Molecular Cell *24*, 257–266, October 20, 2006 ©2006 Elsevier Inc. DOI 10.1016/j.molcel.2006.10.001

Template Misalignment in Multisubunit RNA Polymerases and Transcription Fidelity

Ekaterina Kashkina,[1] Michael Anikin,[1] Florian Brueckner,[2] Richard T. Pomerantz,[3,4,5] William T. McAllister,[1] Patrick Cramer,[2] and Dmitry Temiakov[1,*]

[1] Department of Cell Biology
School of Osteopathic Medicine
University of Medicine and Dentistry of New Jersey
42 East Laurel Road
Stratford, New Jersey 08084

[2] Department of Chemistry and Biochemistry
Gene Center Munich
Ludwig-Maximilians-Universität München
Feodor-Lynen-Strasse 25
81377 Munich
Germany

[3] Department of Microbiology and Immunology and
[4] Graduate Program in Molecular and Cellular Biology
The State University of New York Downstate Medical Center
450 Clarkson Avenue
Brooklyn, New York 11203

Summary

Recent work showed that the single-subunit T7 RNA polymerase (RNAP) can generate misincorporation errors by a mechanism that involves misalignment of the DNA template strand. Here, we show that the same mechanism can produce errors during transcription by the multisubunit yeast RNAP II and bacterial RNAPs. Fluorescence spectroscopy reveals a reorganization of the template strand during this process, and molecular modeling suggests an open space above the polymerase active site that could accommodate a misaligned base. Substrate competition assays indicate that template misalignment, not misincorporation, is the preferred mechanism for substitution errors by cellular RNAPs. Misalignment could account for data previously taken as evidence for additional NTP binding sites downstream of the active site. Analysis of the effects of different template topologies on misincorporation indicates that the duplex DNA immediately downstream of the active site plays an important role in transcription fidelity.

Introduction

RNA polymerase (RNAP) is a central enzyme in gene expression in all organisms. In both prokaryotes and eukaryotes, RNAPs are multisubunit enzymes that exhibit high sequence and structural homology from species to species. Each cycle of transcription by RNAP involves three phases: promoter binding during initiation, processive synthesis of RNA during elongation, and release of the transcript and dissociation of RNAP during termination. During elongation, incorporation of a single-substrate nucleoside triphosphate (NTP) into the transcript is repeated many times. In the so-called nucleotide addition cycle, extension of RNA involves two sites in the RNAP active center (Steitz, 1998). At the beginning of the cycle, the 3′ end of the RNA occupies the product site and the incoming nucleotide is bound in the substrate (insertion) site. Formation of the phosphodiester bond is followed by movement of the newly formed 3′ end of the RNA into the product site (translocation) and release of pyrophosphate (Landick, 2004).

Whereas most enzymes discriminate only between a single cognate substrate and its low-affinity derivatives, RNAPs use a DNA template to select among four similar yet distinct substrates. Based mainly on studies of DNA polymerases, the fidelity of polymerases is generally thought to rely on the induced fit of the polymerase active site to the base pair formed between the template base and the incoming NTP that occupies the insertion site. Binding of the cognate, complementary NTP provides the correct geometry required for catalysis (Doublie and Ellenberger, 1998; Johnson et al., 2003), whereas a noncomplementary (mismatched) base does not lead to correct geometry, and results in a decreased efficiency of phosphodiester bond formation (Johnson and Beese, 2004).

While the induced-fit mechanism addresses substrate recognition and fidelity after the NTP has entered the active site, recent studies with T7 RNAP indicated that there is an earlier level of substrate discrimination that occurs in a "preinsertion site" prior to delivery of the incoming NTP into the substrate (insertion) site (Temiakov et al., 2004). It has been suggested that a similar mechanism of fidelity in multisubunit RNAPs may also involve template-dependent recognition of the incoming NTP (Holmes et al., 2006; Kettenberger et al., 2004; Svetlov et al., 2004; Temiakov et al., 2004). Other studies have suggested that additional substrate binding sites downstream of the active site may also contribute to the fidelity of multisubunit RNAPs (Gong et al., 2005; Nedialkov et al., 2003), but structural evidence to support this model is thus far lacking (Landick, 2005).

The fidelity of transcription may be further improved by proofreading mechanisms that act subsequent to the formation of the phosphodiester bond. A number of studies have indicated that RNAPs may carry out proofreading by means of an intrinsic cleavage activity or as a result of binding of auxiliary factors such as GreA or TFIIS (Erie et al., 1993; Shaw et al., 2002; Thomas et al., 1998). The average misincorporation frequency for multisubunit RNAPs (1 in 10^5) and the effectiveness of selection of ribonucleotide substrates over deoxyribonucleotides are similar to those of the single-subunit T7 RNAP (Blank et al., 1986; Huang et al., 2000; Ninio, 1991; Ozoline et al., 1980; Remington et al., 1998; Rosenberger and Hilton, 1983; Shaw et al., 2002; Springgate and Loeb, 1975; Svetlov et al., 2004). The principal mechanisms that determine fidelity may therefore be similar in cellular and single-subunit RNAPs.

*Correspondence: d.temiakov@umdnj.edu
[5] Present address: Laboratory of DNA Replication, Howard Hughes Medical Institute, The Rockefeller University, 1230 York Avenue, New York, New York 10021.

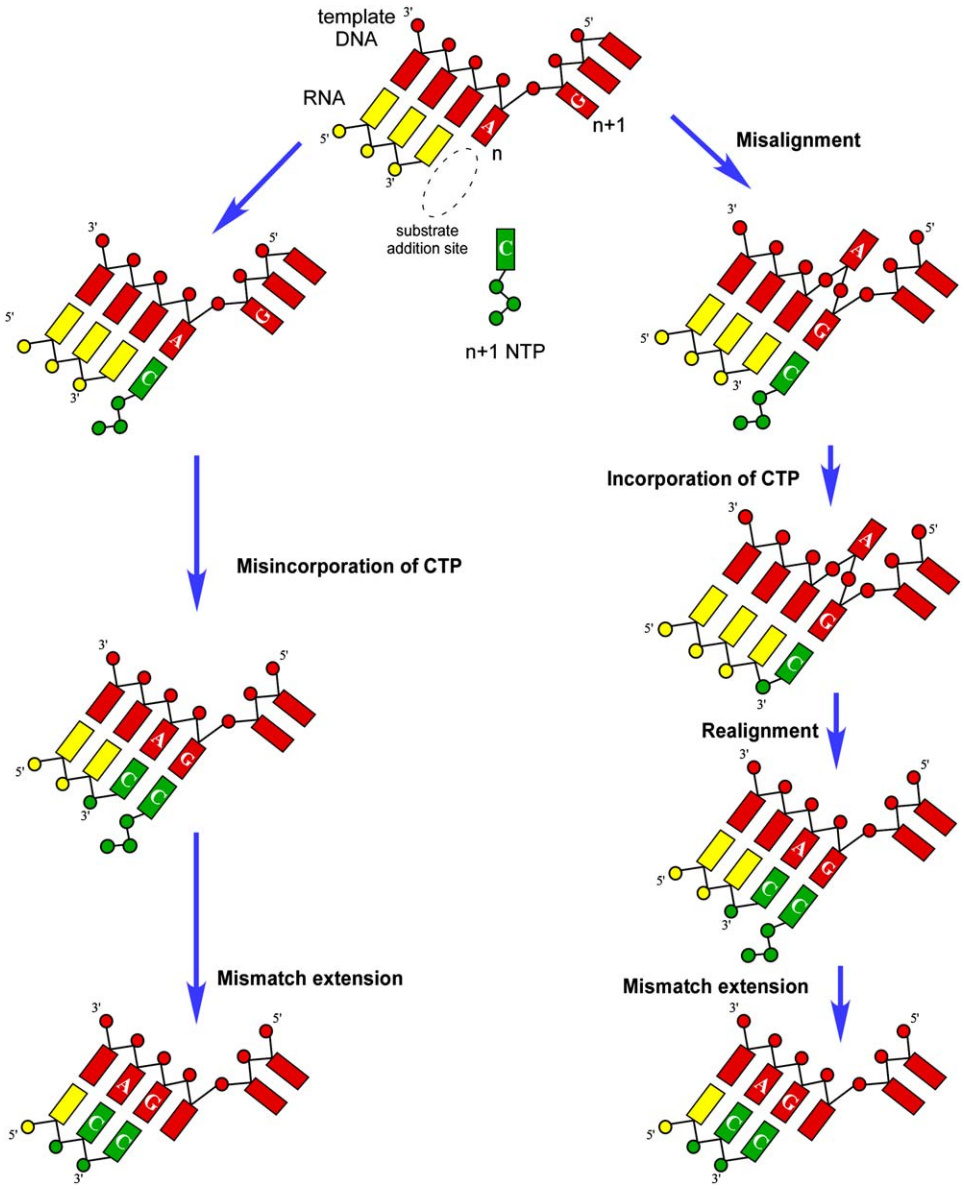

Figure 1. Possible Pathways of n+1 NTP Misincorporation in Multisubunit RNAPs

The template base that is paired with the 3′ end of the RNA is designated as n-1, and the template base in the substrate site as n; template bases further downstream are designated n+1, n+2, etc. The noncognate substrate complementary to the n+1 base is termed the n+1 NTP. Note that in the alternative nomenclature (Landick, 2005), the n-1 and n bases are designated i and i+1, correspondingly. In the misincorporation model (left), the n+1 NTP (CTP) is misincorporated opposite the n base, resulting in a mismatch at the 3′ terminus of the RNA; this is followed by mismatch extension by correct incorporation of CTP opposite the n+1 base. In the misalignment model (right), misalignment of the T strand allows incorporation of CTP opposite the n+1 base; this is followed by realignment of the primer/template and mismatch extension as above. In both cases, the result is a substitution error in the RNA product.

A recent study with T7 RNAP demonstrated a novel mechanism of error production due to template (T)-strand misalignment (Pomerantz et al., 2006 [this issue of *Molecular Cell*]). RNA primer-extension studies showed that a noncognate NTP that is complementary to the n+1 template DNA base (the n+1 NTP) is readily incorporated by T7 RNAP, resulting in an extension of the RNA by two nucleotides (nt). This involves a temporary flipping-out of the n DNA base and a misalignment of the T strand, enabling the n+1 DNA base to pair with the NTP and direct its incorporation (Figure 1, right). The T strand then realigns, resulting in a mismatched RNA end that is further extended.

Here, we show that a similar misalignment mechanism is involved in misincorporation errors by cellular, multisubunit RNAPs from both prokaryotes (*E. coli* and *T. thermophilus*) and eukaryotes (*S. cerevisiae*). These findings provide new insights into understanding the fidelity of transcription and have implications for mechanistic models of substrate selection and incorporation. Our data also reveal a critical role of the nontemplate (NT) strand in the downstream DNA in RNAP fidelity.

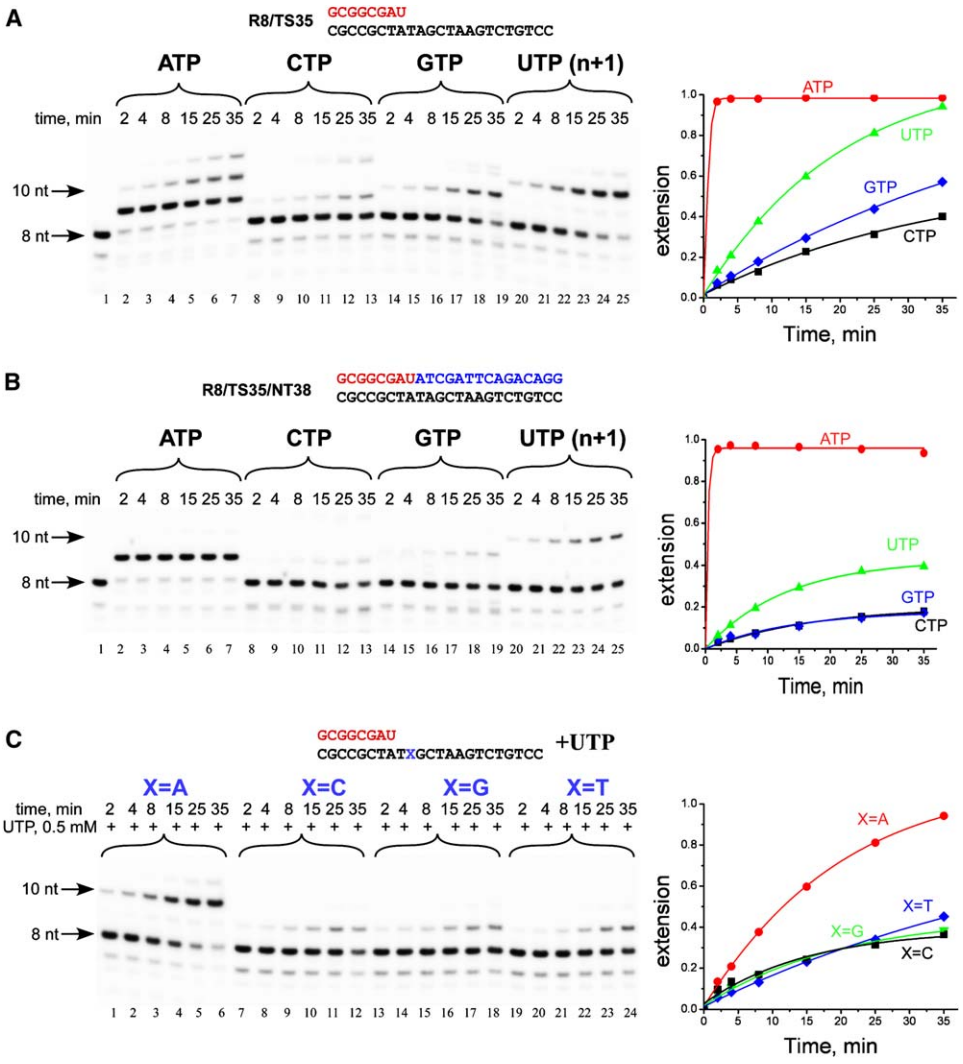

Figure 2. The n+1 NTP Is a Preferable Substrate during Misincorporation by Bacterial RNAP

(A and B) Misincorporation of noncognate NTPs by *E. coli* RNAP. ECs were assembled on scaffolds having single-stranded (A) or double-stranded downstream DNA (B) and 5′ radiolabeled RNA primer as described in Experimental Procedures. ECs were incubated in the presence of 0.5 mM substrate NTPs at 37°C for the times indicated, and the products of the reactions were resolved using 20% PAGE containing 6 M urea and quantified by PhosphorImager. The fraction of the primer extended is plotted as a function of time. Each data point represents an average of at least three independent experiments.

(C) Substrate misincorporation in different sequence context. ECs were incubated with 0.5 mM UTP for the time indicated, and analyzed as above.

Results and Discussion

Misincorporation by Template Misalignment

To monitor nucleotide misincorporation, we utilized a primer-extension assay in which elongation complexes (ECs) were assembled on nucleic acid scaffolds that contained an 8 to 9 nt RNA primer annealed to a DNA T strand that extended 14 to 15 nt downstream. Some scaffolds additionally contained a complementary NT strand downstream of the primer (Figures 2A and 2B).

In the presence of the correct NTP, the 8 nt RNA primer was rapidly extended to 9 nt. A small amount of products >9 nt was also observed during extended incubation, presumably as a result of misincorporation (Figure 2A, lanes 2–7). Extension of the primer in the presence of noncognate NTPs was less efficient. However, the rate of incorporation of UTP, which is the substrate that is complementary to the n+1 base on this template (the n+1 NTP) was notably higher and the primer was extended by 2 nt, rather than by 1 nt, as observed for other noncognate substrates and for the correct substrate (Figure 2A, lanes 20–25). A similar pattern of preferential extension of the primer by 2 nt in the presence of the n+1 NTP was previously observed for T7 RNAP (Pomerantz et al., 2006) and was attributed to a T strand misalignment mechanism (see Figure 1). Preferential misincorporation of the n+1 NTP and extension by 2 nt were also observed when ECs were assembled on a scaffold that contained a complementary NT strand downstream of the RNA primer (Figure 2B). In the latter case, however, the efficiency of extension was lower than in the

absence of the NT strand. Nevertheless, the extension of the primer with the n+1 NTP was still greater than extension due to misincorporation of other noncognate substrates (Figure 2B).

Although the preferential extension of the RNA in the presence of the n+1 UTP shown in Figures 2A and 2B may be accounted for by the misalignment mechanism (Figure 1, right), an alternative explanation is that UTP may be preferentially misincorporated in the particular sequence context utilized in this experiment. To determine whether the preferential incorporation of the n+1 NTP is a general phenomenon, we utilized templates in which each of the four possible bases was present at the n+1 position in the T strand. In each case, the primer was preferentially extended in the presence of the n+1 NTP, and always by 2 nt (see Figure S1A in the Supplemental Data available with this article online).

The misalignment mechanism (Figure 1, right) proposes that preferential incorporation of the n+1 NTP occurs as the result of forming a base pair with the n+1 DNA base in the T strand. We examined this by using a series of templates in which each possible base was present at the n+1 position but the template base at position n remained the same (Figure 2C). Preferential extension of the primer in the presence of UTP was observed only when the complementary base (adenine) was in the n+1 position (Figure 2C, lanes 1–6). This demonstrates that the multisubunit bacterial RNAPs carry out misincorporation via a T strand misalignment mechanism similar to that observed for T7 RNAP. Importantly, we observed an identical pattern of misincorporation on these templates when using yeast RNAP II, suggesting that the misalignment mechanism is universal for all DNA-dependent RNAPs (Figure S1B).

In the presence of the n+1 NTP, the RNA primer was always extended by 2 nt, with little accumulation of an intermediate product extended by only 1 nt. This is consistent with the notion that the rate-limiting step is the initial incorporation event based on template misalignment, and that extension of the RNA having a mismatch 3′ end is fast. We directly examined extension of mismatched RNA primer as shown in Figure S2 and observed that bacterial RNAP from *T. thermophilus* and *E. coli* as well as yeast RNAP II were able to readily extend the mismatched primer.

Template Misalignment Involves an Extrahelical Base

The results of the above experiments indicate that misincorporation of the n+1 NTP in multisubunit RNAPs occurs by a misalignment mechanism (Figure 1, right). In DNAPs, template misalignment requires that one base in the DNA template occupies an extrahelical "flipped-out" position (Bebenek and Kunkel, 1990; Garcia-Diaz et al., 2006; Kobayashi et al., 2002). To verify that template misalignment in bacterial RNAP also involves a similar rearrangement, and to provide direct evidence for the misalignment mechanism, we monitored the chemical environment of the n base in the T strand by fluorescence spectroscopy. We assembled ECs on a scaffold in which the template base at position n (cytidine) was substituted by its fluorescent analog pyrrolocytidine (pyrrolo-C) (Figure 3). Importantly, the fluorescence quantum yield of pyrrolo-C is very sensitive to the local chemical environment, in particular to the relative position of the neighboring bases (Martin et al., 2003). Thus, the quantum yield is low when the base is within an undistorted DNA duplex, but is high when the DNA is melted and the bases are unstacked (Liu and Martin, 2002; Martin et al., 2003). An even higher quantum yield is expected when the base is extrahelical (Kobayashi et al., 2002).

An *E. coli* RNAP EC formed with a scaffold containing pyrrolo-C showed a strong increase in fluorescence emission as opposed to the scaffold alone (data not shown), indicating that the n template base is not involved in base-stacking interactions in the EC, consistent with similar studies of the n template base in DNAP complexes (Kobayashi et al., 2002). Addition of the cognate substrate (GTP) to the pyrrolo-C-containing EC reduced the fluorescence quantum yield. This quenching is apparently due to incorporation of the GMP into the transcript and stacking of the resulting pyrrolo-C-containing nascent base pair within the RNA:DNA hybrid duplex (Figure 3). To monitor changes in the template base position due to substrate binding, we used nonhydrolyzable NTP analogs that cannot be incorporated into the RNA (Figures 3A and 3B). The use of natural NTPs in these experiments was not possible since bacterial RNAP and yeast RNAP II can extend mismatched RNAs (Figure S2), preventing the formation of defined halted complexes under these conditions. Addition of the n+1 AMPcPP increased fluorescence emission, as predicted from a misalignment mechanism that involves flipping-out of the pyrrollo-C to adopt an extrahelical position. In contrast, addition of UTP or UMPcPP, which can base pair with neither the n nor the n+1 template bases, did not affect the quantum yield of pyrrolo-C, providing a negative control (Figure 3). Similar results were obtained with yeast RNAP II (data not shown).

These observations are consistent with findings for DNA polymerase IV, where misalignment resulted in increased fluorescence emission in the presence of the n+1 dNTP (Kobayashi et al., 2002). We conclude that the n template base apparently adopts an extrahelical position during T-strand misalignment, allowing the n+1 base to occupy the site required to direct cognate NTP incorporation.

Structural Considerations

Inspection of the crystal structure of an RNAP II EC (Kettenberger et al., 2004) suggests that there would be space sufficient to accommodate a flipped-out base during misalignment (Figure 4). In considering a possible structure of the misaligned intermediate, we performed modeling using the DNA primer-template duplex with a flipped-out base in the misaligned T strand that was observed in the structure of DNA polymerase λ (PDB code 2BCQ [Garcia-Diaz et al., 2006]). Superposition of this duplex onto the RNA:DNA hybrid duplex in the RNAP II EC in various registers shows that an extrahelical base can be accommodated above the RNAP II protein region switch 2 (Cramer et al., 2001) at a position between registers –1 and +1. The modeling thus suggests that an extrahelical T base can be accommodated in the active center cleft of yeast RNAP II above the active site. The modeling, however, has limitations, since superposition of the B form DNA-DNA duplex onto the

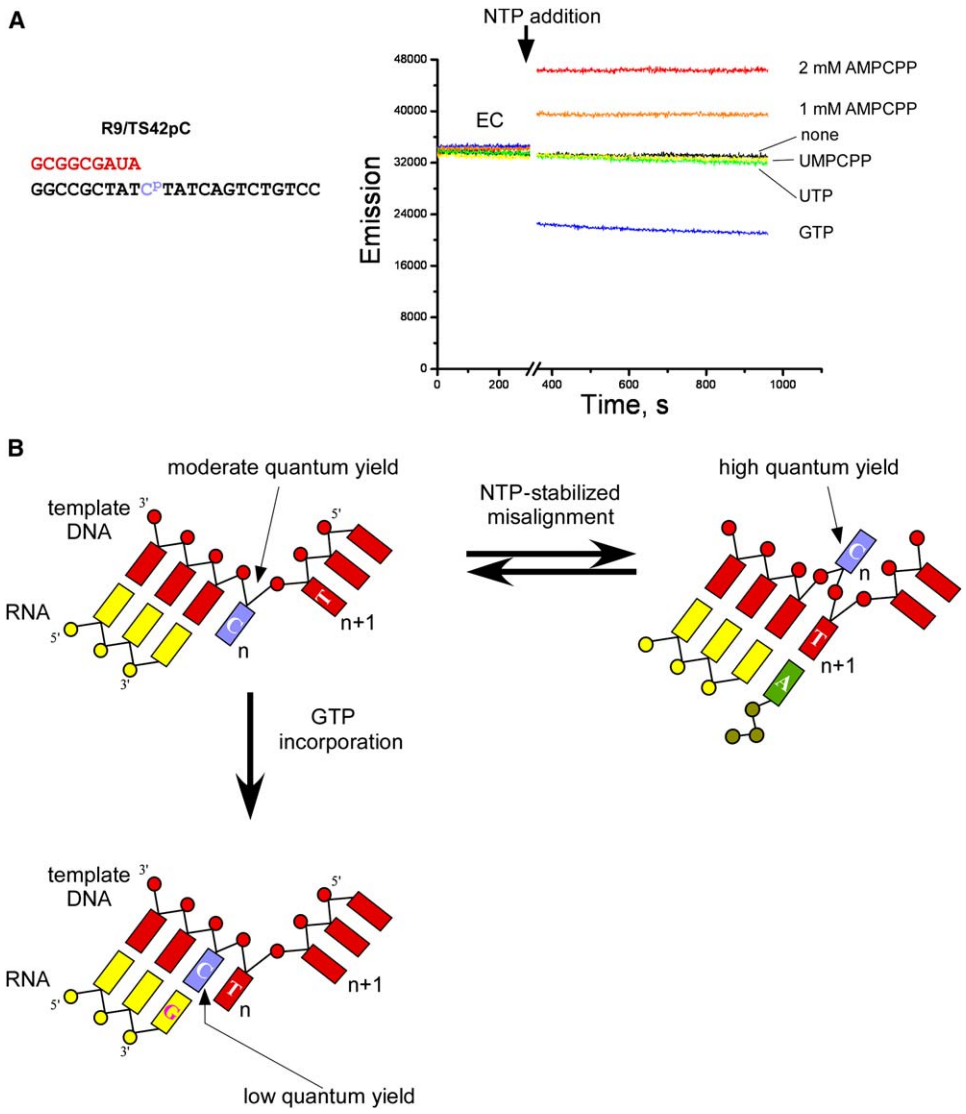

Figure 3. The n Template Base Adopts an Extrahelical Position in the Presence of the n+1 NTP

(A) Fluorescence emission of pyrrolo-C at the n position in the presence of different substrate NTPs. Fluorescent probe, pyrrolo-C, was placed in the T strand at the n position in R9/TS42pC scaffold. *E. coli* core RNAP was added to the scaffold in equimolar concentration (200 nM) and incubated for 10 min at room temperature. Fluorescence emission of the assembled ECs was measured as a function of time with a 1 s interval following addition of cognate (GTP) or noncognate substrates (UTP, UMPcPP, and AMPcPP).

(B) Schematic model of the T-strand misalignment.

near A form RNA:DNA hybrid can only be approximate. Therefore, the exact position of an extrahelical base remains unclear.

The n and n+1 Bases Likely Compete for the Same or Overlapping Binding Sites during Misalignment

The above model predicts that the flipping-out of the template base and template misalignment can be suppressed in the presence of cognate, nonreactive NTP analogs, which should stabilize the normal T-strand conformation (Figure 5). We found that preincubation of the bacterial EC with the cognate GTP analog GMPcPP inhibits misalignment and misincorporation of the n+1 NTP (Figure 5A, lanes 26–31, C), as well as misincorporation of CTP (Figure 5A, lanes 32–37). To verify this effect, we used a scaffold with a different downstream DNA sequence (Figure 5D) and two nonreactive NTP analogs, AMPcPP (cognate) and GMPcPP (noncognate), and monitored misalignment with UTP (n+1 NTP). Again, the presence of the correct nonhydrolyzable NTP efficiently inhibited misalignment, while the presence of an incorrect NTP did not (Figure 5D). Based on these results and the modeling, we speculate that both the n template base in a substrate-bound EC and the n+1 base in an EC with a misaligned template may bind to overlapping sites. This implies that the same NTP binding site could be utilized by RNAP during both normal incorporation and misincorporation due to misalignment.

Misalignment versus NTP Line-Up

In previous work, stimulation of substrate incorporation in the presence of the n+1 NTP was taken as evidence for an additional NTP binding site that does not overlap

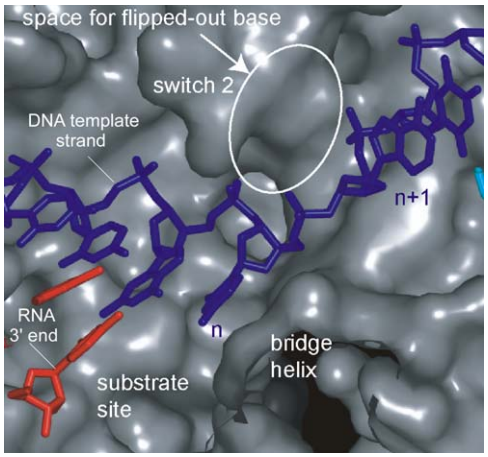

Figure 4. Close-Up of the Active Site of the Yeast RNAP II EC

The protein environment (surface representation, gray) near the n and n+1 bases of the template DNA strand (blue) and the 3′ end of the RNA (red) is shown. The presumed site to accommodate the extrahelical template base during misalignment is indicated.

with the normal substrate insertion site opposite the n base (Nedialkov et al., 2003). This model was extended to suggest that of simultaneous binding of several NTPs to RNAP, leading to a line-up of NTPs downstream before the active site (Gong et al., 2005). We examined whether binding of the n+1 NTP has an effect on the misincorporation rate in bacterial EC by preincubating an EC with AMPcPP, the substrate that is complementary to the n+1 template base (Figures 5A and 5B). We observed that the rate of CTP misincorporation was not increased in the presence of AMPcPP, showing that the presence of the n+1 NTP does not stimulate misincorporation at the n position (Figure 5A, lanes 8–19). When CTP and ATP were used, efficient ATP misincorporation by misalignment was still observed, resulting in extension by AA, whereas little or no extension of the 9-mer RNA primer by C or CA was detected (compare lanes 8–13 and 20–25, Figure 5A). Failure to detect primer extension by CMP (or by CMP and AMP) is most likely due to the much higher rate of misincorporation of the n+1 ATP by misalignment. Alternatively, template rearrangements during misalignment may result in suppression of misincorporation of another noncognate NTP at the n position; however, we did not detect a significant difference in suppression in the experiments involving nonhydrolyzable n+1 NTP analog (Figure 5A, lanes 14–19). A similar effect of the n+1 NTP (diminished accumulation of the products of misincorporation of another noncognate NTP) previously observed in experiments with human RNAP II was interpreted in support of an "NTP line-up mechanism" (Gong et al., 2005). However, our data suggest that template misalignment may also have occurred in this experiment, and lead us to question the NTP line-up model. As our findings suggest that addition of the n+1 NTP to the halted EC may stabilize an alternative, flipped-out conformation of the n base in the RNAP active site and thus affect cognate substrate binding and incorporation, it may be appropriate to review the results of prior experiments that involve the incoming n+1 substrate in light of the misalignment mechanism.

Implications for Overlapping Substrate Sites in RNAPs

It has been suggested that incoming NTP substrates first bind to an entry site (E site) in multisubunit RNAPs, before their delivery to the incorporation site (Westover et al., 2004). The E site lacks direct interactions with the base and ribose moiety of the substrate, and an NTP in the E site does not base pair with a DNA template base (Westover et al., 2004). Our observation that only a cognate NTP can prevent misalignment (Figure 5D) suggests that the E site is not involved in incorporation during template misalignment. This, however, does not exclude the possibility of NTP entry via the E site during normal incorporation. During template misalignment, the NTP is base-paired with the n+1 base of the T strand, and may bind either to the insertion site or to an overlapping, preinsertion-like site (Kettenberger et al., 2004). Since, in both cases, the NTP could form a Watson-Crick base pair with the misaligned template base, our data do not allow us to distinguish between these possibilities.

The Downstream DNA-Binding Site Contributes to Transcription Fidelity

The rates of primer extension by 1 nt due to misincorporation, and extension by 2 nt due to misalignment, were both reduced in the presence of a complementary NT strand immediately downstream of the active site (Figures 2A and 2B), indicating the importance of the NT strand in the fidelity of transcription. Nevertheless, misincorporation of the n+1 NTP by misalignment in such complexes was always greater than misincorporation in the presence of other noncognate NTPs, suggesting that misalignment may be the predominant mechanism by which substitution errors are made by cellular RNAPs.

To add further evidence that misalignment is a commonly occurring phenomenon during transcription by multisubunit RNAPs, we examined the effects of changes in the topology of the nucleic acid scaffold on misincorporation. First, with *E. coli* RNAP we compared misincorporation due to misalignment in scaffolds that either have or lack a complementary downstream NT strand, or have an abasic site in the NT strand at position n+1 (Figure 6A). As before, we found that n+1 CTP misincorporation by misalignment was decreased when the NT strand was present, but the lack of a complementary n+1 base in the NT strand restored efficient misalignment (Figure 6A).

We next examined whether misalignment occurs in ECs halted downstream of a promoter. Three templates containing the T7A1 promoter were used: a fully double-stranded template (Figure 6B, left), a template with a gap in the NT strand commencing 11 nt downstream of the start site (middle), and a template with an abasic site in the NT strand 11 nt downstream of the start site (right). On these templates, *Tth* RNAP forms a halted EC 9 nt downstream of the start site in the presence of ATP, GTP, and CTP. Upon removal of NTPs by gel filtration, the complexes were incubated with the n+1 NTP (ATP). Efficient misalignment was observed when n+1 template DNA base was not base-paired, in the case of templates having either a gap or an abasic site (Figure 6B). The observed misalignment in promoter-originated ECs on templates with abasic sites indicates that this

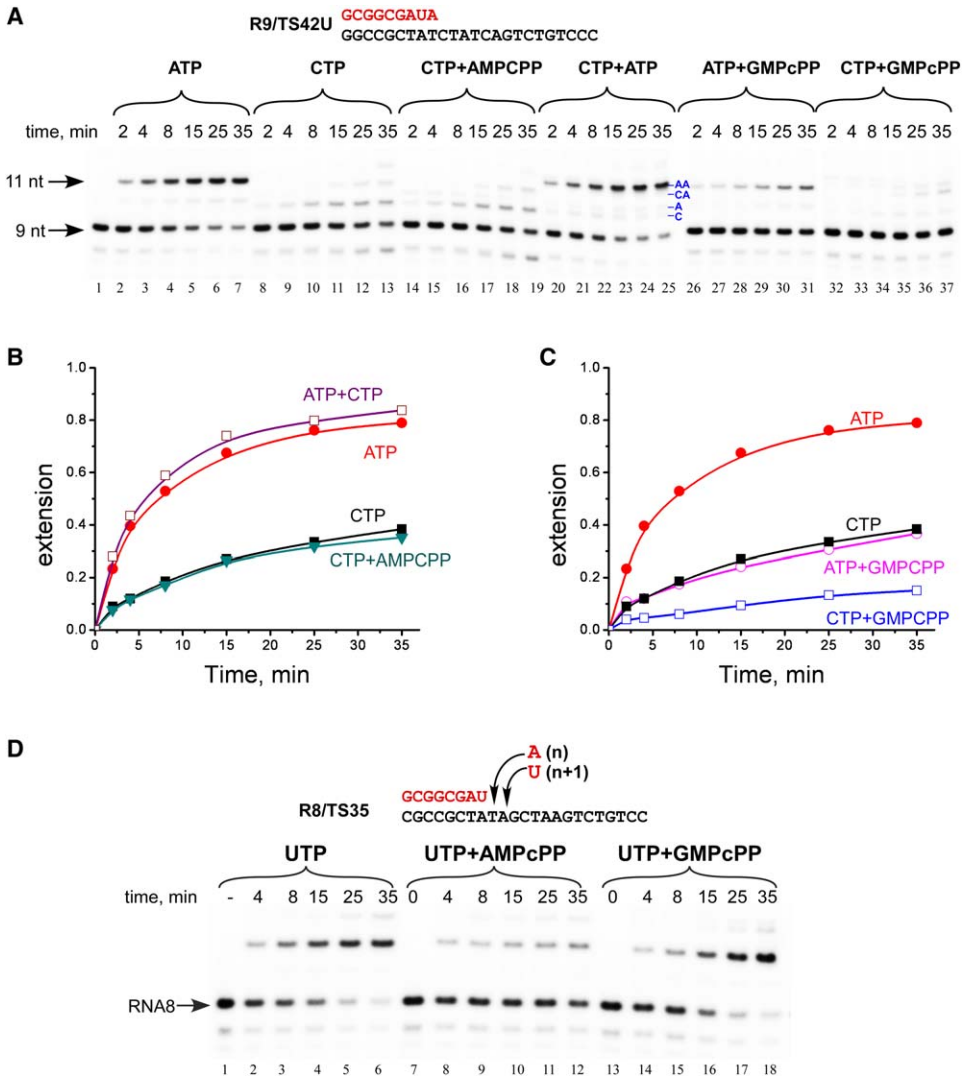

Figure 5. NTP-Stabilized Misalignment

(A) Effect of substrate NTPs on misincorporation and misalignment. *E. coli* ECs were preincubated with nonhydrolyzable NTP analogs (1 mM) for 5 min prior to the addition of the substrate indicated (0.5 mM each).

(B) Effect of the n+1 NTP on misincorporation. The fraction of the primer extended in lanes 2–25 (A) is plotted as a function of time.

(C) Effect of the cognate NTP on misalignment and misincorporation. The fraction of the primer extended in lanes 1–13 and 26–37 (A) is plotted as a function of time.

(D) The misaligned and correctly aligned template bases bind to overlapping sites. *E. coli* EC (R8/TS35) was preincubated with correct (AMPcPP, 1 mM) or noncognate (GMPcPP, 1 mM) substrates for 5 min at room temperature prior to the addition of n+1 UTP (0.5 mM).

is likely to be an important mechanism for producing errors during transcription by multisubunit RNAPs.

We also analyzed the effect of the NT strand on misalignment efficiency in the yeast RNAP II system by examining RNA primer extension by n+1 ATP on scaffolds that differed by the presence or absence of a complementary NT strand downstream of the active site (Figure 6C). In contrast to bacterial RNAP, the rate of n+1 NTP misincorporation by yeast RNAP II was not suppressed in scaffolds having a complementary NT strand in the downstream region (Figure 6C), suggesting that misalignment in yeast RNAP II may occur even in a natural double-stranded context. The observed difference in the efficiency of misalignment on double-stranded templates between RNAP II and bacterial RNAPs may indicate that during elongation RNAP II melts the n+1 base pair in the downstream duplex while bacterial RNAP does not (Holstege et al., 1997; Komissarova and Kashlev, 1998). Taken together, the above experiments suggest that misalignment is a common mechanism by which transcription errors are generated by different multisubunit RNAPs.

It has long been known that there is a DNA-binding site downstream of the active site of RNAPs that contributes to the stability of the EC (Arndt and Chamberlin, 1990; Nudler et al., 1996). Structural studies of core RNAP II ECs revealed contacts of the DNA T strand 1–3 nt downstream of the n site (Gnatt et al., 2001; Westover et al., 2004). The structure of the complete RNAP II EC, which contains duplex DNA in the proximate downstream region, revealed some additional DNA-polymerase contacts that could be responsible for the observed

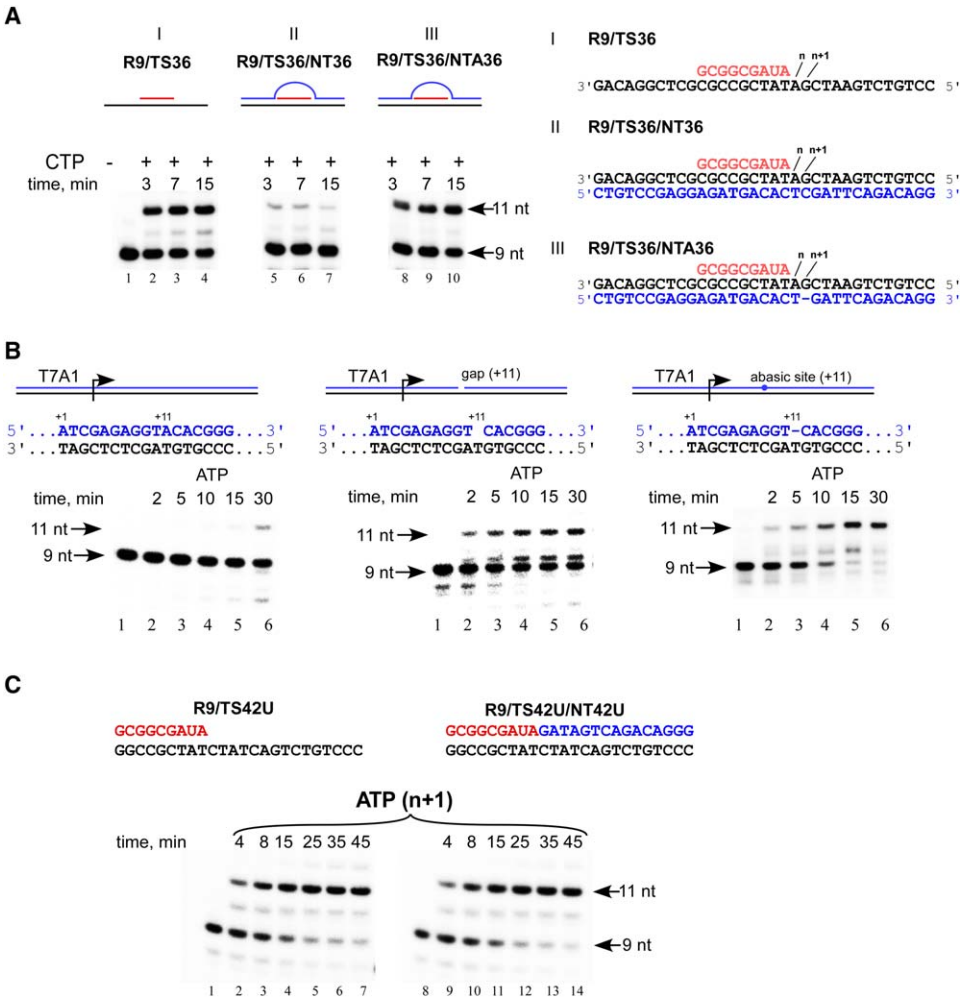

Figure 6. Effect of the NT Strand on Template-Strand Misalignment in Different Types of Multisubunit RNAPs

(A) Misincorporation by misalignment is enhanced in the absence of the NT strand and in the presence of an abasic site at the n+1 position in NT DNA strand in *E. coli* ECs. Nucleic acid scaffolds used are indicated. ECs were incubated with the n+1 CTP (0.5 mM) at 37°C for the times indicated.

(B) The absence of the complementary n+1 base in the NT strand stimulates misincorporation by misalignment in *Tth* ECs halted downstream of the promoter. ECs were incubated with the n+1 ATP (0.5 mM) at 60°C for the times indicated.

(C) The rate of misincorporation by misalignment is not affected by the presence of the NT strand in yeast RNAP II EC. ECs were assembled using yeast RNAP II and the scaffolds indicated and incubated with the n+1 ATP (0.5 mM) for 4–45 min at 30°C.

stabilizing effect (Kettenberger et al., 2004). The data presented here suggest that the proximate downstream DNA grip is important to prevent misalignment errors and to position the templating base accurately in the RNAP active site. Loosening the downstream grip, as exemplified by a lack of the NT strand, leads to error production due to misincorporation and template misalignment.

Comparison with Misalignment in DNAPs

Our data, together with recent studies of T7 RNAP (Pomerantz et al., 2006), suggest that template misalignment is a universal mechanism for generation of transcription errors by both single-subunit and multisubunit RNAPs. Our studies also reveal two fundamental differences to the mechanism of misalignment in DNAPs (Bebenek and Kunkel, 1990; Bell et al., 1997; Garcia-Diaz et al., 2006; Kobayashi et al., 2002). First, in DNAPs, as revealed by recent structural data (Garcia-Diaz et al., 2006), the extrahelical base can be positioned between the n-1 and the n-2 base upstream of the active site, whereas our data so far reveal only that a position between the n-1 and the n base can be occupied in RNAPs. Second, in DNAPs, the extrahelical base can apparently persist and be translocated along the growing template-primer duplex, resulting in a deletion mutation (Bebenek and Kunkel, 1990); but in RNAPs, our data so far indicate that the extrahelical base is not translocated into the next position but undergoes realignment, leading to a substitution mutation. The modeling shows that an extrahelical base between positions n-2 and n-1 clashes with two protein regions in yeast RNAP II, switch 2 and switch 3 (Cramer et al., 2001). This steric constraint explains why an extrahelical base accommodated just before that position during misalignment is apparently not translocated but is rather realigned, resulting in the formation of a mismatched base pair at the end of the RNA:DNA hybrid.

In contrast to error production by DNAPs, misincorporation errors by RNAPs likely result only in phenotypic mutations in the cell, as they do not change the genetic material and are therefore not inherited. It is known that errors at the transcription level can give rise to the synthesis of toxic proteins as demonstrated for the variants of β-amyloid precursor protein and ubiquitin B found in neurons of Alzheimer's and Down's syndrome patients (van Leeuwen et al., 2006). Such erroneous transcription of nonmutated genes has been termed "misreading" and results in frameshift mutations (van Leeuwen et al., 2000). It remains to be seen whether erroneous RNA resulting from template misalignment, which may be facilitated at certain DNA sequences, gives rise to aberrations in protein function in living cells. To understand whether and how RNAPs can generate deletion errors via a misalignment mechanism requires additional studies.

Concluding Remarks

Our data on cellular multisubunit RNAPs, together with recent studies of T7 RNAP (Pomerantz et al., 2006), suggest that template misalignment is a universal mechanism by which substitution errors are generated during transcription. In vitro, errors due to T-strand misalignment appear to dominate over errors due to misincorporation, independent of the nucleic acid scaffold or the nature of the EC. Template-strand misalignment should therefore be considered as a factor that may affect experiments that involve kinetics of NTP binding, incorporation, and misincorporation. In vivo, it is likely that misalignment is enhanced when abasic sites or lesions are encountered by RNAP before the DNA-repair machinery fixes the damages. Although the presence of the correct NTP inhibits misalignment, the high efficiency of this mechanism as opposed to misincorporation suggests that a considerable number of substitution errors in RNA might be generated by misalignment. Indeed, experiments in vivo demonstrated that when substrates are present in nonequimolar concentrations, misincorporation by DNAP due to misalignment is increased 7- to 15-fold (Bebenek and Kunkel, 1990). An important aim for the future will be to investigate the role of template misalignment in the fidelity of cellular RNAPs in vivo.

Experimental Procedures

Purification of RNAP and Polymerase Activity Assay

WT *Tth* core and holo RNAP were purified from cell biomass obtained from HB8 strain as described (Vassylyeva et al., 2002). WT *E. coli* core polymerase was purified from cell biomass obtained from MRE600 strain as described (Nudler et al., 2003). The purity of these RNAPs was greater then 99.8% as judged by SDS-PAGE. Polymerase activity was measured by the ability to extend a ^{32}P-labeled RNA primer by 1 nt in a reaction where the concentrations of nucleic acid scaffold (R8/TS35/NT35) and RNAP were equimolar (see transcription conditions below). Yeast RNAP II was prepared as described (Armache et al., 2003).

RNA and DNA Oligonucleotides

The following synthetic oligonucleotides were used (all sequences are 5′ to 3′): RNA oligomers (Dharmacon), (R8) GCGGCGAU and (R9) GCGGCGAUA; DNA oligomers (IDT DNA), (TS35) CCTGTC TGAATCGATATCGCCGC, (TS36) CCTGTCTGAATCGATATCGCCG CGCTCGGACAG, (TS42U) CCCTGTCTGACTATCTATCGCCGG, (TS42pC, **C^{p}** indicates pyrrolo-C) CCTGTCTGACTAT**C^{p}**TATCGCC GG, (TS40) CCTGTCTGAATCGTCATCGCCGC, (TS44G) CCTGTCTG AATACGTATCGCCGC, (NT42U) GATAGTCAGACAGGG, (NT35) TCG ATTCAGACAGG, (NT38) ATCGATTCAGACAGG, (NT36) CTGTCCGA GGAGATGACACTCGATTCAGACAGG, and (NTA36, "-" indicates abasic site) CTGTCCGAGGAGATGACACTC-ATTCAGACAGG.

For promoter-based experiments, the following DNA oligos containing T7A1 promoter were used: EK33 (T strand, TTCCCCGTGTAC CTCTCGATGGCTGTAAGTATCCTATAGGTTAGACTTTAAGTCAATA CT), EK32 (NT strand, AGTATTGACTTAAAGTCTAACCTATAGGATA CTTACAGCCATCGAGAGGTACACGGGGAA), EK36 (NT strand with abasic [-] site) AGTATTGACTTAAAGTCTAACCTATAGGATACTTAC AGCCATCGAGAGGT-CACGGGGAA), EK34 (NT strand, AGTATTGA CTTAAAG TCTAACCTATAGGATACTTACAGCCATCGAGAGG), and EK35 (NT strand, CACGGGGAA)

Substrate NTPs

Nonhydrolyzable NTP analogs were from Sigma (AMPcPP), Jena Scientific (GMPcPP), and BIOLOG Life Science Institute (UMPcPP). All other NTP substrates were from GE Health.

Assembly of ECs and Transcription Conditions

Nucleic acid scaffolds were assembled by annealing equimolar concentrations of complementary RNA and DNA oligomers as previously described (Temiakov et al., 2002). RNA and DNA oligomers were labeled at their 5′ ends using γ-^{32}P ATP and polynucleotide kinase (New England Biolabs) prior to assembly. To assemble ECs, core RNAP (0.2–1 μM) was incubated with an equimolar concentration of scaffold for 5 min at room temperature in 10 μl of transcription buffer (40 mM Tris [pH 7.9] at 25°C, 100 mM NaCl, 5 mM $MgCl_2$, and 5 mM 2-mercaptoethanol). Primer extension was achieved by incubation of complexes with substrate NTPs (500 μM) for 1–60 min at 60°C (*Tth* RNAP), 37°C (*E. coli* RNAP), or 30°C (yeast RNAP II).

ECs halted downstream of a promoter were formed by incubating *Tth* holo RNAP (40 nM) with linear DNA template containing the T7A1 promoter (40 nM), ApU (200 μM), and appropriate NTPs (50 μM) on ice in 10 μl of transcription buffer. The reactions were incubated at 60°C for 10 min, and NTPs were removed by gel filtration on Quick Spin columns (G50, Roche). Reactions were terminated by the addition of 90% formamide and 50 mM EDTA solution, and products were resolved by electrophoresis in 20% PAGE in the presence of 6 M urea and visualized by PhosphorImager (GE Health).

Fluorescence Spectroscopy

Fluorescence measurements were carried out using a Fluoromax-3 fluorometer (HORIBA Jobin Yvon) equipped with an F-3004 thermostated sample holder and an LFI-3751 temperature controller (Wavelength Electronics). Quartz cells (30 μl) with 1.5 × 1.5 mm optical length paths (Hellma) were used. Samples were excited at 350 nm, and fluorescence emission was monitored at 440 nm with both excitation and emission slits set at 5 nm. The nucleic acid scaffold carrying a fluorescent probe (pyrrolo-C) was prepared as described above by annealing together oligonucleotides R9 and TS42pC. The ECs were prepared by preincubating 200 nM scaffold with an equimolar amount of *E. coli* RNAP or yeast RNAP II in transcription buffer containing 5 mM $MgCl_2$ for 10 min at 23°C. The resulting fluorescent complexes were incubated with hydrolyzable (0.5 mM) or nonhydrolyzable (1–2 mM) NTPs for 5 min at 23°C, and emission was monitored for 5–10 min as a function of time. All measurements were performed at 23°C. The effect of protein and NTP absorption on the emission of the pyrrolo-C was determined to be negligible.

Supplemental Data

Supplemental Data include two figures and can be found with this article online at http://www.molecule.org/cgi/content/full/24/2/257/DC1/.

Acknowledgments

We thank Dr. Laurie K. Read, SUNY Buffalo School of Medicine, for the generous gift of UMPcPP. This work was supported by NIH grant GM38147 (to W.T.M.) and by a UMDNJ Foundation grant (to D.T.). P.C. is supported by the Deutsche Forschungsgemeinschaft.

Received: July 28, 2006
Revised: September 27, 2006
Accepted: October 3, 2006
Published: October 19, 2006

References

Armache, K.J., Kettenberger, H., and Cramer, P. (2003). Architecture of initiation-competent 12-subunit RNA polymerase II. Proc. Natl. Acad. Sci. USA *100*, 6964–6968.

Arndt, K.M., and Chamberlin, M.J. (1990). RNA chain elongation by Escherichia coli RNA polymerase. Factors affecting the stability of elongating ternary complexes. J. Mol. Biol. *213*, 79–108.

Bebenek, K., and Kunkel, T.A. (1990). Frameshift errors initiated by nucleotide misincorporation. Proc. Natl. Acad. Sci. USA *87*, 4946–4950.

Bell, J.B., Eckert, K.A., Joyce, C.M., and Kunkel, T.A. (1997). Base miscoding and strand misalignment errors by mutator Klenow polymerases with amino acid substitutions at tyrosine 766 in the O helix of the fingers subdomain. J. Biol. Chem. *272*, 7345–7351.

Blank, A., Gallant, J.A., Burgess, R.R., and Loeb, L.A. (1986). An RNA polymerase mutant with reduced accuracy of chain elongation. Biochemistry *25*, 5920–5928.

Cramer, P., Bushnell, D.A., and Kornberg, R.D. (2001). Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. Science *292*, 1863–1876.

Doublie, S., and Ellenberger, T. (1998). The mechanism of action of T7 DNA polymerase. Curr. Opin. Struct. Biol. *8*, 704–712.

Erie, D.A., Hajiseyedjavadi, O., Young, M.C., and von Hippel, P.H. (1993). Multiple RNA polymerase conformations and GreA: control of fidelity of transcription. Science *262*, 867–873.

Garcia-Diaz, M., Bebenek, K., Krahn, J.M., Pedersen, L.C., and Kunkel, T.A. (2006). Structural analysis of strand misalignment during DNA synthesis by a human DNA polymerase. Cell *124*, 331–342.

Gnatt, A.L., Cramer, P., Fu, J., Bushnell, D.A., and Kornberg, R.D. (2001). Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 A resolution. Science *292*, 1876–1882.

Gong, X.Q., Zhang, C., Feig, M., and Burton, Z.F. (2005). Dynamic error correction and regulation of downstream bubble opening by human RNA polymerase II. Mol. Cell *18*, 461–470.

Holmes, S.F., Santangelo, T.J., Cunningham, C.K., Roberts, J.W., and Erie, D.A. (2006). Kinetic investigation of Escherichia coli RNA polymerase mutants that influence nucleotide discrimination and transcription fidelity. J. Biol. Chem. *281*, 18677–18683.

Holstege, F.C., Fiedler, U., and Timmers, H.T. (1997). Three transitions in the RNA polymerase II transcription complex during initiation. EMBO J. *16*, 7468–7480.

Huang, J., Brieba, L.G., and Sousa, R. (2000). Misincorporation by wild-type and mutant T7 RNA polymerases: identification of interactions that reduce misincorporation rates by stabilizing the catalytically incompetent open conformation. Biochemistry *39*, 11571–11580.

Johnson, S.J., and Beese, L.S. (2004). Structures of mismatch replication errors observed in a DNA polymerase. Cell *116*, 803–816.

Johnson, S.J., Taylor, J.S., and Beese, L.S. (2003). Processive DNA synthesis observed in a polymerase crystal suggests a mechanism for the prevention of frameshift mutations. Proc. Natl. Acad. Sci. USA *100*, 3895–3900.

Kettenberger, H., Armache, K.J., and Cramer, P. (2004). Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIS. Mol. Cell *16*, 955–965.

Kobayashi, S., Valentine, M.R., Pham, P., O'Donnell, M., and Goodman, M.F. (2002). Fidelity of Escherichia coli DNA polymerase IV. Preferential generation of small deletion mutations by dNTP-stabilized misalignment. J. Biol. Chem. *277*, 34198–34207.

Komissarova, N., and Kashlev, M. (1998). Functional topography of nascent RNA in elongation intermediates of RNA polymerase. Proc. Natl. Acad. Sci. USA *95*, 14699–14704.

Landick, R. (2004). Active-site dynamics in RNA polymerases. Cell *116*, 351–353.

Landick, R. (2005). NTP-entry routes in multi-subunit RNA polymerases. Trends Biochem. Sci. *30*, 651–654.

Liu, C., and Martin, C.T. (2002). Promoter clearance by T7 RNA polymerase. Initial bubble collapse and transcript dissociation monitored by base analog fluorescence. J. Biol. Chem. *277*, 2725–2731.

Martin, C.T., Ujvari, A., and Liu, C. (2003). Evaluation of fluorescence spectroscopy methods for mapping melted regions of DNA along the transcription pathway. Methods Enzymol. *371*, 13–33.

Nedialkov, Y.A., Gong, X.Q., Hovde, S.L., Yamaguchi, Y., Handa, H., Geiger, J.H., Yan, H., and Burton, Z.F. (2003). NTP-driven translocation by human RNA polymerase II. J. Biol. Chem. *278*, 18303–18312.

Ninio, J. (1991). Transient mutators: a semiquantitative analysis of the influence of translation and transcription errors on mutation rates. Genetics *129*, 957–962.

Nudler, E., Avetissova, E., Markovtsov, V., and Goldfarb, A. (1996). Transcription processivity: protein-DNA interactions holding together the elongation complex. Science *273*, 211–217.

Nudler, E., Gusarov, I., and Bar-Nahum, G. (2003). Methods of walking with the RNA polymerase. Methods Enzymol. *371*, 160–169.

Ozoline, O.N., Oganesjan, M.G., and Kamzolova, S.G. (1980). On the fidelity of transcription by Escherichia coli RNA polymerase. FEBS Lett. *110*, 123–125.

Pomerantz, R.T., Temiakov, D., Anikin, M., Vassylyev, D.G., and McAllister, W.T. (2006). A mechanism of nucleotide misincorporation during transcription due to template-strand misalignment. Mol. Cell. *24*, this issue, 245–255.

Remington, K.M., Bennett, S.E., Harris, C.M., Harris, T.M., and Bebenek, K. (1998). Highly mutagenic bypass synthesis by T7 RNA polymerase of site-specific benzo[a]pyrene diol epoxide-adducted template DNA. J. Biol. Chem. *273*, 13170–13176.

Rosenberger, R.F., and Hilton, J. (1983). The frequency of transcriptional and translational errors at nonsense codons in the lacZ gene of Escherichia coli. Mol. Gen. Genet. *191*, 207–212.

Shaw, R.J., Bonawitz, N.D., and Reines, D. (2002). Use of an in vivo reporter assay to test for transcriptional and translational fidelity in yeast. J. Biol. Chem. *277*, 24420–24426.

Springgate, C.F., and Loeb, L.A. (1975). On the fidelity of transcription by Escherichia coli ribonucleic acid polymerase. J. Mol. Biol. *97*, 577–591.

Steitz, T.A. (1998). A mechanism for all polymerases. Nature *391*, 231–232.

Svetlov, V., Vassylyev, D.G., and Artsimovitch, I. (2004). Discrimination against deoxyribonucleotide substrates by bacterial RNA polymerase. J. Biol. Chem. *279*, 38087–38090.

Temiakov, D., Anikin, M., and McAllister, W.T. (2002). Characterization of T7 RNA polymerase transcription complexes assembled on nucleic acid scaffolds. J. Biol. Chem. *277*, 47035–47043.

Temiakov, D., Patlan, V., Anikin, M., McAllister, W.T., Yokoyama, S., and Vassylyev, D.G. (2004). Structural basis for substrate selection by T7 RNA polymerase. Cell *116*, 381–391.

Thomas, M.J., Platas, A.A., and Hawley, D.K. (1998). Transcriptional fidelity and proofreading by RNA polymerase II. Cell *93*, 627–637.

van Leeuwen, F.W., Fischer, D.F., Kamel, D., Sluijs, J.A., Sonnemans, M.A., Benne, R., Swaab, D.F., Salehi, A., and Hol, E.M. (2000). Molecular misreading: a new type of transcript mutation expressed during aging. Neurobiol. Aging *21*, 879–891.

van Leeuwen, F.W., van Tijn, P., Sonnemans, M.A., Hobo, B., Mann, D.M., Van Broeckhoven, C., Kumar-Singh, S., Cras, P., Leuba, G., Savioz, A., et al. (2006). Frameshift proteins in autosomal dominant forms of Alzheimer disease and other tauopathies. Neurology *66*, S86–S92.

Vassylyeva, M.N., Lee, J., Sekine, S.I., Laptenko, O., Kuramitsu, S., Shibata, T., Inoue, Y., Borukhov, S., Vassylyev, D.G., and Yokoyama, S. (2002). Purification, crystallization and initial crystallographic analysis of RNA polymerase holoenzyme from *Thermus thermophilus*. Acta Crystallogr. *D58*, 1497–1500.

Westover, K.D., Bushnell, D.A., and Kornberg, R.D. (2004). Structural basis of transcription: nucleotide selection by rotation in the RNA polymerase II active center. Cell *119*, 481–489.