Ruffner, D. E., Dahm, S. C., & Uhlenbeck, O. C. (1989) *Gene 82*, 31–41.
Sampson, J. R., Sullivan, F. X., Behlen, L. S., Direnzo, A. B., & Uhlenbeck, O. C. (1987) *Cold Spring Harbor Symp. Quant. Biol. 52*, 267–275.
Sampson, J. R., Direnzo, A. B., Behlen, L. S., & Uhlenbeck, O. C. (1990) *Biochemistry 29*, 2523–2532.
Schneider, T. D., Stormo, G. D., Gold, L., & Ehrenfeucht, A. (1986) *J. Mol. Biol. 188*, 415–431.
Sheldon, C. C., & Symons, R. H. (1989) *Nucleic Acids Res. 17*, 5679–5685.
Simoncsits, A., Brownlee, G. G., Brown, R. S., Rubin, J. R., & Guilley, H. (1977) *Nature 269*, 833–836.
Uhlenbeck, O. C. (1987) *Nature 328*, 596–600.
Yu, K.-I., Hsu, S.-P., Heiss, R. E., Jr., & Hasiuk, L. Z. (1987) *Quest 4*, 5–19.
Zaug, A. J., Grosshans, C. A., & Cech, T. R. (1988) *Biochemistry 27*, 8924–8931.

Heterogeneous Initiation Due to Slippage at the Bacteriophage 82 Late Gene Promoter in Vitro†

Hwai-Chen Guo‡ and Jeffrey W. Roberts*
Section of Biochemistry, Molecular and Cell Biology, Biotechnology Building, Cornell University, Ithaca, New York 14853
Received May 4, 1990; Revised Manuscript Received August 6, 1990

ABSTRACT: RNAs synthesized in vitro by purified *Escherichia coli* RNA polymerase from a bacteriophage 82 promoter are heterogeneous at the 5′ end. We show that this heterogeneity results from variable addition of extra adenine residues, allowed by slippage of the initial oligonucleotide pppAAA-OH against its DNA template sequence TTT. Slippage backward by one base allows another A to be added, giving pppAAAA-OH, and this cycle can continue more than 20 times before it is ended by incorporation of UMP encoded by the fourth template base A. Slippage is abolished by mutation of the TTT template sequence to TGT and is sensitive to the concentrations of UTP and ATP in the reaction mixture. Analysis of deletions, substitutions, and point mutants implies that the slippage reaction requires only the existence of TTT at the initiation site of the template strand, although changes in neighboring nucleotides slightly affect its efficiency.

Initiation of an RNA chain by *Escherichia coli* RNA polymerase is thought to comprise a defined sequence of steps: (i) RNA polymerase locates a promoter and binds specifically to form an inactive intermediate, the "closed complex"; (ii) the closed complex isomerizes into an "open complex" in which about 16 nucleotides of DNA are melted to expose the template strand; (iii) the template-directed synthesis of an RNA chain begins; and (iv) RNA polymerase clears the promoter and enters a distinct elongation mode [for reviews, see von Hippel et al. (1984) and McClure (1985)].

Successful initiation requires that promoter binding be tight enough to form the open complex, but not too tight to allow the escape of polymerase from the promoter to elongate mRNA. This inherent conflict appears to be expressed in certain irregularities of the initiation process. For example, nascent transcripts up to nine nucleotides long can be released from the RNA polymerase–promoter complex without dissociation of the polymerase from the promoter, whereupon the enzyme starts another chain from the beginning (Johnston & McClure, 1976; Carpousis & Gralla, 1980, 1985; Reznikoff et al., 1982; Spassky, 1986; Levin et al., 1987; Krummel & Chamberlin, 1989). Such "abortive initiation" occurs in vitro at a variety of *E. coli* promoters, and may reflect an activity common to both prokaryotic and eukaryotic RNA polymerase (Yamakawa et al., 1981; Cowie et al., 1982; Luse & Jacob, 1987). For the bacterial enzyme, conversion from a recycling abortive complex to a stable elongation complex has been correlated with loss of the σ subunit after transcription of 9–11 nucleotides (Hansen & McClure, 1980; Straney & Crothers, 1985). To account for abortive cycling and for some heterogeneity in transcription initiation sites, Carpousis et al. (1982) proposed that promoter-bound RNA polymerase chooses its initiating nucleotide and synthesizes short RNA chains from a small initiation region without breaking the original strong contacts that form the open complex.

Here we demonstrate a further variant of the initiation process that produces extensive 5′ heterogeneity of the late gene promoter transcript of bacteriophage 82: the initial trinucleotide pppAAA of the RNA slips against the template TTT, beginning a slippage cycle through which the transcript acquires an untemplated initial sequence of poly(A) that may exceed 20 nucleotides. A similar slippage reaction in a mutant promoter was reported by Harley et al. (1990). This reaction may be similar enzymatically to less natural reactions in which a DNA homopolymer, an RNA homopolymer, or a short homopolymeric DNA sequence within a larger chemically random DNA sequence can serve as template to make a long complementary-sequence RNA homopolymer (Falashi et al., 1963; Stevens, 1964; Chamberlin & Berg, 1964; Fox & Weiss, 1964; Krakow & Karstadt, 1967).

Experimental Procedures

Materials. All materials were of the highest purity available and were purchased from commercial sources, unless otherwise stated. NusA protein was purified by the method of Schmidt and Chamberlin (1984b) with the modification described by Goliger and Roberts (1987). RNA polymerase was purified in this lab by the methods of Burgess and Jendrisak (1975) as modified by Lowe et al. (1979). Radiolabeled nucleoside triphosphates were purchased from Amersham; regular and "ultrapure" (grade 3) nucleoside triphosphates were from

† This work was supported by Grant GM 21941 from the National Institutes of Health.
‡ Present address: Department of Biochemistry and Molecular Biology, 7 Divinity Ave., Cambridge, MA 02138.

0006-2960/90/0429-10702$02.50/0

Pharmacia; oligoribonucleotide primers were from Sigma; and a synthetic oligonucleotide used to prime cDNA synthesis was from the Cornell University Biotechnology Program Oligonucleotide Synthesis Facility.

Bacterial Strain. Plasmids were constructed and maintained in the *E. coli* strain HB101 [F^-, *hsdS20* (r_B, m_B), *recA13*, *ara-14*, *proA2*, *lacY1*, *galK2*, *rpsL20* (Sm^r), *xyl-5*, *mtl-1*, *supE44*, λ^-].

Plasmids. The point mutations were made by standard methods of oligonucleotide-directed mutagenesis (Kunkel et al., 1987). Mutation designations in Figures 4b and 7 indicate the mutant base in the RNA-like DNA strand, relative to the wild-type sequence of Figure 1b. Phage 82 late gene promoter fusions $\Delta 82(+35)t_0$ and $\Delta 82(+51)t_0$ were described previously (Goliger & Roberts, 1989). Construction of the promoter fusion *qut*λ82 is described elsewhere (H.-C. Guo, Ph.D. Thesis, Cornell University, 1990).

In Vitro Transcription. Transcription in vitro was performed as described by Grayhack and Roberts (1982); 150 nM NusA protein was present in all reactions, although it does not detectably affect the heterogeneity of transcripts. Grade 3 ("ultrapure") nucleoside triphosphates were used for the experiments of Figures 2 and 4.

For experiments of Figure 1, transcription reactions containing 50 mM Tris-HCl (pH 7.9), 0.1 mM EDTA, 1 mM DTT, 80 mM KCl, 200 μM ATP, 200 μM GTP, 200 μM CTP, 25 μM UTP, 2.5 μCi of [α-^{32}P]UTP, approximately 1–2 nM DNA template, 10–20 nM purified RNA polymerase holoenzyme, and 150 nM NusA protein were preincubated 10 min at 37 °C. Transcription was initiated by the addition of $MgCl_2$ to a final concentration of 5 mM, and two aliquots were removed and stopped at 40 and 60 s after initiation. Rifampicin was added at 70 s to a final concentration of 10 μg/mL, and two more aliquots were removed at 80 and 100 s. We found that the +2C mutation decreases promoter strength (data not shown); therefore, rifampicin was added about a minute after initiation by Mg^{2+} to prevent continued initiation, rather than simultaneously, as is usually done to produce a single round of transcription from the wild-type promoter. Transcription was stopped by adding 75 μL of transcription stop buffer [0.66 M Tris-HCl (pH 7.9), 15 mM EDTA, and 0.25 mg/mL tRNA] and chilling on ice.

The reactions of Figure 2 contained 50 mM Tris-HCl (pH 7.9), 0.1 mM EDTA, 1 mM DTT, 80 mM KCl, 5 mM $MgCl_2$, 1 or 5 μM ATP, no GTP (in order to stall transcription at +6), 5 μM CTP and UTP, 2.5 μCi of [α-^{32}P]UTP, 100 μM oligoribonucleotide ApU, UpA, or ApApU as indicated, approximately 2 nM DNA template, 10–20 nM *E. coli* RNA polymerase holoenzyme, and 150 nM NusA protein. After an incubation of 10 min at 37 °C, the reaction mixture was divided; one portion was made 0.38 M in KCl and incubated 2 min further ("salt challenge"), and both were chased by adding all four rNTPs and rifampicin to final concentrations of 200 μM and 10 μg/mL, respectively. Reactions were stopped at 0.25 min after chasing. Since transcription complexes at the early initiation stage (abortive cycle) are sensitive to rifampicin, addition of rifampicin to the +6 stalled complexes (instead of at the beginning) reduces inhibition of transcription by rifampicin, particularly with ATP in the low concentration that is necessary for selective primer initiation, while still allowing a single round of synchronized transcription. This manipulation results in a 5-fold increase in productive transcription over that obtained by using the regular initiation protocol in which rifampicin and Mg^{2+} are added together (data not shown).

For the reactions with low [UTP]/[ATP] (2 μM/300 μM) (Figure 4a), a mixture containing 50 mM Tris-HCl (pH 7.9), 0.1 mM EDTA, 1 mM DTT, 40 mM KCl, 300 μM ATP, 200 μM GTP, 200 μM CTP, 2 μM UTP, 2.5 μCi of [α-^{32}P]UTP, approximately 2 nM DNA template, 10–20 nM *E. coli* RNA polymerase holoenzyme, and 150 nM NusA protein was preincubated 8 min at 37 °C. Transcription was initiated by the addition of $MgCl_2$ to a final concentration of 5 mM, and one aliquot was removed and stopped at 0.3 min after initiation. Rifampicin was added at 0.8 min to a final concentration of 10 μg/mL, and further aliquots were removed at the indicated times. For reactions with high [UTP]/[ATP] (400 μM/20 μM) shown in Figure 4a, transcription was performed as above, except with 20 μM ATP, 400 μM UTP, and 2.5 μCi of [α-^{32}P]ATP. Reactions shown in Figure 4b were performed like those of Figure 2, except with 15 μM ATP, no GTP, 10 μM CTP, 300 μM UTP, no oligoribonucleotide primer, and 2.5 μCi of [α-^{32}P]ATP, and the incubation was 6.5 min at 37 °C. The stalled transcription complexes were chased by adding GTP, CTP, and rifampicin to final concentrations of 200 μM, 200 μM, and 10 μg/mL, respectively. Aliquots were removed and added to stop buffer at the indicated times after the chase began.

Reaction mixtures for Figure 7 contained 80 μM ATP, 80 μM GTP, 80 μM CTP, 5 μM UTP, and 2.5 μCi of [α-^{32}P]-UTP; after incubation for 6.5 min at 37 °C, synthesis was initiated by the simultaneous addition of $MgCl_2$ and rifampicin to final concentrations of 5 mM and 10 μg/mL, respectively.

For reactions shown in Figure 8, transcription was performed as described for Figure 1 except that that rifampicin was added at 0.7 min and aliquots were removed at the indicated times after initiation.

RNA was extracted with 100 μL of phenol/chloroform/isoamyl alcohol (1/1/0.04), ethanol precipitated, and resuspended in 3 μL of formamide loading dye containing 0.01% bromophenol blue and 0.01% xylene cyanol. RNA was resolved by electrophoresis through 0.4-mm-thick 50% urea–polyacrylamide sequencing gels and visualized by autoradiography.

cDNA Synthesis and Sequencing. *Hin*dIII–*Sma*I restriction fragments of DNAs $\Delta 82(+35)t_0$ and $\Delta 82(+51)t_0$ (Goliger et al., 1989) containing the phage 82 late gene promoter and t_0 terminator were used as templates in 50-μL reaction mixtures containing different [UTP]/[ATP] ratios as indicated. Transcription reactions were performed as described for Figure 4a, but without rifampicin or radiolabeled nucleoside triphosphates, and stopped after 20 min by phenol extraction. DNase treatment of template DNA before cDNA synthesis was not necessary. A ^{32}P-5′-end-labeled oligonucleotide (5′-CACCACAGAAAGGTCG-3′) complementary to a region of the t_0 terminator segment in the terminated transcripts was used as primer to synthesize cDNAs from the $\Delta 82(+35)t_0$ or $\Delta 82(+51)t_0$ transcripts as described by Maniatis et al. (1982). The cDNAs (about 90 nucleotides) were resolved by polyacrylamide gel electrophoresis, and the cluster ranging from the 4th to the 10th band (counting from the shortest cDNA) was excised from the gel, extracted in TE buffer, subjected to three Tris/ethanol precipitations, and sequenced according to Maxam and Gilbert (1977). Note that the uncleaved shorter cDNAs appear in all reactions and therefore mask the sequencing pattern of longer cDNAs in Figure 6.

Results

Transcript 5′ Heterogeneity at the Bacteriophage 82 Late Gene Promoter Requires the Template Sequence TTT. RNA initiated at the phage 82 late gene promoter p_{82}—a constitutive

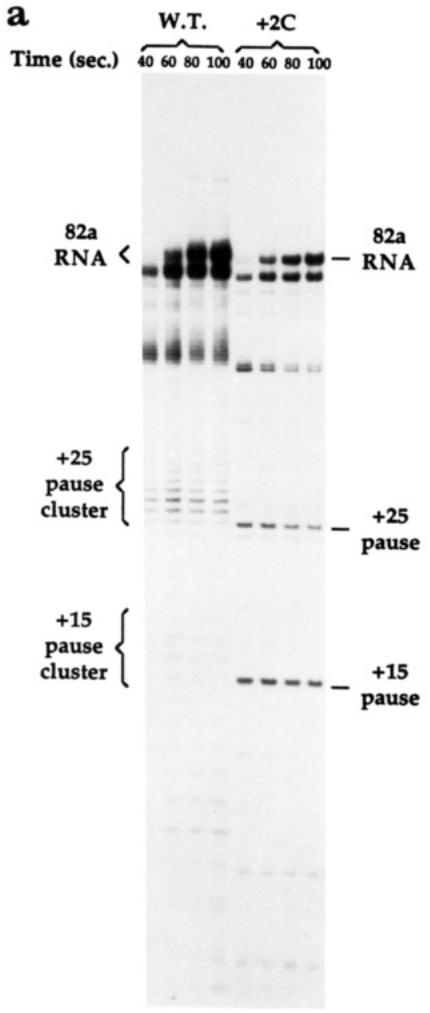

FIGURE 1: (a) Heterogeneity of phage 82 late gene promoter transcripts and the effect of a +2 point mutation on heterogeneity. An autoradiogram of transcription time courses from the native phage 82 (W.T.) or +2C mutant DNA template is shown. Products of the early pauses (+15 and +25) and the terminated RNAs (82a) are indicated. Other higher molecular weight bands are paused transcription complexes and are prominent only at short times (compare Figures 2 and 4a). Faint smaller RNAs are abortively initiated products. (b) Nucleotide sequence of the 82 late gene promoter region (RNA-like strand) showing oligonucleotides used to prime RNA synthesis. Oligonucleotides in parentheses prime less efficiently than does UA.

promoter that is actively transcribed in the absence of activators in vitro and presumably in vivo (Goliger & Roberts, 1987)—is resolved by polyacrylamide gel electrophoresis as clusters of bands, both at the two early transcriptional pauses and at the terminator (Goliger & Roberts, 1987; Yang et al., 1989; Figure 1a). Such heterogeneity could reflect initiation spread across the three dT's (or other nucleotides) in the template strand at the initiation site (Figure 1b), or heterogeneity at the pause or termination site, or internal imprecision of templating, or a combination of these processes. Since end-labeled RNAs cut at discrete sites by RNase T1 still consist of clusters of bands (Goliger & Roberts, 1987), heterogeneous initiation must be at least partly responsible.

If transcription initiates variously at the three adjacent T's [the first A residue of the RNA was designated "+1" by Goliger and Roberts (1987); see Figure 1b], a point mutation

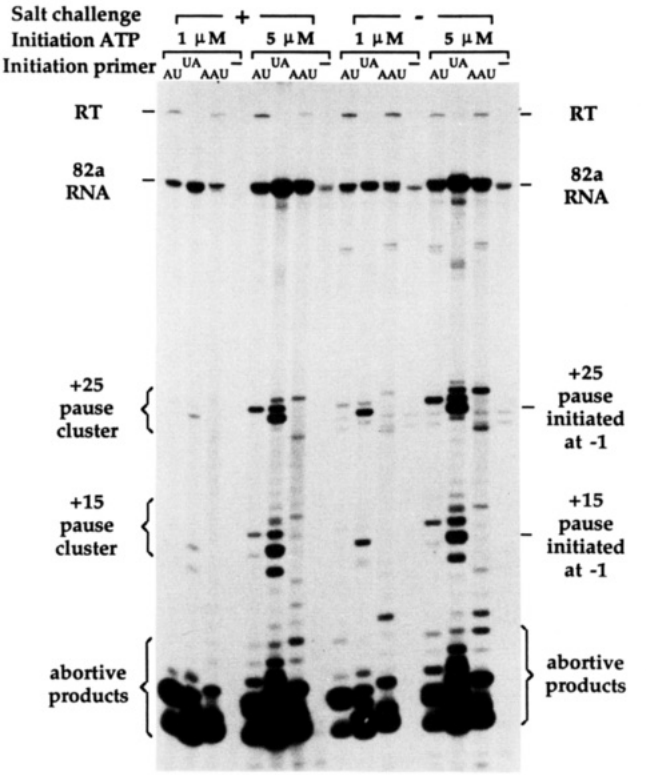

FIGURE 2: Mapping of the transcriptional initiation site of the phage 82 late gene promoter. An autoradiogram of RNAs synthesized from the native phage 82 DNA template is shown. Transcription was initiated either without or with oligoribonucleotide primer ApU, UpA, or ApApU in the presence of 1 or 5 μM ATP, and was stalled at +6 by omitting GTP from the transcription mixture. One set of stalled transcription complexes was made 0.38 M in KCl and incubated 2 min ("salt challenge"), and elongation was continued by chasing with high concentrations of all four ribonucleoside triphosphates and rifampicin. RNAs were sampled 0.25 min after chasing. Abortive transcripts, the +15 and +25 transcriptional pause clusters resulting from heterogeneous initiation (marked on the left), terminated RNAs (82a), readthrough (RT) RNAs, and the +15 and +25 paused transcripts initiated at −1 with primer UpA (marked on the right) are indicated. For both the +25 and +15 pauses, the most intense bands in the UpA-initiated set consist of RNAs starting at −1.

of the middle T (+2) should alter the pattern of transcript heterogeneity. For the experiment of Figure 1, we transcribed a restriction fragment containing either the natural late gene promoter p_{82} with its associated terminator t_{82}, yielding the 80-nucleotide 82a RNA, or an identical fragment carrying a +2C mutation (named according to the RNA-like strand). It is clear that the +2 mutation does not simply change one of the multiple transcripts, but instead completely eliminates the heterogeneity of abortive, paused, and terminated transcripts. Transcript heterogeneity therefore must depend upon nucleotide 2, and possibly upon the existence of three adjacent T's in the template strand.

This experiment also establishes that there are single pause sites at +15 and +25. Assignment of the first pause to +15 is based on an extended pause induced at +14 by limitation of GTP in vitro, and the +25 position is assigned by counting (faint) bands in the natural ladder between the pause sites (H.-C. Guo, Ph.D. Thesis, Cornell University, 1990).

Transcription Initiates at the First of Three Consecutive dT's in the Template Strand. To determine the precise sites of initiation, we compared the natural ATP-initiated transcripts of p_{82} to those selectively initiated with oligoribonucleotide primers (Figure 2). Synthesis was initiated with or without primers at a low and highly limiting concentration of initiator ATP (1 or 5 μM). Radioactive triphosphates were included,

and synthesis was stalled at +6 by omitting GTP from the transcription mixture; elongation was continued by chasing with high concentrations of all four cold ribonucleoside triphosphates and rifampicin to block further initiation. RNAs were sampled shortly (0.25 min) after addition of unlabeled rNTP chase solution (Figure 2). As expected, abortive initiation occurred in the +6 stalled complexes. In the absence of primer, the normal transcripts appeared faintly in reactions not challenged with a salt jump, particularly at +25; this slight background initiation with ATP could be eliminated by challenging the +6 stalled complex with 0.38 M KCl before chasing with all four NTPs, but was useful for determining transcript sizes. The dinucleotide UpA, which should selectively initiate transcription at −1 (Figure 1b), primes synthesis more efficiently than ApU or ApApU. The fact that heterogeneity of paused transcripts is reduced by selective primer initiation, particularly at the lower concentration of ATP (1 μM) (Figure 2; also see below), is consistent with the existence of single pause sites at +15 and +25. We discuss below how slippage could account for the remaining heterogeneity of transcripts initiated with primer UpA at 5 μM ATP.

Transcripts initiated with oligonucleotide primers lack the charged 5′-triphosphate group and, therefore, are expected to migrate more slowly in electrophoresis than natural transcripts of the same number of nucleotides. The major UpA-initiated RNA migrates somewhat slower than the second of the natural +25 paused RNAs (counting upward), so that the second RNA appears to have the same length as an RNA initiated at −1; this is most evident in the rightmost set of reactions in Figure 2. Thus, the shortest transcript in each cluster of bands initiated by ATP begins at the first of three consecutive dT's, the +1 position of the template strand. The primers ApU and ApApU appear not to work at the three consecutive dT's in the template strand; instead, ApU primes inefficiently at −2 and ApApU at −3 and +2. Comparison of transcripts initiated at −3 and +2 with ApApU in the presence of 5 μM ATP shows that initiation at −3 is more sensitive to salt jump challenge than initiation at +2 (most evident at the +25 pause in Figure 2); this is consistent with the fact that one mismatch at −3 is involved in initiation at an upstream site with ApApU. The different activities of these primers at the phage 82 late gene promoter are consistent with the report by Ruetsch and Dennis (1987) that an oligonucleotide containing the preformed third phosphodiester bond does not function as a primer, as long as the +2 and +3 T template bases cannot function as initiation sites.

A Slippage Model To Account for Heterogeneous Transcription Initiation at p_{82}. End labeling of the 82a RNA by initiating transcription with [γ-^{32}P]ATP reveals approximately the same heterogeneity in transcript length as internal labeling (Goliger & Roberts, 1987), indicating that all RNAs start with ATP. Since the shortest transcript in each cluster is initiated from the first of three consecutive dT's in the template strand, +1 (see above), the existence of multiple initiation sites would require that the longer transcripts be initiated upstream of +1. However, the pattern of A residues upstream of +1 does not match the heterogeneity pattern.

Instead, we suggest a model in which heterogeneity arises from slippage of the initial trinucleotide within the open promoter complex (Figure 3). According to this model, RNA polymerase starts transcription at the first of the three consecutive dTs in the template strand (+1). After two phosphodiester bonds are formed, resulting in three rA·dT base pairs in the transcription bubble, the initiation complex can take either of two alternative routes: (i) the initiated transcript

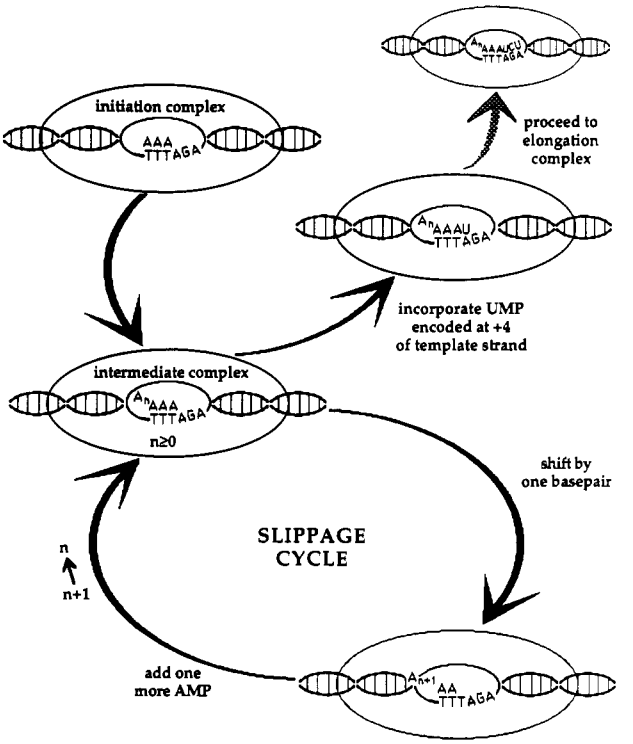

FIGURE 3: Slippage model to explain 5′ heterogeneity of the phage 82 late gene RNA. After two phosphodiester bonds are formed, resulting in three rA·dT base pairs in the transcription bubble, the intermediate complex can take one of two alternative routes: (i) slippage of the initiated transcript one base pair toward the 5′ end, and addition of one more AMP encoded by the third of three consecutive dT's in the template strand, to complete the slippage cycle; or (ii) incorporation of UMP, encoded by position +4 of the template strand, to proceed to the normal elongation complex. The initiation complexes can go through different numbers of slippage cycles before they switch into elongation complexes, resulting in different numbers of A residues at the 5′ end of transcripts.

can slip one base pair toward the 5′ end, so that one more AMP can be templated by the third of three consecutive dT's in the template strand, to complete a slippage cycle; or (ii) the UMP encoded by position +4 of the template strand can be incorporated, to establish the normal elongation sequence. Slippage could also (or instead) occur from the dinucleotide stage, although the failure of a promoter encoding only two A's (see below) to support slippage argues against the possibility. We suggest that initiation complexes can go through numerous slippage cycles before they switch into elongation complexes, resulting in numerous A's at the 5′ ends of transcripts. Slippage would be abolished by preventing the shift of oligo(A) RNAs across the face of the first three template nucleotides, as occurs with the +2 point mutation. Clearly, this postulated ability of an initiation complex to form again after slippage is consistent both with the existence of the abortive initiation reaction and with the activity of exogenous template-pairing oligonucleotides in priming initiation.

Initiation Heterogeneity Depends upon [UTP] and [ATP] in the Reaction Mixture. According to the slippage model (Figure 3), the number of slippage cycles that occur in the initiation complex should depend upon the elongation rate, particularly between template bases 3 and 4, and this rate, in turn, may depend upon the substrate concentration: high [UTP] should favor incorporation of UMP at +4 and thus suppress slippage and heterogeneity, whereas low [UTP] should favor slippage and heterogeneity. Furthermore, high [ATP] might favor the reiterative polymerization of A residues into the transcripts by promoting elongation through the dT

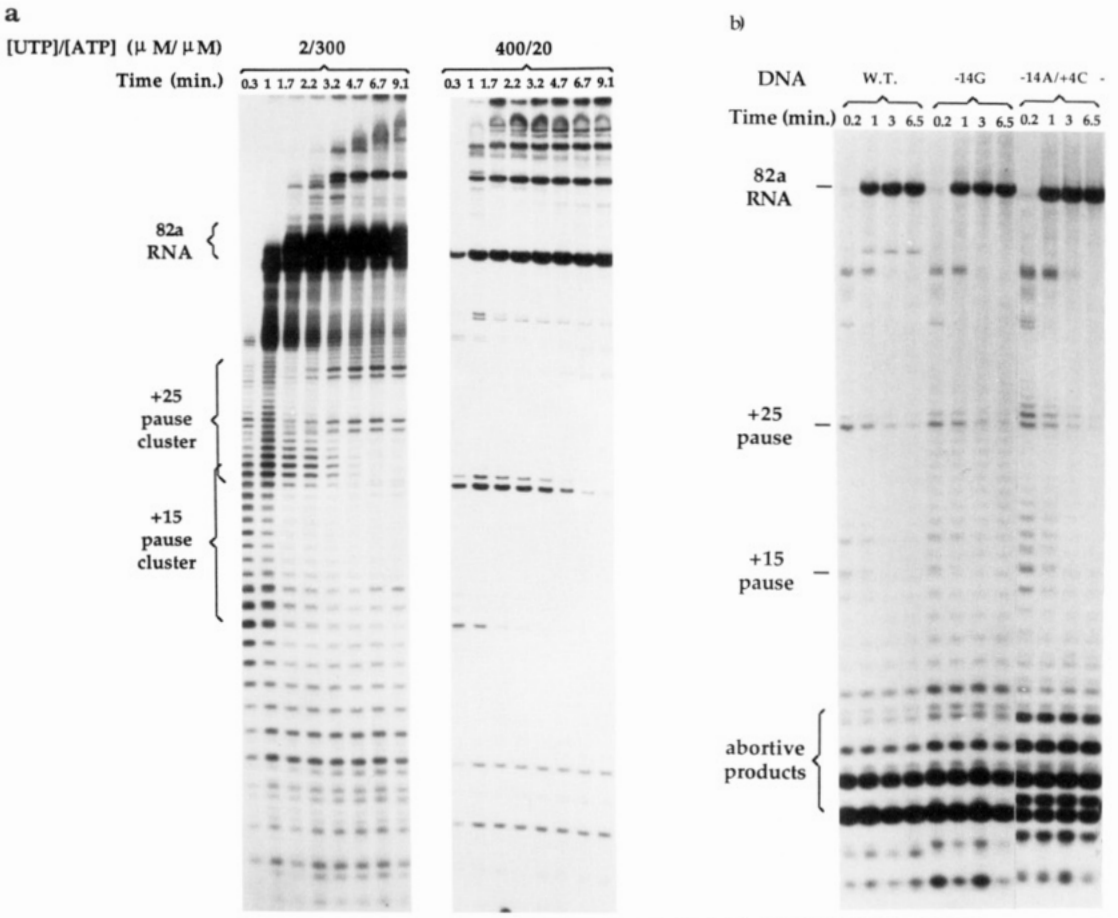

FIGURE 4: Effect of a +4 mutation on the suppression of transcript heterogeneity by high [UTP]/[ATP]. (a) Time course of early paused RNAs transcribed from native phage 82 DNA template using different ratios of [UTP]/[ATP]. Products of the early pauses (+15 and +25) and the terminated RNAs (82a) are indicated. (b) Time course of early paused RNAs transcribed from native phage 82 or its mutant derivative DNA templates using high [UTP]/[ATP] (300 μM/15 μM). Abortive products, products of the early pauses (+15 and +25), and the terminated RNAs (82a) are indicated.

segment and thus reducing abortive loss of the RNA. To test these predictions of the model, we transcribed the natural p_{82}–t_{82} template with different ratios of [UTP] to [ATP] (Figure 4a). Transcripts were sampled at various times up to 9 min and resolved by electrophoresis. As predicted by the model, lower [UTP]/[ATP] produces strikingly heterogeneous transcripts at the +15 and +25 pauses, at the terminator, and at other sites along the course of transcription. In contrast, higher [UTP]/[ATP] suppresses heterogeneity. The suppression is even underestimated by the visualization of Figure 4a, because the reaction with higher [UTP]/[ATP] was labeled with [α-^{32}P]ATP which overlabels RNAs with 5′-poly(A) (see below). We also may see evidence of the substrate effect on the heterogeneity of transcripts initiated with the oligonucleotide UpA (Figure 2): when 5 μM ATP was present during initiation with primer UpA, transcripts appeared more heterogeneous than in the presence of 1 μM ATP. This result is consistent with the slippage model since one unstable rU·dA base pair (Martin & Tinoco, 1980) at −1 might not be strong enough to prevent the transcript from slipping.

A +4 Mutation Abolishes Suppression of Initiation Slippage by High [UTP]/[ATP]. Since the model proposes that suppression of slippage by high [UTP]/[ATP] is due to fixation of nascent transcripts on the template by incorporation of U at position +4, a mutation at +4 should abolish the effect of high [UTP]/[ATP]. Figure 4b shows an autoradiogram of RNAs transcribed from wild-type 82 or mutant DNA templates using high [UTP]/[ATP] (300 μM/15 μM). Heterogeneity was greatly reduced under this condition for the wild type and most mutants (compare panels a and b of Figure 4). However, a −14A/+4C double mutation (the only available source of a +4 change) gave heterogeneous abortive and paused transcripts. This effect is almost certainly not due to the −14 mutation, since another −14 mutation (C to G) does not cause slippage in this condition. Thus, as predicted by the model, a +4C mutation abolishes suppression of slippage by high [UTP]/[ATP].

Direct Evidence for Heterogeneous 5′-Poly(A) on 82a Transcripts. To show directly that transcript heterogeneity represents variable numbers of 5′-adenines, we determined the 5′-terminal nucleotide sequence of longer transcripts. RNAs were synthesized from two independently prepared templates containing p_{82}, using different ratios of [UTP] to [ATP] similar to those of Figure 4a. A ^{32}P-5′-end-labeled oligonucleotide complementary to a region of the t_0 terminator segment in the terminated transcripts was used as primer to make cDNA (about 90 nucleotides long), which was resolved by gel electrophoresis (Figure 5). More than 20 species of cDNA were visible in a darker exposure for the most heterogeneous transcripts (3 μM UTP/300 μM ATP). Because sequencing total cDNA with heterogeneous 5′ ends by either chemical cleavage or dideoxy chain termination would give a complex pattern, we isolated clusters in a narrow size range of the longer cDNAs, prepared as in Figure 5 (lower [UTP]/[ATP]), and sequenced them according to Maxam and Gilbert (1977). Figure 6 shows sequence analyses of cDNAs derived from the two sets of transcripts. Since the uncleaved species with a

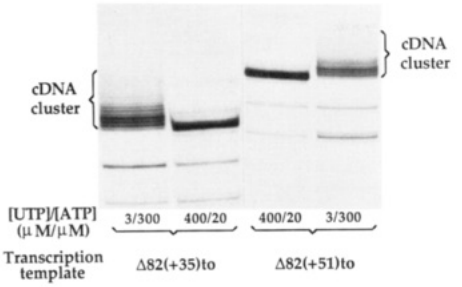

FIGURE 5: 5′ heterogeneity of RNAs transcribed by using different ratios of [UTP] to [ATP] detected by primer extension. RNAs were synthesized from $\Delta 82(+35)t_0$ or $\Delta 82(+51)t_0$ DNA template by using different [UTP]/[ATP] ratios as indicated. These DNAs have the natural sequence through base 35 or base 51 of the 82 late RNA coding segment and are equivalent for the purpose of this experiment. The autoradiogram shows the pattern of cDNAs synthesized from the above RNAs using a ^{32}P-5′-end-labeled primer hybridized to an internal region of the transcripts.

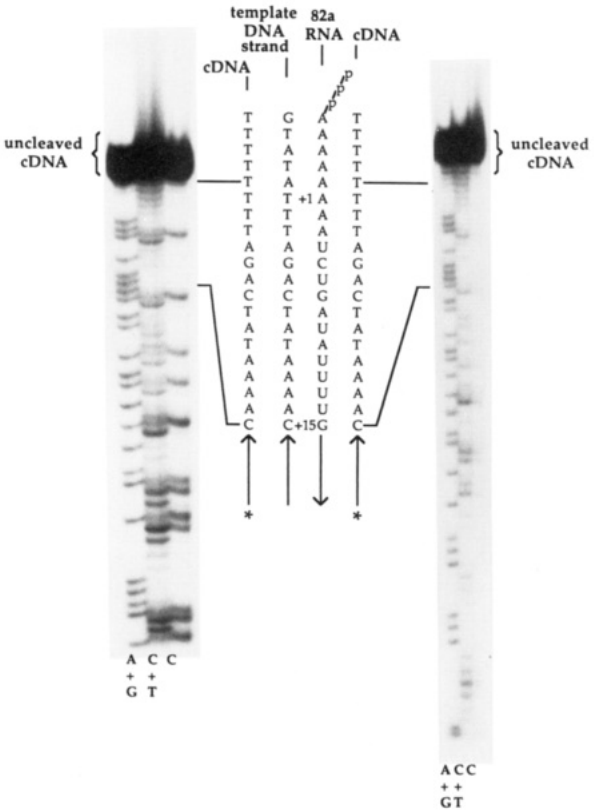

FIGURE 6: DNA sequence of cDNAs of heterogeneous RNAs. The cDNAs ranging from the 4th to the 10th band (counting from the shortest transcript) in the clusters shown in Figure 5 were isolated and sequenced as described. The autoradiograms show independent experiments for cDNAs derived from $\Delta 82(+35)t_0$ (left) or $\Delta 82(+51)t_0$ (right) templates. Also shown are sequences of the template DNA strand, a transcript with five extra A residues at 5′ end, and a cDNA derived from the transcript. The 5′ → 3′ direction is indicated by arrows, and a star denotes the ^{32}P-5′-end-label of the oligonucleotide primer used for cDNA synthesis.

shorter 3′-poly(dT) sequence appear in all sequencing reactions, the 3′-end sequences of heterogeneous cDNA cannot be read to a great length. Nonetheless, it is clear that heterogeneous transcripts contain a homopolymer tract of more than four A residues at the 5′ end that does not match the template DNA sequence, as the model predicts (Figure 3). Furthermore, it can be deduced from the homogeneity of the pattern up to +1 (particularly in the purine track) that there can be no substantial sequence heterogeneity after +1.

Point Mutations in the Nontranscribed Region of the Phage 82 Late Gene Promoter Affect Slippage in Transcription Initiation. The suppression of heterogeneity by UTP above 5 μM implies a significantly high K_M for interaction of UTP with the initiating RNA polymerase to incorporate UMP at position +4 of the template strand. It has been suggested that the binding constant for a particular substrate is determined by the sequence-dependent conformation of the transcription complex, a proposal consistent with the fact that the K_M for initiation is generally higher than that for elongation (McClure et al., 1978). To test the implication that the heterogeneity might be affected by base changes near the active template site, we examined the effects of point mutations in the promoter region on slippage of initiation. Figure 7 shows a transcription time course from wild-type 82 and several mutant DNAs at 80 μM ATP and 5 μM UTP. Comparison of the patterns shows that the −5G mutation enhances slippage during initiation, whereas the −5A and −2G mutations reduce slippage; this result is consistent with the existence of different Michaelis constants for elongation imposed by variant structures of the elongating complex.

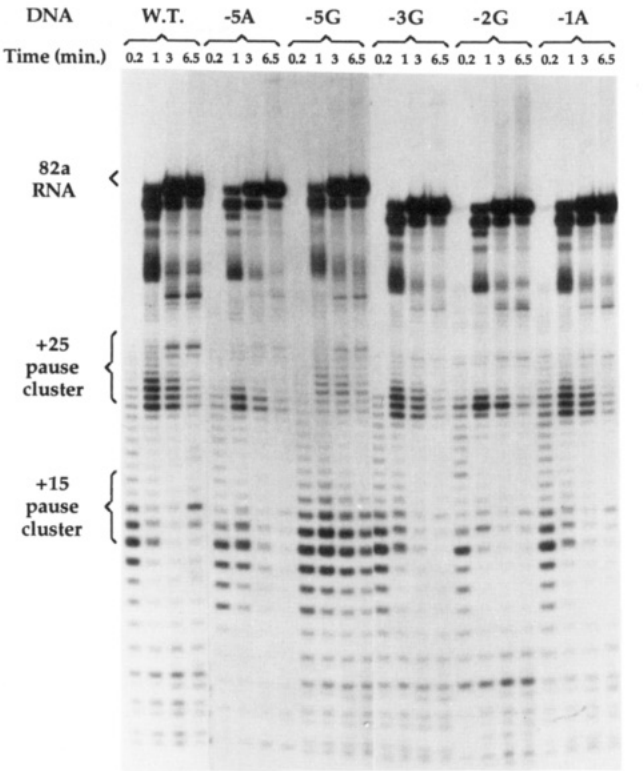

FIGURE 7: Effects of point mutations in the nontranscribed region on slippage during initiation. Autoradiogram of a transcription time course from wild-type phage 82 or its mutant derivative DNA templates. Products of the early pauses (+15 and +25) and the terminated RNAs (82a) are indicated.

Slippage Is Not Promoter-Specific. Although some mutations in the nontranscribed region of the 82 late gene promoter affect the extent of slippage (see above), no available mutation in the nontranscribed region (−16, −15, −14, −13, −5, −3, −2, −1) completely abolishes slippage during initiation. To test if the early transcribed region of the 82a RNA coding sequence alone suffices to support slippage, we transcribed a template in which this segment is fused to the λ late gene promoter $p_{R'}$. The promoter fusion template *qut*λ82 contains sequences of $\lambda p_{R'}$ from −39 to −2 and sequences of $\Delta 82(+17)t_{82}$ to the right of −2; the $\Delta 82(+17)t_{82}$ template, which is a linker substitution of nucleotides 18–25, behaves essentially like wild-type 82 in encoding pauses and supporting initiation slippage. Figure 8 shows that the *qut*λ82 template still allows slippage of initiation to occur. Its transcripts are somewhat

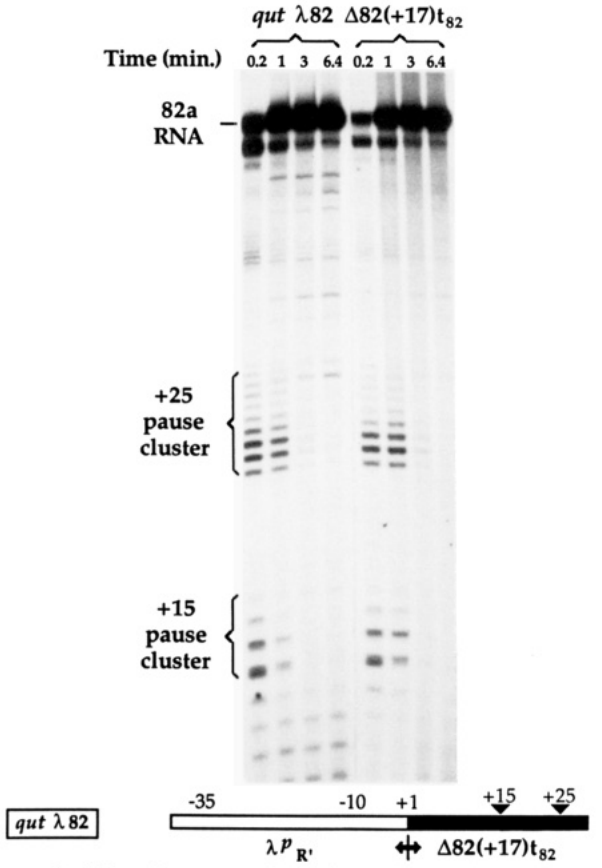

FIGURE 8: Effect of promoter substitution on slippage during initiation. Autoradiogram of a transcription time course from DNA of Δ82-(+17)t_{82}, which is a linker substitution of nucleotides 18–25, and supports slippage during initiation like the wild-type p_{82} segment (Goliger & Roberts, 1989), and the promoter substitution *qut*λ82. The latter DNA contains the Δ82(+17)t_{82} sequences to the right of −2 and λ$p_{R'}$ sequences to the left of −1; its structure is shown at bottom of the figure with the pause sites indicated as downward triangles. The sequence of *qut*λ82 from −35 through −1, with the −35 and −10 hexamers underlined, is TTGACTTATTGAATAAAATTGGG-TAAATTGACTT. Products of the early pauses (+15 and +25) and the terminated RNAs (82a) are indicated.

more broadly dispersed at the +25 pause than those of Δ82-(+17)t_{82}, perhaps because *qut*λ82 has a G at −5 (see above). We conclude that the presence of the early transcribed sequence of 82a RNA at its natural location can confer on a foreign promoter the ability to template a poly(A) leader sequence via the slippage mechanism.

Discussion

We have shown that the sequence TTT in the DNA template strand at the initiation site of the phage 82 late gene promoter supports reiterative copying through slippage of initial oligo(A) RNAs across the face of the template strand (Figure 3). Slippage is prevented by incorporation of UMP encoded at position +4 of the template strand; incorporation of the UMP, in turn, is enhanced by a higher ratio of UTP to ATP in the in vitro transcription reaction. Although the TTT template sequence appears to be sufficient to support slippage, the extent of slippage is slightly affected by the sequences surrounding the initiation site. It is possible that surrounding sequences affect the binding constant for the rNTP substrate by determining the detailed conformation of the transcription complex, since we showed that slippage is sensitive to the concentration of UTP and ATP. Alternatively, the stability of the emergent RNA–DNA duplex, and thus its tendency to dissociate and allow slippage, might be sensitive to the structure of the complex.

A mechanism similar to but not identical with that described here was proposed by Cowie et al. (1982) to explain heterogeneous ends of polyoma virus early RNA. The polyoma mechanism requires multiple base shifts of the oligonucleotide and involves mispairing; although slippage can occur many times (more than 20) for p_{82}, only one base shift each time is required to add another AMP. Furthermore, in contrast to the polyoma mechanism, there are always two nucleotides at the 3′ end of the p_{82} "slipped" transcripts associated with the template sequence at the start site. The slippage of initiation described here is also different from that observed by Chamberlin and Berg (1964) in at least one respect: the natural double-stranded DNA template at the promoter is used instead of single-stranded DNA. Slippage reactions essentially identical with that described here were proposed for a phage T4 promoter by Kassavetis et al. (1986), by Martin et al. (1988) for phage T7 RNA polymerase acting at a phage promoter, by Harley et al. (1990) for a mutant *tet* promoter, and by Jacques and Susskind (1990) for a phage P22 promoter.

The behavior of DNA deletions of the early transcribed segment shows that sequences downstream of +5 of the 82a RNA coding region are not required for slippage during initiation (Goliger & Roberts, 1989). Activity of the promoter fusion template *qut*λ82 shows that sequences upstream of −1 of the phage 82 promoter segment are not required, except to provide an active promoter. Point mutational analysis shows that −1T is not essential for slippage (Figure 7), although a −1C mutation reduces slippage, probably by shifting the initiation site to this position (data not shown). The fact that both +4G (data not shown) and +4C (Figure 4b) mutations still allow slippage to occur establishes that a T residue after the stretch of A's at the initiation site also is not necessary for slippage of initiation. Thus, initiation of RNA polymerase at the first A residue of the TTT sequence in the template strand apparently suffices to allow reiterative copying of the T's. In agreement, the promoter fusion *qut*λ82 template, which contains the λ$p_{R'}$ promoter fused to the 82 AAAU-encoding sequence at the initiation site, supports slippage, whereas the natural λ$p_{R'}$ promoter that encodes AACG does not. It is not known if a stretch of more than three A residues at the initiation site also will allow slippage to occur.

Slippage during initiation may occur at other promoters in the initiation complexes of both prokaryotic and eukaryotic RNA polymerases. 5′ heterogeneity is observed for RNAs initiated at the *deo* promoter P2, which has the sequence AAAC at the initiation site (Valentin-Hansen et al., 1982). A 5′-poly(A) leader sequence different from the template DNA sequence CAAAT (the nucleotide corresponding to the major 5′ end of the transcript is underlined) has been reported for bacteriophage φX174 mRNAs (Grohmann et al., 1975). 5′-Polyadenylation also is observed in synthesis of vaccinia virus late mRNA (Wright & Moss, 1987), for which the majority of mRNA molecules contain a 5′-poly(A) sequence longer than 30 nucleotides. Interestingly, AAAT is present in a conserved sequence (TAAAT) that is necessary for vaccinia virus late gene transcription and is appropriately located to function in slippage like the p_{82} sequence. Wright and Moss (1987) further point out that (i) a linear DNA segment containing only 37 bp of DNA upstream of the TAAAT sequence can synthesize 5′-poly(A) sequence in vitro, (ii) a correctly initiated and 5′-polyadenylated product is synthesized in vitro within 5 min after transcription reactions are initiated, suggesting that the 5′ end of the mRNA is not generated by a processing

mechanism, and (iii) a single base change at +1 that converts the TAAAT sequence of the vaccinia virus late gene promoter to TGAAT abolishes transcription. All of these observations are consistent with the slippage model of transcription we describe (Figure 3). The observation that prokaryotic and eukaryotic RNA polymerases have homologous core subunits (Sweetser et al., 1987) suggests that the slippage mechanism might be similar in transcription initiation complexes of organisms of all levels.

ACKNOWLEDGMENTS

We thank members of the laboratory for advice and helpful discussions and John Lis for criticizing the manuscript.

Registry No. ATP, 56-65-5; UTP, 63-39-8; RNA polymerase, 9014-24-8.

REFERENCES

Burgess, R. R., & Jendrisak, J. J. (1975) *Biochemistry 14*, 4634–4638.

Carpousis, A. J., & Gralla, J. D. (1980) *Biochemistry 19*, 3245–3253.

Carpousis, A. J., & Gralla, J. D. (1985) *J. Mol. Biol. 183*, 165–177.

Carpousis, A. J., Stefano, J. E., & Gralla, J. D. (1982) *J. Mol. Biol. 157*, 619–633.

Chamberlin, M. J., & Berg, P. (1964) *J. Mol. Biol. 8*, 708–726.

Cowie, A., Jat, P., & Kamen, R. (1982) *J. Mol. Biol. 159*, 225–255.

Falashi, A., Adler, J., & Khorana, H. G. (1963) *J. Biol. Chem. 238*, 3080–3085.

Fox, C. F., & Weiss, S. B. (1964) *J. Biol. Chem. 239*, 175–185.

Goliger, J. A., & Roberts, J. W. (1987) *J. Biol. Chem. 262*, 11721–11725.

Goliger, J. A., & Roberts, J. W. (1989) *J. Mol. Biol. 210*, 461–471.

Grayhack, E. J., & Roberts, J. W. (1982) *Cell 30*, 637–648.

Grohmann, K., Smith, L. H., & Sinsheimer, R. L. (1975) *Biochemistry 14*, 1951–1955.

Hansen, U. M., & McClure, W. R. (1980) *J. Biol. Chem. 255*, 9564–9570.

Harley, C. B., Lawrie, J., Boyer, H. W., & Hedgpeth, J. (1990) *Nucleic Acids Res. 18*, 547–552.

Jacques, J.-P., & Susskind, M. M. (1990) *Genes Dev.* (in press).

Johnston, D. E., & McClure, W. R. (1976) in *RNA Polymerase* (Losick, R., & Chamberlin, M., Eds.) pp 413–428, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Kassavetis, G. A., Zentner, P. G., & Geiduschek, E. P. (1986) *J. Biol. Chem. 261*, 14256–14265.

Krakow, J. S., & Karstadt, M. (1967) *Proc. Natl. Acad. Sci. U.S.A. 58*, 2094–2101.

Krummel, B., & Chamberlin, M. J. (1989) *Biochemistry 28*, 7829–7842.

Kunkel, T. A., Roberts, J. D., & Zakour, R. A. (1987) *Methods Enzymol. 154*, 367–382.

Levin, J. R., Krummel, B., & Chamberlin, M. J. (1987) *J. Mol. Biol. 196*, 85–100.

Lowe, P. A., Hager, D. A., & Burgess, R. R. (1979) *Biochemistry 18*, 1344–1352.

Luse, D. S., & Jacob, G. A. (1987) *J. Biol. Chem. 262*, 14990–14997.

Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Martin, C. T., Muller, D. K., & Coleman, J. E. (1988) *Biochemistry 27*, 3966–3974.

Martin, F. H., & Tinoco, I., Jr. (1980) *Nucleic Acids Res. 8*, 2295–2299.

Maxam, A. M., & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 560–565.

McClure, W. R. (1985) *Annu. Rev. Biochem. 54*, 171–204.

McClure, W. R., Cech, C. L., & Johnston, D. E. (1978) *J. Biol. Chem. 253*, 8941–8948.

Reznikoff, W. S., Maquat, L. E., Munson, L. M., Johnson, R. C., & Mandecki, W. (1982) in *Promoters: Structure and Function* (Rodriguez, R. L., & Chamberlin, M. J., Eds.) pp 80–95, Praeger, New York.

Ruetsch, N., & Dennis, D. (1987) *J. Biol. Chem. 262*, 1674–1679.

Spassky, A. (1986) *J. Mol. Biol. 188*, 99–103.

Stevens, A. (1964) *J. Biol. Chem. 239*, 204–209.

Straney, D. C., & Crothers, D. M. (1985) *Cell 43*, 449–459.

Sweetser, D., Nonet, M., & Young, R. A. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 1192–1196.

Valentin-Hansen, P., Aiba, H., & Schumperli, D. (1982) *EMBO J. 1*, 317–322.

von Hippel, P. H., Bear, D. G., Morgan, W. D., & McSwiggen, J. A. (1984) *Annu. Rev. Biochem. 53*, 389–446.

Wright, C. F., & Moss, B. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 8883–8887.

Yamakawa, M., Furuichi, Y., Nakashima, K., LaFiandra, A. J., & Shatkin, A. J. (1981) *J. Biol. Chem. 256*, 6507–6514.

Yang, X., Goliger, J. A., & Roberts, J. W. (1989) *J. Mol. Biol. 210*, 453–460.