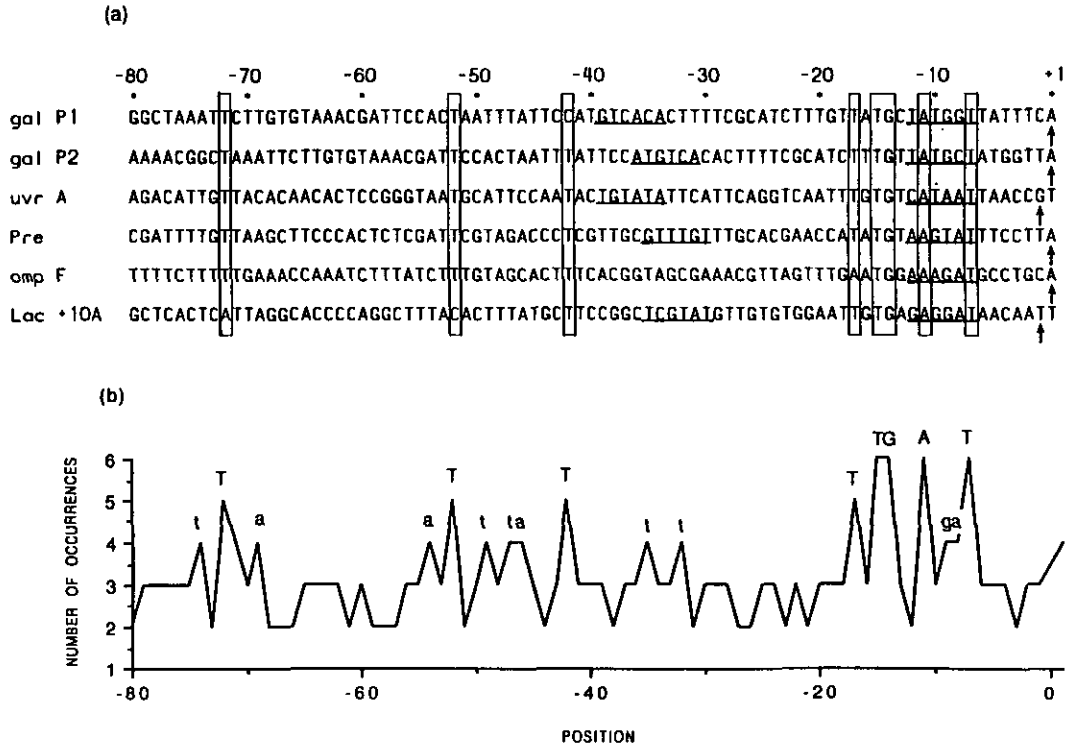

Figure 9. Sequence comparison of a subclass of *E. coli* sigma 70-dependent promoters without specific −35 regions. (a) DNA sequences of the promoters *gal* P1, *gal* P2, *uvr* A, λ Pre, *omp*F and *lac* +10A are listed. The transcription initiation start points are indicated by (↑). The −10 and −35 hexamers are underlined. The highly conserved bases among these 6 promoters are boxed. (b) The diagram displays the number of occurrences of the most prevalent base at each position. The bases that occur in at least 5 out of 6 promoters are capitalized. The bases that occur in 4 out of the 6 promoters are indicated by lower-case letters.

We thank Russell Karls for providing some of the +10A promoter mutations and for helpful discussions. We also thank C. Gross, R. Gourse and P. Kiley for helpful discussions and comments on the manuscript. This work was supported by grant DMB9020517 from the N.S.F.

References

Air, G. M. (1979). Nucleotide sequence coding for the "signal peptide" and N terminus of the hemagglutinin from an Asian (H2N2) strain of influenza virus. *Virology*, **97**, 468–472.

Backendorf, C., Brandsma, J. A., Kaitasova, T. & van de Putte, P. (1983). *In vivo* regulation of the *uvr*A gene: role of the "−10" and "−35" promoter regions. *Nucl. Acids Res.* **11**, 5795–5810.

Busby, S. & Chan, B. (1989). Recognition of nucleotide sequences at the *Escherichia coli* galactose operon P1 promoter by RNA polymerase. *Gene*, **84**, 227–236.

Busby, S., Spassky, A. & Chan, B. (1987). RNA polymerase makes important contacts upstream from base pair −49 at the *Escherichia coli* galactose operon P1 promoter. *Gene*, **53**, 145–152.

Cao, G. J. & Sarkar, N. (1992). Poly(A) RNA in *Escherichia coli*: nucleotide sequence at the junction of the Ipp transcript and the polyadenylate moiety. *Proc. Nat. Acad. Sci., U.S.A.* **89**, 7546–7550.

Carpousis, A. J. & Gralla, J. D. (1980). Cycling of ribonucleic acid polymerase to produce oligonucleotides during initiation *in vitro* at the *lac* UV5 promoter. *Biochemistry*, **19**, 3245–3253.

Carpousis, A. J. & Gralla, J. D. (1985). Interaction of RNA polymerase with *lac* UV5 promoter DNA during mRNA initiation and elongation. *J. Mol. Biol.* **183**, 165–177.

Chamberlin, M. & Berg, P. (1964). Mechanism of RNA polymerase action: characterization of the DNA-dependent synthesis of polyadenylic acid. *J. Mol. Biol.* **8**, 708–726.

DeBorde, D. C., Naeve, C. W., Louise Herlocher, M. L. & Maassab, H. F. (1986). Resolution of a common RNA sequencing ambiguity by terminal deoxynucleotidyl transferase. *Anal. Biochem.* **157**, 275–282.

Fowler, T., Taylor, L. & Thompson, R. (1983). The control region of the F plasmid transfer operon: DNA sequence of the *traJ* and *traY* genes and characterization of the *traY*-Z promoter. *Gene*, **26**, 79–89.

Guo, H. C. & Roberts, J. F. (1990). Heterogeneous initiation due to slippage at the bacteriophage 82 late gene promoter *in vitro*. *Biochemistry*, **29**, 10702–10709.

Harley, C. B. & Reynolds, R. P. (1987). Analysis of *E. coli* promoter sequences. *Nucl. Acids Res.* **15**, 2343–2361.

Harley, C. B., Lawrie, J., Boyer, H. W. & Hedgpeth, J. (1990). Reiterative copying by *E. coli* RNA polymerase during transcription initiation of mutant pBR322 *tet* promoters. *Nucl. Acids Res.* **18**, 547–552.

Hawley, D. K. & McClure, W. R. (1983). Compilation and analysis of *Escherichia coli* promoter DNA sequences. *Nucl. Acids Res.* **11**, 2237–2255.

Inokuchi, K., Furukawa, H., Nakamura, K. & Mizushima, S. (1984). Characterization by deletion mutagenesis *in vitro* of the promoter region of *omp*F, a positively regulated gene of *Escherichia coli*. *J. Mol. Biol.* **178**, 653–668.

Jacques, J. P. & Susskind, M. M. (1990). Pseudo-templated transcription by *Escherichia coli* RNA polymerase at a mutant promoter. *Genes Develop.* **4**, 1801–1810.

Karls, R., Schulz, V., Jovanovich, S. B., Flynn, S., Pak, A. & Reznikoff, W. S. (1989). Pseudorevertants of a *lac* promoter mutation reveal overlapping nascent promoters. *Nucl. Acids Res.* **17**, 3927–3939.

Keilty, S. & Rosenberg, M. (1987). Constitutive function of a positively regulated promoter reveals new sequences essential for activity. *J. Biol. Chem.* **262**, 6389–6395.

Machida, C., Machida, Y. & Ohtsubo, E. (1984). Both inverted repeat sequences located at the ends of IS1 provide promoter functions. *J. Mol. Biol.* **177**, 247–267.

Maquat, L. E. & Reznikoff, W. S. (1978). *In vitro* analysis of the *Escherichia coli* RNA polymerase interaction with wild-type and mutant lactose promoters. *J. Mol. Biol.* **125**, 467–490.

Maquat, L. E. & Reznikoff, W. S. (1980). *Lac* promoter mutation $P^{r}115$ generates a new transcription initiation point. *J. Mol. Biol.* **139**, 551–556.

Martin, C. T., Muller, D. K. & Coleman, J. E. (1988). Processivity in early stages of transcription by T7 RNA polymerase. *Biochemistry*, **27**, 3966–3974.

Martin, F. H. & Tinoco, I., Jr (1980). DNA-RNA hybrid duplexes containing oligo(dA : rU) sequences are exceptionally unstable and may facilitate termination of transcription. *Nucl. Acids Res.* **8**, 2295–2299.

Miller, J. H. (1972). In *Experiments in Molecular Genetics*, pp. 352–356, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Moyle, H., Waldburger, C. & Susskind, M. M. (1991). Hierarchies of base pair preferences in the P22 *ant* promoter. *J. Bacteriol.* **173**, 1944–1950.

Mulligan, M. E., Hawley, D. K., Entriken, R. & McClure, W. R. (1984). *Escherichia coli* promoter sequences predict *in vitro* RNA polymerase selectivity. *Nucl. Acids Res.* **12**, 789–801.

Munson, L. M. & Reznikoff, W. S. (1981). Abortive initiation and long ribonucleic acid synthesis. *Biochemistry*, **20**, 2081–2085.

Ponnambalam, S., Webster, C., Bingham, A. & Busby, S. (1986). Transcription initiation at the *Escherichia coli* galactose operon promoters in the absence of the normal −35 region sequences. *J. Biol. Chem.* **261**, 16043–16048.

Rothmel, R. K. & LeClerk, J. E. (1989). Mutational analysis of the *lac* regulatory region: second-site changes that activate mutant promoters. *Nucl. Acids Res.* **17**, 3909–3925.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*, pp. 15.53–15.59, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Wright, C. F. & Moss, B. (1987). *In vitro* synthesis of vaccinia virus late mRNA containing a 5′ poly(A) leader sequence. *Proc. Nat. Acad. Sci., U.S.A.* **84**, 8883–8887.

Xiong, X. F., de la Cruz, N. & Reznikoff, W. S. (1991). Downstream deletion analysis of the *lac* promoter. *J. Bacteriol.* **173**, 4570–4577.

Edited by M. Gottesman