Molecular Cell
Article

Structural Basis of Transcription: Mismatch-Specific Fidelity Mechanisms and Paused RNA Polymerase II with Frayed RNA

Jasmin F. Sydow,[1] Florian Brueckner,[1,3] Alan C.M. Cheung,[1] Gerke E. Damsma,[1] Stefan Dengl,[1] Elisabeth Lehmann,[1] Dmitry Vassylyev,[2] and Patrick Cramer[1,*]
[1]Department of Chemistry and Biochemistry, Gene Center Munich and Center for Integrated Protein Science CIPSM, Ludwig-Maximilians-Universität München, Feodor-Lynen-Strasse 25, 81377 Munich, Germany
[2]Department of Biochemistry and Molecular Genetics, University of Alabama at Birmingham, Schools of Medicine and Dentistry, 720 20th Street South, Birmingham, AL 35294, USA
[3]Current address: Membrane Protein Laboratory, Imperial College London, Diamond Light Source Ltd., Harwell Science and Innovation Campus, Didcot OX11 0DE, UK
*Correspondence: cramer@lmb.uni-muenchen.de
DOI 10.1016/j.molcel.2009.06.002

SUMMARY

We show that RNA polymerase (Pol) II prevents erroneous transcription in vitro with different strategies that depend on the type of DNA·RNA base mismatch. Certain mismatches are efficiently formed but impair RNA extension. Other mismatches allow for RNA extension but are inefficiently formed and efficiently proofread by RNA cleavage. X-ray analysis reveals that a T·U mismatch impairs RNA extension by forming a wobble base pair at the Pol II active center that dissociates the catalytic metal ion and misaligns the RNA 3′ end. The mismatch can also stabilize a paused state of Pol II with a frayed RNA 3′ nucleotide. The frayed nucleotide binds in the Pol II pore either parallel or perpendicular to the DNA-RNA hybrid axis (fraying sites I and II, respectively) and overlaps the nucleoside triphosphate (NTP) site, explaining how it halts transcription during proofreading, before backtracking and RNA cleavage.

INTRODUCTION

Structural and functional studies of DNA polymerases revealed that the fidelity of DNA-dependent DNA synthesis is achieved by a high selectivity of the polymerase for the correct dNTP substrate, but also by the ability of the polymerase to detect misincorporation and invoke proofreading (Kunkel and Bebenek, 2000; McCulloch and Kunkel, 2008). Comparatively little is known about the mechanisms that govern the fidelity of RNA synthesis by DNA-dependent RNA polymerases (Pols), although it is thought that transcription fidelity prevents formation of erroneous mRNAs and mutant proteins with impaired function (Saxowsky and Doetsch, 2006).

Mechanisms that underlie transcription fidelity were suggested by enzymatic studies of Pols in vitro (Erie et al., 1993; Thomas et al., 1998). These studies showed that misincorporation leads to slow addition of the next nucleotide, and that a mismatched RNA 3′ end can be removed with factors that stimulate the polymerase cleavage activity. In a bacterial elongation complex (EC), a mismatched RNA 3′ nucleotide induces an unactivated state and is removed by cleavage-stimulatory Gre factors (Erie et al., 1993). In human Pol II, a mismatched RNA 3′ nucleotide causes slow addition of the next nucleotide, and RNA cleavage is stimulated by TFIIS (Thomas et al., 1998).

Transcriptional fidelity and proofreading, however, do not require cleavage-stimulatory factors, but are intrinsic properties of Pols. Yeast Pol II fidelity does not depend on TFIIS in vivo (Nesser et al., 2006; Shaw et al., 2002), but requires the Pol II subunit Rpb9 (Nesser et al., 2006). A mutation in the Pol II active center trigger loop leads to decreased transcription fidelity in vitro and a requirement for TFIIS in vivo (Kireeva et al., 2008). The eukaryotic Pol I and Pol III very efficiently cleave RNAs with a mismatched 3′ nucleotide (Alic et al., 2007; Kuhn et al., 2007). Proofreading by Pol III is so efficient that misincorporation can only be detected with a cleavage-deficient polymerase isoform (Alic et al., 2007). As shown for a bacterial Pol, the mismatched 3′ nucleotide itself stimulates RNA cleavage by assisting the active center in the hydrolysis of the penultimate phosphodiester bond (Zenkin et al., 2006).

To study the molecular mechanisms underlying transcription fidelity, we reconstituted complete yeast Pol II ECs and carried out a systematic, quantitative analysis of the three reactions that determine fidelity: misincorporation, mismatch extension, and cleavage of mismatched RNA 3′ ends. These studies reveal different transcription fidelity strategies for different types of mismatches. Exemplary erroneous transcription events are rationalized with X-ray structures of T·U mismatch-containing ECs. These studies reveal accommodation of a T·U wobble base pair (bp) at the active center, mismatch-induced disruption of the catalytic site, and mismatch-stabilized fraying of the RNA 3′ end, which underlies polymerase pausing and likely occurs before backtracking during proofreading.

RESULTS

Misincorporation Efficiency Is Mismatch Specific

To determine the efficiency of misincorporation by Pol II, we performed RNA extension assays with reconstituted ECs

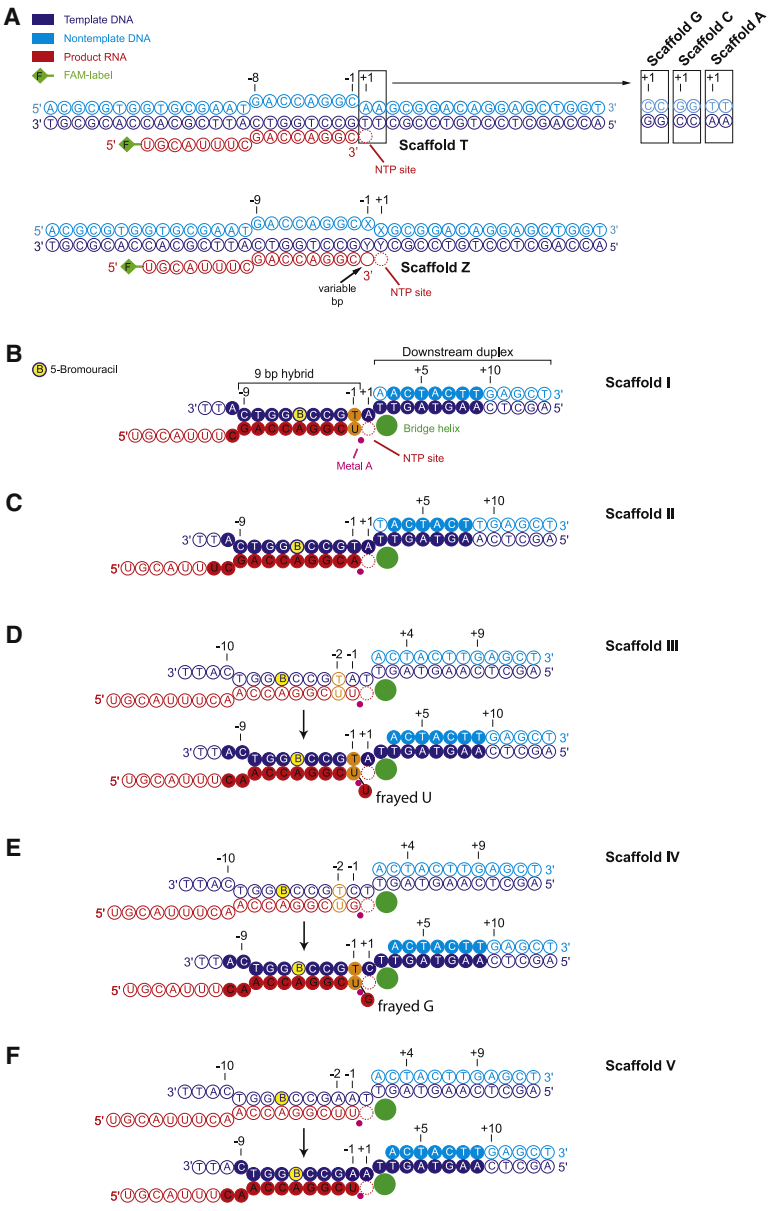

Figure 1. Nucleic Acid Scaffolds for Reconstitution of Pol II ECs

(A) Scaffold T and related scaffolds G, C, and A were used in incorporation assays. Scaffold Z is variable and was used for extension and cleavage assays. The variable bp (black arrow) was one of the sixteen different matched or mismatched bps to mimic the result of all 16 possible (mis)incorporation events obtained with scaffolds T, G, C, and A.

(B) Scaffold I contains a T•U mismatch at position −1 (orange) and was used for structural analysis. Filled circles denote nucleotides with interpretable electron density.

(C) Scaffold II contains a T-A match at position −1.

(D) Scaffold III as designed (top), containing a T•U mismatch at position −2 (orange) and an A-U match at position −1, and as observed in the crystal (bottom) with a frayed 3′ uridine.

(E) Scaffold IV as designed (top), containing a T•U mismatch at position −2 (orange) and a G-C match at position −1, and as observed in the crystal (bottom) with a frayed 3′ guanine.

(F) Scaffold V as designed (top), containing A-U matches at positions −2 and −1, and as observed in the crystal (bottom).

(Brueckner et al., 2007; Kireeva et al., 2003). The nucleic acid scaffolds contained fully complementary DNA strands, 18 bps of downstream DNA, 15 bps of upstream DNA, an eight bp DNA-RNA hybrid, and eight nucleotides of exiting RNA labeled with 6-carboxyfluoresceine (FAM) at its 5′ end (Figure 1). The scaffolds T, G, C, and A differed in the +1 nucleotide opposite the NTP site (Figure 1A). The +1 and +2 nucleotides were identical to prevent misincorporation by template misalignment (Kashkina et al., 2006). To compare the efficiency of all 16 incorporation events (four correct incorporations and 12 misincorporations), the four scaffolds were assembled with Pol II into ECs that were incubated with 0.1 mM of each NTP. Reactions were stopped at 0.5, 1, or 5 min, and product RNAs were separated by gel electrophoresis and quantified with a fluorimager (Figure 2; Experimental Procedures). The relative amounts of misincorporation with respect to correct incorporation are provided in Figure 2B.

Misincorporations generating a purine•purine mismatch occurred with low efficiency, whereas those generating a pyrimidine•pyrimidine mismatch were more efficient, except for the C•C mismatch (Figure 2B, DNA•RNA mismatches are indicated with a dot). No general rule could be derived for misincorporations resulting in purine•pyrimidine and pyrimidine•purine mismatches. Misincorporations resulting in T•G or G•U mismatches were inefficient, but those resulting in C•A or A•C mismatches were efficient. To determine first order rate constants, we performed time course experiments for three types of misincorporations that were representative for low (G•A), medium (T•U), and high (A•C) efficiencies and for their corresponding correct incorporations (Figure 2C). Compared to correct incorporations, the misincorporations leading to G•A, T•U, and A•C mismatches were 4300-, 3400-, and 2000-fold slower, respectively (Experimental Procedures). Thus, Pol II misincorporation efficiencies depend on the type of the resulting mismatch.

Transcript Extension Efficiency Is Mismatch Specific

To investigate the efficiency of RNA extension after misincorporation, ECs were reconstituted that contained the 12 different mismatches at position −1 (scaffold Z, Figure 1A). These ECs mimic the situation after misincorporation and allow monitoring the addition of the next nucleotide. For RNA extension, we added the next complementary NTP and stopped reactions at 1 or 5 min (Figure 2D; Experimental Procedures). To prevent extension after RNA dinucleotide cleavage as a side reaction, the nucleotides at −2 and +1 were different. Incorporation of the next nucleotide was always less efficient when a mismatch was present at −1 instead of a match (Figures 2D and 2E).

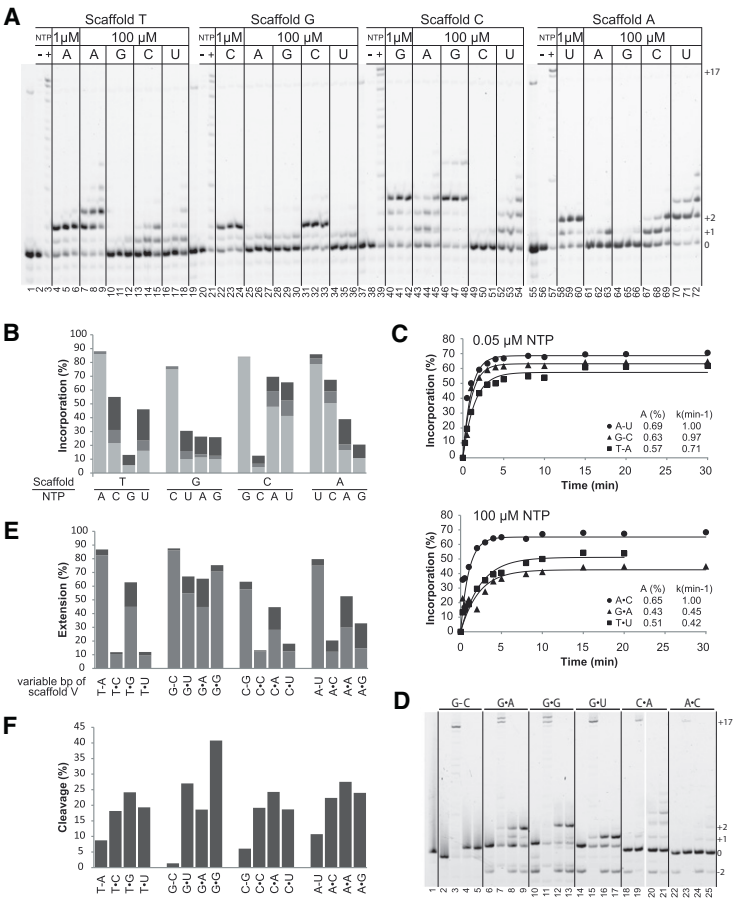

Figure 2. Systematic Quantitative Analysis of Misincorporation, RNA Extension, and RNA Cleavage

(A) Representative gel electrophoresis separation of RNA products obtained in incorporation assays. Lanes 1, 19, 37, and 55 show the fluorescently labeled reactant RNA. ECs of samples shown in lanes 2, 20, 38, and 56 were incubated for 5 min in transcription buffer without addition of NTPs (Experimental Procedures). Run-off controls after incubation with 100 μM NTPs for 5 min are shown in lanes 3, 21, 39, and 57. In the other lanes, the scaffolds were incubated with the indicated NTPs for 0.5, 1, and 5 min (left to right).

(B) Summary of incorporation efficiencies determined by addition of 100 μM of the indicated NTP to the EC. Light gray, gray, and dark gray bars represent the 0.5 min, 1 min, and 5 min time points, respectively. Average values are shown for two independent experiments that generally resulted in very similar values, indicating the high reproducibility.

(C) Time-course experiments for selected incorporation reactions. For correct incorporations, 0.05 μM NTPs were used. For misincorporations, 100 μM NTPs were used. The pre-exponential factor A and the rate constant k were calculated with the program OriginPro 8 (ADDITIVE GmbH) using the equation $c(t) = A \times (1 - \exp[-k \times t])$. For comparison of rate constants of correct incorporation and misincorporation, a dilution factor of 2000 was applied, assuming that reduction of NTP concentration (from 100 to 0.05 μM) leads to equivalent reduction of the rate constant, as described (Alic et al., 2007).

(D) Representative electrophoretic separation of RNA products resulting from RNA extension and cleavage. Six examples are shown for which the bp at position −1 of scaffold Z (Figure 1) is given. Lane 1 shows the fluorescently labeled reactant RNA. Each block of four lanes shows from left to right the cleavage experiment, the run-off control, and extension experiments stopped after 1 and 5 min of incubation. For RNA extension, ECs were incubated with 100 μM of the corresponding next correct NTP.

(E) Summary of RNA extension efficiencies. Grey and dark gray bars represent 1 min and 5 min time points, respectively. Average values for two independent experiments are shown.

(F) Summary of RNA cleavage efficiencies. For these experiments, ECs were incubated in transcription buffer for 5 min. Average values for two independent experiments are shown.

Purine•purine mismatches were more efficiently extended than pyrimidine•pyrimidine mismatches (Figure 2E). Extensions with a mismatched guanine in the RNA were all efficient, generally consistent with results for a bacterial Pol (Zenkin et al., 2006). Among the pyrimidine•purine and purine•pyrimidine mismatches, extension was more efficient for T•G and G•U and less efficient for C•A and A•C. Extensions with a guanine at template position −1 were all efficient. Control experiments showed that the efficiency of incorporating a nucleotide following a matched bp was very similar for the different bps (Figure 2E). Thus, the efficiency of RNA extension is always lower in the presence of a mismatch but varies with the type of mismatch.

Pol II Accommodates a T·U Wobble Pair

To unravel the molecular basis of fidelity mechanisms for one type of mismatch, we determined structures of complete Pol II ECs containing a T•U mismatch. Complete Pol II was cocrystallized with a scaffold containing the mismatch at position –1 (scaffold I, Figure 1B). For the resulting EC I, diffraction data of very high quality were obtained, and the register of nucleic acids was defined by bromine labeling (Table 1; Figure 3; Experimental Procedures). With the use of zonal scaling (Vassylyev et al., 2007a), the structure was refined to a free R factor of 25.2% at 3.2 Å resolution, the highest resolution for a complete Pol II structure (Figures 3A–3C; Table 1). The structure showed that EC I adopts the posttranslocation state and accommodates the T•U mismatch at the active center at position −1 (Figures 3A, 3C, and 3D). The mismatch adopts a wobble bp that is stabilized by two hydrogen bonds formed between the N3 and the O2 atoms of the uracil and the O4 and N3 atoms, respectively, of the template thymine (Figure 3D). The accommodation of a wobble pair may explain why uridine misincorporation opposite a template thymine is efficient (Figure 2B) and supports our previous proposal that uridine misincorporation opposite a thymine within a DNA photolesion results from wobble formation (Brueckner et al., 2007).

Active Site Disruption Explains Impaired RNA Extension

To detect the structural changes in EC I that result from the T•U mismatch, we solved a reference structure that contained a matched T-A bp at position −1 (scaffold II, EC II, Figure 1C; Table 1). The overall structures of ECs I and II did not deviate, but in the mismatched EC I, the 3′-terminal RNA nucleotide at position −1 and its 5′-flanking phosphate were shifted away from the active site by over 2 Å (Figure 3E). Thus, the T•U wobble triggers misalignment of the nucleophilic RNA 3′ end with the

Table 1. Diffraction Data and Refinement Statistics

	EC I	EC II	EC III	EC IV	EC V	EC VI
Data Collection						
Space group	C2	$C222_1$	$C222_1$	$C222_1$	$C222_1$	$C222_1$
Unit cell axes (Å)	394.3, 221.6, 283.4	222.3, 393.4, 283.1	222.7, 396.0, 283.5	221.4, 393.8, 281.8	221.6, 393.7, 282.6	222.1, 392.7, 282.4
Unit cell β angle (°)	90.9	90	90	90	90	90
Wavelength (Å)	0.9189	0.9190	0.9188	0.9188	0.9188	0.9177
Resolution range (Å)	40–3.20	50–3.50	50–3.60	50–3.65	50–3.65	50–3.40
Unique reflections	372,166[a] (32,852)[b]	155,150 (21,507)	144,009 (19,441)	135,977 (18,105)	136,470 (18,185)	168,339 (24,019)
Completeness (%)	95.6 (84.7)	99.9 (100)	99.9 (99.9)	99.9 (100)	99.9 (100)	99.9 (100)
Redundancy	3.0 (2.2)	7.3 (7.2)	7.3 (7.3)	7.6 (7.8)	7.5 (7.4)	7.5 (7.9)
Mosaicity (°)	0.38–0.72[c]	0.11	0.14	0.12	0.09	0.08
R_{sym} (%)	7.5 (37.5)	9.5 (52.9)	9.2 (75.0)	7.7 (63.6)	6.6 (52.0)	6.4 (50.4)
I/σ (I)	20.7 (2.6)	15.8 (4.7)	17.6 (3.2)	22.0 (3.7)	22.9 (4.3)	24.1 (5.0)
Refinement						
Nonhydrogen atoms	63,666	31,778	31,877	31,962	31,935	31,804
RMSD bonds	0.010	0.010	0.011	0.010	0.011	0.010
RMSD angles	1.60	1.59	1.65	1.61	1.65	1.61
R_{cryst} (%)	23.3	21.0	21.4	21.0	21.2	21.6
R_{free} (%)	25.2	22.6	25.4	25.3	25.0	25.4
Br peak in anom. Fourier (σ)	8.6	8.6	10.9	8.1	8.7	9.6
Ramachandran Statistics						
Core (%)	71.8[d]/72.0[e]	71.1	72.7	72.6	70.6	74.0
Allowed (%)	23.4/23.3	24.1	22.6	22.9	24.6	21.7
Generally allowed (%)	3.1/3.0	3.2	3.2	2.6	3.2	3.1
Disallowed (%)	1.7/1.7	1.6	1.5	1.9	1.6	1.2

[a] Friedel pairs are merged.
[b] Values in parentheses are for highest resolution shell.
[c] Refined for batches of images.
[d] Molecule 1 of the asymmetric unit.
[e] Molecule 2 of the asymmetric unit.

catalytic site and NTP, and a deviation from the optimum geometry for catalysis, a collinear inline attack during an S_N2 reaction (Figure 3E). In addition, the active site aspartate loop lost the catalytic metal ion A (Figure 3F). The three metal-binding aspartate side chains in Rpb1 changed conformation (Figure 3G). The D481 carboxylate is mobile, and the side chains of D483 and D485 could both form a hydrogen bond with the RNA 3′ hydroxyl (Figure 3F). Metal A is apparently lost due to the disruption of the active site by the wobble bp since it is observed in EC II and in a published EC structure obtained under the same conditions (Brueckner and Cramer, 2008). Thus, the low efficiency of RNA extension after a T•U mismatch can be explained by disruption of the catalytic site that involves loss of the catalytic metal A and a shift of the RNA 3′ end.

Mismatch Extension and RNA 3′ Fraying

To investigate RNA extension past the mismatch, we prepared a scaffold with the T•U mismatch at position −2 and an A-U bp at position −1 (scaffold III, Figure 1D). In the resulting EC III structure (Table S1 available online), the hybrid was similar to that in EC I, including the T•U wobble bp at position −1, and downstream DNA was slightly shifted as previously observed (Brueckner et al., 2007). The 3′-terminal RNA uridine, however, did not form a bp with the template adenine as designed but was flipped away from the template, creating a frayed RNA end (Figures 4A and 4B). The frayed uracil was oriented parallel to the axis of the DNA-RNA hybrid and occupied a site in the pore ("fraying site I," Figure 4C). A frayed RNA 3′ nucleotide was shown biochemically to be the hallmark of a common elongation intermediate, the elemental pause, that occurs during polymerase pausing and before transcription arrest and termination (Artsimovitch and Landick, 2000; Chan et al., 1997; Toulokhonov et al., 2007). The frayed nucleotide overlaps the tip of the closed trigger loop and the NTP in the insertion site (Figure 5), and contacts Rpb2 residues R766 and R1020, which also bind the NTP triphosphate (Table S1). This explains how the frayed RNA end interferes with nucleotide binding and incorporation.

In EC III, fork loop 2 adopts a new conformation (Figure 6). Fork loop 2 residues have moved by up to 6 Å toward the DNA nontemplate strand at the downstream edge of the transcription

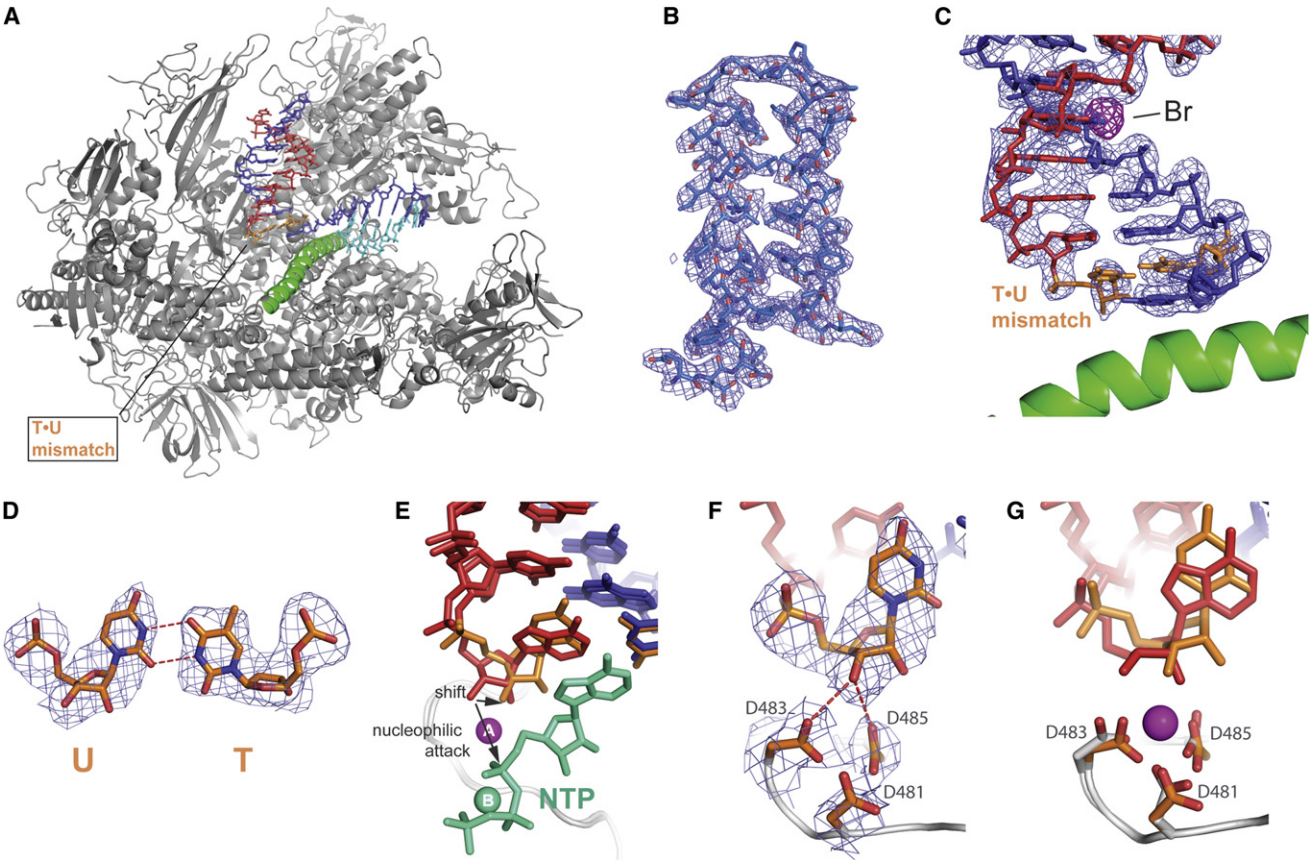

Figure 3. Structure of EC I Reveals a T·U Wobble Pair at 3.2 Å Resolution

(A) Structure of the T•U mismatch-containing Pol II EC I. Pol II is shown from the side as a ribbon model in silver, with the bridge helix highlighted in green and a portion omitted for clarity. The nucleic acids are shown as stick models using the same color code as in Figure 1. The T•U mismatch is shown in orange throughout.

(B) Representative protein electron density. The final $2F_o$-F_c density is shown as a blue mesh, contoured at 1.1σ. Depicted is the clamp coiled coil, an exposed part of subunit Rpb1.

(C) Electron density of part of the DNA-RNA hybrid ($2F_o$-F_c map contoured at 1.8σ). A peak in the anomalous difference Fourier map (magenta, contoured at 4.3σ) reveals the location of the bromine atom at position −5 of the template strand, defining the posttranslocated state.

(D) T•U wobble base pair in the Pol II active center. The final $2F_o$-F_c electron density map is shown in blue, contoured at 1.0σ. Hydrogen bonds are indicated by red dashed lines.

(E) Superposition of the mismatched EC I with the matched EC II (at 3.5 Å resolution, Table 1) reveals a 2 Å shift of the RNA 3′-hydroxyl (horizontal arrow, the mismatched terminal RNA U residue is shown in orange). As a consequence, the nucleophilic RNA 3′ end is no longer in a position suited for an in-line nucleophilic attack (vertical arrow) of the phosphodiester bond between the α and β phosphates of the incoming NTP substrate (green cyan, taken from PDB-code 2O5J [Vassylyev et al., 2007b]). The structures EC I and 2O5J were superimposed by least-squares fitting of Rpb1 residues A478-A487 to β′ residues D745-D736 and RNA residues in positions −1, −2, and −3. Metal ion A is from EC II, and metal ion B is from 2O5J. For NTP modeling, we used the bacterial NTP complex structure rather than the yeast core Pol II NTP complex since it contains an intact RNA 3′ hydroxyl group.

(F) Loss of metal ion A in the active site of EC I. The final $2F_o$-F_c electron density map is contoured at 1.0σ.

(G) Comparison of the RNA 3′ nucleotide and the catalytic aspartate loop in EC I (orange) and EC II (gray). Metal A (pink sphere) is only present in EC II.

bubble (Figure 6A). The guanidinium head group of Rpb2 residue R504 forms two hydrogen bonds to N7 and O6 of the template guanine at +4 (Figure 6B). R504 is invariant among Pol II enzymes and bacterial and archaeal Pols, but not conserved in Pol I and III (Figure 6D) (Jasiak et al., 2006; Kuhn et al., 2007; Naji et al., 2008. This arginine is important for promoter-dependent transcription and normal elongation (Naji et al., 2008. It is possible that the observed fork loop 2 downstream DNA interaction, or alternative contacts of the flexible arginine (Figure 6C) with other nearby bases in DNA, contribute to the stability of the paused state as suggested (Toulokhonov et al., 2007).

Two RNA Fraying Sites

To test whether the fraying was dependent on the stability of the bp at the end of the hybrid, we replaced the A-U bp in scaffold III with a C-G bp (scaffold IV, Figure 1E). The resulting EC IV structure was very similar to that of EC III, including the T•U wobble bp (Figures 4D and 4E). The RNA 3′ nucleotide was again frayed, but was oriented perpendicular to the hybrid axis, occupying a different site in the pore ("fraying site II," Figure 4F). Fraying sites I and II are both lined by Rpb1 residues K987 and D483 but are separated by Rpb2 residue Y769, which stacks against the frayed guanine (Figure 4I). The frayed guanine contacts Rpb2 residue

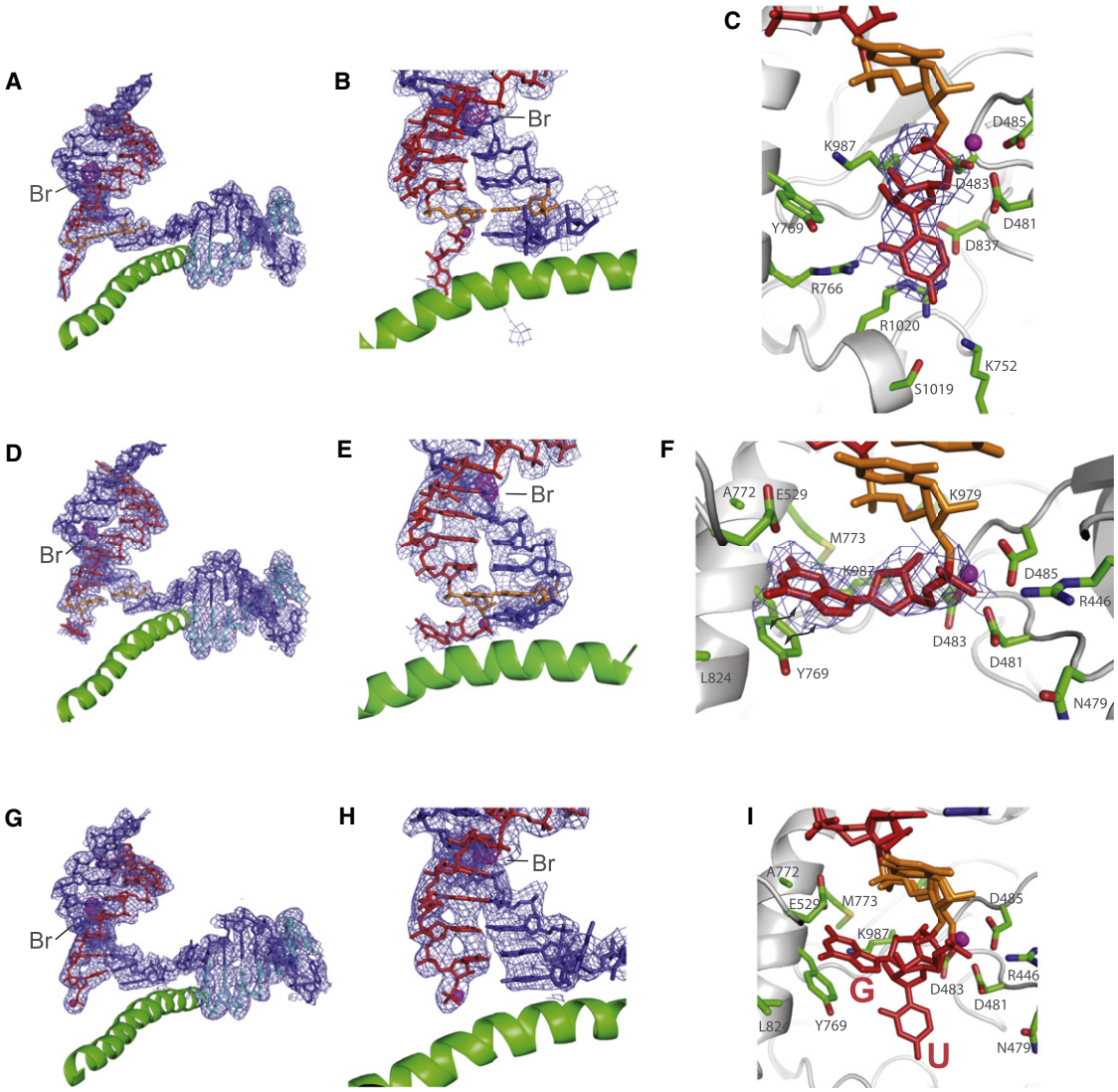

Figure 4. Two Frayed States of the RNA 3′ Nucleotide

(A) Structure of Pol II EC III reveals a frayed 3′-terminal RNA uridine at 3.6 Å resolution. The final $2F_o$-F_c electron density of the nucleic acids is shown as a blue mesh, contoured at 1.2σ. The location of the bromine atom at position −5 defines the register (the anomalous difference Fourier map is shown in magenta, contoured at 4.2σ).

(B) Detailed view of the electron density map in (A) near the active center.

(C) Fraying site I. Depicted are Pol II residues contacting the frayed 3′-terminal RNA uridine. The final $2F_o$-F_c density is shown for the frayed nucleotide, contoured at 0.9σ.

(D) Nucleic acids structure of EC IV reveals a frayed 3′-terminal RNA guanine at 3.65 Å resolution. The final $2F_o$-F_c is shown as a blue mesh, contoured at 1.0σ. The bromine peak at position −5 defines the register (anomalous difference Fourier, magenta, contoured at 4.5σ).

(E) Detailed view of the electron density map in (D) near the active center.

(F) Fraying site II. Depicted are Pol II residues contacting the frayed 3′-terminal RNA guanine. The final $2F_o$-F_c density is shown for the frayed nucleotide, contoured at 1.0σ. Stacking interactions are indicated by two-headed arrows.

(G) Structure of EC V at 3.65 Å resolution reveals a mobile 3′-terminal nucleotide. The final $2F_o$-F_c electron density is contoured at 1.0 σ (blue) and shows only the phosphate group of the 3′-nucleotide. The location of the bromine atom at position −5 defines the register (anomalous difference Fourier, magenta, contoured at 4.3σ).

(H) Detailed view of the electron density map in (G).

(I) Superposition of the structures of ECs III and IV allows for comparison of the two frayed RNA 3′-nucleotides that are either oriented parallel (U, fraying site I) or perpendicular (G, fraying site II) to the axis of the DNA-RNA hybrid (vertical in this view).

E529 in a region called βDloopII in bacterial Pol (Table S1). Thus, the RNA 3′ nucleotide can occupy at least three alternative sites, the pretranslocated position, which preserves base pairing with the template, and two alternative fraying sites in the pore, in which this base pairing is disrupted. The frayed nucleotide is either oriented along the hybrid axis and approaches the NTP triphosphate-binding site (fraying site I), or it is oriented perpendicular to the hybrid axis and approaches βDloopII (fraying site II).

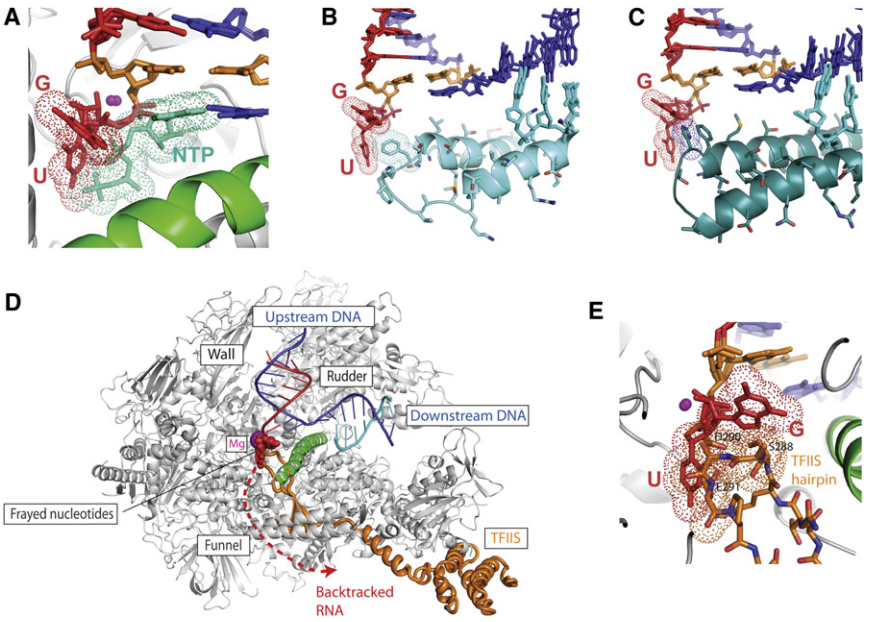

Figure 5. Frayed Nucleotides Overlap the NTP, Closed Trigger Loop, and the TFIIS Hairpin

(A) Frayed nucleotides overlap the NTP bound to the insertion site (green cyan, taken from bacterial Pol EC, PDB-code 2O5J [Vassylyev et al., 2007b]). Van der Waals radii are represented by colored dots. All structures were superimposed with their active site regions.

(B and C) Frayed nucleotides overlap the closed trigger loop (cyan) at residue F1084 (B, taken from the Pol II EC, PDB-code 2E2H [Wang et al., 2006]) and/or at residue H1242 (C, bacterial Pol EC, PDB-code 2O5J [Vassylyev et al., 2007b]).

(D) Frayed nucleotides overlap the tip of the hairpin of the cleavage-stimulatory factor TFIIS. The structures of EC III, EC IV, and the Pol II-TFIIS complex (PDB-code 1PQV, [Kettenberger et al., 2003]) were superimposed with their active center regions. TFIIS is shown in orange. The canonical side view is used.

(E) Detailed view of the superposition in (D) around the active site, revealing a potential clash of the TFIIS acidic hairpin with the frayed nucleotides.

To test whether fraying was due to the T•U mismatch, we replaced the mismatch in scaffold III by a correct A-U bp (Figure S1). The resulting EC V structure (Table 1, Figure 4G) revealed electron density for the RNA −1 uridine and for the phosphate of the RNA nucleotide at position +1 (Figure 4H), but not for the terminal uracile base and ribose at register +1, which are mobile. These observations suggest that two uridine residues at the RNA 3′ terminus, which are present at canonical pause sites, destabilize the bp at +1 and favor a frayed state, which can be stabilized at specific locations by a T•U mismatch at position −1 and can then be observed crystallographically.

Nucleotide-Specific Cleavage of Mismatched RNA Ends

The above results rationalize slow mismatch extension, which is a prerequisite for RNA cleavage during proofreading (Erie et al., 1993; Thomas et al., 1998). To investigate Pol II cleavage efficiency for different mismatches, we incubated the ECs used for extension assays with standard transcription buffer containing 8 mM magnesium ions (Figures 2D and 2F). Cleavage of dinucleotides was generally observed (Figure 2D) and confirmed by MALDI mass spectrometry of the RNA products (data not shown). Most efficient cleavage was observed for G•G, A•A, G•U, T•G, A•G, A•C, and G•A mismatches. RNA in the mismatched ECs was always more efficiently cleaved than in the matched ECs, and cleavage was very efficient for those mismatches that support extension (Figures 2E and 2F).

To further investigate efficient cleavage of a mismatch that is efficiently extended, we included a G•A mismatch at the end of the hybrid and solved the structure of the resulting EC VI (scaffold VI, Figure S1; Table 1). The overall structure was similar to the Pol II EC that contains the same nucleic acid scaffold except a matched G-C bp at position −1 (Kettenberger et al.,

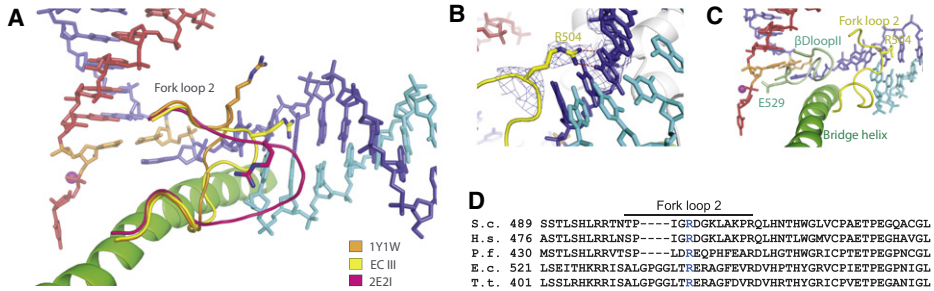

Figure 6. Fork Loop 2-Downstream DNA Contact

(A) Comparison of the conformation of fork loop 2 in EC III with that in previous Pol II EC structures (PDB-codes 1Y1W (Kettenberger et al., 2004) and 2E2I (Wang et al., 2006).

(B) Interaction of the side chain of fork loop 2 Rpb2 residue R504 with the guanine base at position +4 of downstream DNA. The final $2F_o$-F_c electron density is shown in blue, contoured at 0.7 σ.

(C) Interaction of regions in EC III that may be involved in pausing, including the frayed nucleotide, βDloopII, the bridge helix, fork loop 2, and downstream DNA.

(D) Multiple sequence alignment of fork loop 2 and surrounding Rpb2 residues from *S. cerevisiae*, *H. sapiens*, *P. furiosus*, *E. coli*, and *T. thermophilus* (CLUSTAL W). The conserved R504 from *S. cerevisiae* is highlighted in blue.

2004). However, a bromine label revealed that Pol II had apparently backtracked by two steps, although this required accommodation of A•A mismatches at positions −7 and +4 (Figure S1). Backtracking moved the templating G of the G•A mismatch from the designed position −1 to the downstream position +2. There was only fragmented electron density for the two backtracked terminal RNA nucleotides, indicating that dinucleotide cleavage had occurred prior to crystal analysis. Thus, impaired accommodation of the purine•purine mismatch in the active center apparently favors backtracking and creates the state of the EC that is prone to dinucleotide RNA cleavage, which is observed in functional assays (Figures 2D and 2F).

DISCUSSION

Mismatch-Specific Transcription Fidelity Mechanisms

Transcription fidelity relies on the abilities of Pols to select the correct nucleotide for incorporation, to impair RNA extension beyond a mismatch, and to cleave a mismatched RNA 3′ end. Here, we analyzed these three reactions in a systematic and quantitative way. We show that Pol II evolved mismatch-specific fidelity mechanisms. Mismatches that efficiently form impair RNA elongation, and mismatches that do not strongly impair RNA elongation are not formed efficiently (Figure 7A). Mismatches that are efficiently extended are also cleaved efficiently (Figure 7B), and this can be followed by efficient re-extension (Figure 7C), providing the basis for proofreading. Our misincorporation efficiencies are consistent with those reported recently (Kireeva et al., 2008) and with misincorporation opposite a template cytidine by Pol III (Alic et al., 2007). The efficiencies of misincorporation, mismatch extension, and cleavage are apparently dominated by the type of mismatch, and sequence context has a minor influence, as seen for a DNA polymerase (Joyce et al., 1992).

We also report structures of mismatch-containing Pol II ECs, which suggest three mechanisms of how misincorporation impairs RNA extension. First, a mismatch may stably bind to Pol II and disrupt the catalytically competent active site conformation. For example, a T•U mismatch can bind to the −1 position and cause loss of the catalytic metal ion A and misalignment of the RNA 3′ end. Second, a mismatch may facilitate backtracking and RNA cleavage. For example, a G•A mismatch results in a backtracked state in a crystal and in RNA dinucleotide cleavage in vitro. Third, misincorporation may result in an offline state of the EC with a frayed RNA 3′ end. Consistently, several ways of how mismatches can disrupt DNA elongation were revealed with structural studies of a DNA polymerase (Johnson and Beese, 2004). Disruption of DNA elongation by a T•T mismatch resembles the disruption of RNA elongation by a T•U mismatch observed here. Both mismatches form a wobble bp at the active center and shift the RNA 3′ end away from the catalytic site. Despite this similarity, many fidelity-determining mechanisms apparently differ for DNA and RNA polymerases, since the efficiencies of misincorporation and mismatch extension strongly differ between Pol II and DNA polymerases (Bebenek et al., 1990; Joyce et al., 1992; Kwok et al., 1990; Lai and Beattie, 1988; Mendelman et al., 1989; Mendelman et al., 1990; Perrino and Loeb, 1989; Perrino et al., 1989).

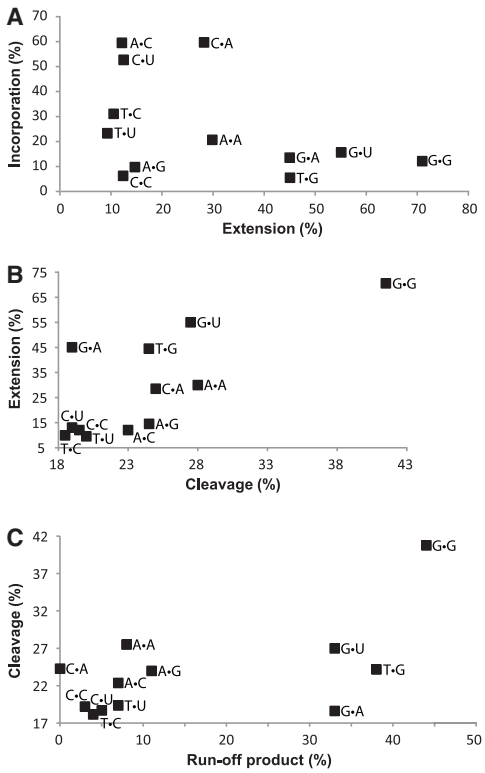

Figure 7. Correlations between Fidelity Reaction Efficiencies

(A) Correlation between misincorporation and mismatch extension efficiencies. Mean values of 1 min time points shown in Figures 2B and 2E are plotted.

(B) Correlation between efficiencies of mismatch extension (1 min time point mean values, Figure 2E) and RNA cleavage (Figure 2F).

(C) Correlation between efficiencies of RNA cleavage and run-off product formation. For run-off experiments, mismatch-containing ECs with scaffolds Z (Figure 1A) were incubated for 5 min with 100 μM of a mixture of all NTPs, and the bands corresponding to run-off products were quantified (see Experimental Procedures).

RNA Fraying, Polymerase Pausing, Backtracking, and Proofreading

Our work also provided unexpected insights into polymerase pausing, which is the first step in backtracking, proofreading, and termination and is a focal point for gene regulation (Artsimovitch and Landick, 2000; Gusarov and Nudler, 1999; Kireeva et al., 2005; Komissarova and Kashlev, 1997; Landick, 2006; Nudler et al., 1997; Palangat et al., 1998; Park et al., 2004; Reeder and Hawley, 1996). The elemental pause state is the common unactivated intermediate during elongation that results from a rearrangement of the EC that inhibits nucleotide addition without backtracking (Artsimovitch and Landick, 2000; Herbert et al., 2006; Neuman et al., 2003; Toulokhonov et al., 2007; Zhang et al., 2003). Site-directed crosslinking and mutagenesis revealed that the elemental pause contains a frayed RNA 3′-terminal nucleotide (Artsimovitch and Landick, 2000; Chan et al., 1997; Toulokhonov et al., 2007). It is possible that the elemental pause also underlies Pol II pausing at the beginning of genes that emerges as a global regulatory mechanism (Core and Lis, 2008; Erie, 2002; Gilmour, 2009; Guenther et al., 2007; Landick, 2006; Muse et al., 2007).

Structural insights into the elemental pause state are obtained from ECs III and IV. A frayed RNA 3′ nucleotide binds in two different sites in the pore that are lined by conserved residues (Table S1). Both sites overlap the NTP site and the tip of the closed trigger loop, explaining how pausing prevents NTP-coupled translocation and nucleotide addition. The EC V structure further suggests that two A-U bps at the end of the hybrid, which are obtained by transcription of a canonical pause sequence, result in a nontranslocated, nonbacktracked paused state. All structures reveal a mobile trigger loop and do not elucidate the proposed paused conformation of the trigger loop (Toulokhonov et al., 2007). EC III also reveals specific contact of R504 in the Pol II fork loop 2 and a guanine in downstream DNA, suggesting how downstream sequences may contribute to pausing (Artsimovitch and Landick, 2000; Chan et al., 1997; Holmes and Erie, 2003; Landick, 1997; Lee et al., 1990; Palangat et al., 2004; Palangat and Landick, 2001; Wang et al., 1995).

In a DNA polymerase, the 3′ cleavage rate is governed by the rate of fraying (Morales and Kool, 2000), suggesting that RNA fraying occurs during transcriptional proofreading. Since RNA cleavage generally occurs in dinucleotide steps (Izban and Luse, 1993), the polymerase must backtrack by one step after fraying. Backtracking allows the terminal nucleotide to contribute catalytic groups to the active site for cleavage stimulation (Zenkin et al., 2006). Backtracking and dinucleotide cleavage is stimulated by TFIIS, which may trigger release of a frayed nucleotide since its acidic hairpin overlaps the fraying sites (Figures 5D and 5E). TFIIS may also suppress fraying and, thus, prevent pausing by keeping the EC in the pretranslocated online state.

While this manuscript was under consideration, structures of Pol II ECs with backtracked RNA were reported (Wang et al., 2009). Thus, Pol II ECs were trapped in four different states to date: the pre- and posttranslocation states, the frayed, nonbacktracked state (this work), and the backtracked state (Wang et al., 2009). In the backtracked structures, the RNA nucleotide at position +2 occupies a "proofreading" site (whereas the terminal RNA nucleotides at register +1 in structures reported here occupy the fraying sites). Comparison of the frayed and backtracked structures suggests that the fraying site II reported here is similar to the proofreading site reported by Wang et al. This suggests that proofreading generally begins with conversion of the pretranslocated EC to a nonbacktracked, paused EC with a frayed RNA end. Subsequent backtracking by one step moves the bp of the penultimate RNA nucleotide to register −1 but may maintain the terminal nucleotide in or near the fraying site. This enables cleavage of an RNA dinucleotide, resulting in a new 3′-hydroxyl group at the catalytic site from which elongation can resume.

Nucleotide Selectivity

In the future, the molecular basis of nucleotide selectivity may be analyzed with EC structures containing mismatched NTP substrates, although such structures are very difficult to obtain for Pol II (Brueckner et al., 2007) and also for single-subunit DNA polymerases (Batra et al., 2008). NTP selectivity is likely governed by similar mechanisms in DNA and RNA polymerases because it involves an induced fit and closure of the active center in both cases (Arora et al., 2005; Kaplan et al., 2008; Kireeva et al., 2008; Krahn et al., 2004; Kuchta et al., 1988; Kuchta et al., 1987; Sawaya et al., 1997; Washington et al., 2001; Wong et al., 1991). The closed active center may accommodate small template-NTP pyrimidine•pyrimidine mismatches, explaining the facilitated formation of T•U, T•C, and C•U mismatches (Figure 2B). Other misincorporations may, however, occur via an EC intermediate that lacks a DNA base in the templating site, as suggested by recent structures of a DNA polymerase with a mismatched NTP (Batra et al., 2008). Indeed, an empty templating site was observed in a recent Pol II EC intermediate structure (Brueckner and Cramer, 2008), and Pol II can misincorporate opposite an abasic template site (Damsma et al., 2007) and likely also opposite an empty templating site that results from a failure to translocate a bulky DNA lesion into the active site (Brueckner et al., 2007; Damsma et al., 2007).

EXPERIMENTAL PROCEDURES

Improved Pol II Purification Protocol

S. cerevisiae Pol II containing a hexahistidine-tagged Rpb3 subunit (strain kindly provided by the laboratory of M. Kashlev) was purified as described (Kireeva et al., 2003), but with several significant modifications. Briefly, 150 g of cell pellet were resuspended in 50 mM Tris-HCl (pH 7.9), 1 mM EDTA, 10 μM $ZnCl_2$, 10% v/v glycerol, 1% v/v DMSO, 10 mM DTT, 1× protease inhibitors (100 × protease inhibitor mix: 1.42 mg Leupeptin, 6.85 mg Pepstatin A, 850 mg PMSF, 1650 mg benzamidine; dry ethanol ad 50 ml), and were lysed by bead beating for 80 min using intervals of 30 s followed by 90 s pauses. The lysate was cleared by centrifugation and ultracentrifugation. Pol II was precipitated by the addition of 50% saturated ammonium sulfate solution. The pellet was dissolved in Ni buffer (20 mM Tris-HCl [pH 7.9], 150 mM KCl, 10 μM $ZnCl_2$, 10% v/v glycerol, 10 mM DTT, and 1× protease inhibitors) and subjected to Ni-NTA affinity chromatography (2 × 8 ml fresh Ni-NTA) using gravity flow. After washing with high-salt buffer (Ni buffer with 1000 mM KCl and 7 mM imidazole), and with Ni7 buffer (20 mM Tris-HCl [pH 7.9], 150 mM KCl, 7 mM imidazole, 10 μM $ZnCl_2$, 10 mM DTT, and 1× protease inhibitors), the protein was eluted with Ni7 buffer containing 100 mM imidazole and no protease inhibitors. The eluted protein was diluted with MonoQ buffer (20 mM Tris-acetate [pH 7.9], 0.5 mM EDTA, 10 μM $ZnCl_2$, 10% v/v glycerol, and 10 mM DTT) and subjected to anion exchange chromatography (MonoQ, GE healthcare) using a gradient from 150 mM to 1500 mM KOAc. The last elution peak (at a conductivity of 50 mS/cm) was collected and concentrated. The concentrated Pol II was precipitated by the addition of 50% ammonium sulfate, and the pellets were stored at −80°C.

EC Assembly

For the bead-based assays, the ECs containing complete complementary scaffolds were assembled essentially as described (Kireeva et al., 2003). Briefly, the DNA nontemplate was 5′-end-labeled with Biotin with the use of a TTTTT linker. The RNA was 5′-end-labeled with 6-carboxyfluoresceine (FAM). For EC assembly, Pol II was incubated with a hybrid of the DNA template strand annealed to the RNA (2-fold excess) in transcription buffer (TB, 20 mM HEPES [pH 7.6], 60 mM $[NH_4]_2SO_4$, 8 mM $MgSO_4$, 10 μM $ZnCl_2$, 10% v/v glycerol, and 10 mM DTT) for 15 min at 20°C, subsequently with the biotinylated nontemplate DNA strand (4-fold excess) for 10 min at 25°C, and then with recombinant Rpb4/7 (5-fold excess) for 10 min at 25°C.

Bead-Based RNA Extension and Cleavage Assays

Bead-based assays were carried out as described with some modifications (S.D. and P.C., unpublished data). Briefly, beads (Dynabeads MyOne Streptavidin T1 from Invitrogen) were added to ECs for assembly and incubated for 30 min at 25°C. Beads were subsequently washed with TB containing 0.1% Triton-X, TB containing 0.2 M $(NH_4)_2SO_4$, and with TB. Beads were resuspended in TB. For RNA extension assays including time course experiments, different amounts of NTPs (Jena Bioscience) were added, the mixture was

incubated at 28°C and reactions were stopped at different time points by addition of an equal volume of 100 mM EDTA, essentially as described (Brueckner et al., 2007). For cleavage assays, the bead-coupled ECs were incubated at 28°C in TB for 5 min and stopped as described for extension assays. The beads were transferred into urea loading buffer, and samples were heated to 95°C and loaded on a 20% polyacrylamid gel containing 7 M Urea. The FAM 5′-labeled RNA products were visualized with a Typhoon 9400 scanner (GE Healthcare). Gel bands were quantified using ImageQuant (GE healthcare). In case more than one product was observed (A), the amounts of different RNA products were added up. For MALDI-TOF analysis, the reaction was incubated, stopped and analyzed as described (Brueckner et al., 2007). NTP samples were analyzed by reverse phase HPLC analysis, and no cross-contamination with other NTPs was detected. We also requested analytic data from the supplier, which showed that the NTPs are 99.8% pure and the remaining impurities are NDPs and NMP of the same kind, but not other types of NTPs. Since the NTPs are synthesized de novo and are not derived from fractionation of an NTP pool, cross-contamination cannot occur. We are, therefore, certain that misincorporation took place.

Crystal Structure Determinations

The match- and mismatch-containing scaffolds were cocrystallized and the structures were determined essentially as described (Brueckner et al., 2007), with minor changes. The crystallization solution lacked magnesium ions (200 mM ammonium acetate, 300 mM sodium acetate, 50 mM HEPES [pH 7.0], 4%–6% w/v PEG 6000, and 5 mM TCEP). Diffraction data of EC I were collected at the beamline X06A of the Swiss Light Source using a mar225 CCD detector, whereas data of ECs II-VI were collected using a PILATUS 6M pixel detector (Broennimann et al., 2006) (Table 1). Raw data of EC I were processed with HKL2000, data of ECs II-VI with XDS (Kabsch, 1993). The structure of EC I and VI were solved by molecular replacement with the program PHASER (McCoy et al., 2005), using the structure of the complete 12-subunit Pol II without nucleic acids as a search model (PDB 1Y1W) (Kettenberger et al., 2004). The higher resolution of the EC I crystal produced a superior model of the protein compared with 1Y1W, as judged by the quality of the Ramachandran plot and R_{cryst}/R_{free} values. When used as a search model for molecular replacement, EC I resulted in better quality models in the determination of the EC II, III, IV, and V structures. The molecular-replacement solution was subjected to rigid-body refinement with CNS version 1.2 (Brünger et al., 1998). Model building was done with Coot (Emsley and Cowtan, 2004) and Moloc (Gerber Molecular Design, Switzerland, http://www.moloc.ch). The nucleic acids were built stepwise into unbiased F_o-F_c electron density. The register of the nucleic acids was unambiguously defined by bromine labeling as described (Brueckner et al., 2007). Refinement of ECs II-VI was monitored with the free R-factor calculated from the same set of excluded reflections as in the refinement of the complete Pol II complex (Armache et al., 2005) and the complete Pol II EC (Brueckner et al., 2007; Damsma et al., 2007; Kettenberger et al., 2004). Due to the different space group and higher resolution of EC I, a new test set of reflections was generated.

ACCESSION NUMBERS

Coordinates and structure factors have been deposited in the Protein Data Bank with accession codes 3HOU, 3HOV, 3HOW, 3HOZ, 3HOX, and 3HOY for ECs I, II, III, IV, V, and VI, respectively.

SUPPLEMENTAL DATA

Supplemental Data include one table and one figure and can be found with this article online at http://www.cell.com/molecular-cell/supplemental/S1097-2765(09)00393-1.

ACKNOWLEDGMENTS

We thank Thomas Fröhlich for help with MALDI experiments. We thank Dirk Kostrewa and other members of the Cramer laboratory. J.F.S. was supported by a Cusanuswerk fellowship. P.C. was supported by the Deutsche Forschungsgemeinschaft, the Sonderforschungsbereich SFB646, the SFB TR5, the EU research grant network 3D Repertoire, LMU*excellent*, the Nanosystems Initiative Munich NIM, the Elitenetzwerk Bayern graduate programme NanoBioTechnology, and the Fonds der chemischen Industrie. D.V. was supported in part by NIH grants GM74252 and GM74840 (to D.G.V.). Part of this work was conducted at the protein crystallography beamline PXI of the Swiss Light Source (SLS) at the Paul Scherrer Institute, Villigen, Switzerland. We thank C. Schulze-Briese and his team at the SLS for help.

Received: March 16, 2009
Revised: May 5, 2009
Accepted: June 5, 2009
Published: June 25, 2009

REFERENCES

Alic, N., Ayoub, N., Landrieux, E., Favry, E., Baudouin-Cornu, P., Riva, M., and Carles, C. (2007). Selectivity and proofreading both contribute significantly to the fidelity of RNA polymerase III transcription. Proc. Natl. Acad. Sci. USA *104*, 10400–10405.

Armache, K.-J., Mitterweger, S., Meinhart, A., and Cramer, P. (2005). Structures of complete RNA polymerase II and its subcomplex Rpb4/7. J. Biol. Chem. *280*, 7131–7134.

Arora, K., Beard, W.A., Wilson, S.H., and Schlick, T. (2005). Mismatch-induced conformational distortions in polymerase beta support an induced-fit mechanism for fidelity. Biochemistry *44*, 13328–13341.

Artsimovitch, I., and Landick, R. (2000). Pausing by bacterial RNA polymerase is mediated by mechanistically distinct classes of signals. Proc. Natl. Acad. Sci. USA *97*, 7090–7095.

Batra, V.K., Beard, W.A., Shock, D.D., Pedersen, L.C., and Wilson, S.H. (2008). Structures of DNA polymerase beta with active-site mismatches suggest a transient abasic site intermediate during misincorporation. Mol. Cell *30*, 315–324.

Bebenek, K., Joyce, C.M., Fitzgerald, M.P., and Kunkel, T.A. (1990). The fidelity of DNA synthesis catalyzed by derivatives of Escherichia coli DNA polymerase I. J. Biol. Chem. *265*, 13878–13887.

Broennimann, C., Eikenberry, E.F., Henrich, B., Horisberger, R., Huelsen, G., Pohl, E., Schmitt, B., Schulze-Briese, C., Suzuki, M., Tomizaki, T., et al. (2006). The PILATUS 1M detector. J. Synchrotron Radiat. *13*, 120–130.

Brueckner, F., and Cramer, P. (2008). Structural basis of transcription inhibition by alpha-amanitin and implications for RNA polymerase II translocation. Nat. Struct. Mol. Biol. *15*, 811–818.

Brueckner, F., Hennecke, U., Carell, T., and Cramer, P. (2007). CPD damage recognition by transcribing RNA polymerase II. Science *315*, 859–862.

Brünger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. (1998). Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Crystallogr. D Biol. Crystallogr. *54*, 905–921.

Chan, C.L., Wang, D., and Landick, R. (1997). Multiple interactions stabilize a single paused transcription intermediate in which hairpin to 3′ end spacing distinguishes pause and termination pathways. J. Mol. Biol. *268*, 54–68.

Core, L.J., and Lis, J.T. (2008). Transcription regulation through promoter-proximal pausing of RNA polymerase II. Science *319*, 1791–1792.

Damsma, G.E., Alt, A., Brueckner, F., Carell, T., and Cramer, P. (2007). Mechanism of transcriptional stalling at cisplatin-damaged DNA. Nat. Struct. Mol. Biol. *14*, 1127–1133.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. *60*, 2126–2132.

Erie, D.A. (2002). The many conformational states of RNA polymerase elongation complexes and their roles in the regulation of transcription. Biochim. Biophys. Acta *1577*, 224–239.

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

Gilmour, D.S. (2009). Promoter proximal pausing on genes in metazoans. Chromosoma *118*, 1–10.

Guenther, M.G., Levine, S.S., Boyer, L.A., Jaenisch, R., and Young, R.A. (2007). A chromatin landmark and transcription initiation at most promoters in human cells. Cell *130*, 77–88.

Gusarov, I., and Nudler, E. (1999). The mechanism of intrinsic transcription termination. Mol. Cell *3*, 495–504.

Herbert, K.M., La Porta, A., Wong, B.J., Mooney, R.A., Neuman, K.C., Landick, R., and Block, S.M. (2006). Sequence-resolved detection of pausing by single RNA polymerase molecules. Cell *125*, 1083–1094.

Holmes, S.F., and Erie, D.A. (2003). Downstream DNA sequence effects on transcription elongation. Allosteric binding of nucleoside triphosphates facilitates translocation via a ratchet motion. J. Biol. Chem. *278*, 35597–35608.

Izban, M.G., and Luse, D.S. (1993). SII-facilitated transcript cleavage in RNA polymerase II complexes stalled early after initiation occurs in primarily dinucleotide increments. J. Biol. Chem. *268*, 12864–12873.

Jasiak, A.J., Armache, K.J., Martens, B., Jansen, R.P., and Cramer, P. (2006). Structural biology of RNA polymerase III: subcomplex C17/25 X-ray structure and 11 subunit enzyme model. Mol. Cell *23*, 71–81.

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

Joyce, C.M., Sun, X.C., and Grindley, N.D. (1992). Reactions at the polymerase active site that contribute to the fidelity of Escherichia coli DNA polymerase I (Klenow fragment). J. Biol. Chem. *267*, 24485–24500.

Kabsch, W. (1993). Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. J. Appl. Cryst. *26*, 795–800.

Kaplan, C.D., Larsson, K.M., and Kornberg, R.D. (2008). The RNA polymerase II trigger loop functions in substrate selection and is directly targeted by alpha-amanitin. Mol. Cell *30*, 547–556.

Kashkina, E., Anikin, M., Brueckner, F., Pomerantz, R.T., McAllister, W.T., Cramer, P., and Temiakov, D. (2006). Template misalignment in multisubunit RNA polymerases and transcription fidelity. Mol. Cell *24*, 257–266.

Kettenberger, H., Armache, K.-J., and Cramer, P. (2003). Architecture of the RNA polymerase II-TFIIS complex and implications for mRNA cleavage. Cell *114*, 347–357.

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.

Kireeva, M.L., Hancock, B., Cremona, G.H., Walter, W., Studitsky, V.M., and Kashlev, M. (2005). Nature of the Nucleosomal Barrier to RNA Polymerase II. Mol. Cell *18*, 97–108.

Kireeva, M.L., Lubkowska, L., Komissarova, N., and Kashlev, M. (2003). Assays and affinity purification of biotinylated and nonbiotinylated forms of double-tagged core RNA polymerase II from Saccharomyces cerevisiae. Methods Enzymol. *370*, 138–155.

Kireeva, M.L., Nedialkov, Y.A., Cremona, G.H., Purtov, Y.A., Lubkowska, L., Malagon, F., Burton, Z.F., Strathern, J.N., and Kashlev, M. (2008). Transient reversal of RNA polymerase II active site closing controls fidelity of transcription elongation. Mol. Cell *30*, 557–566.

Komissarova, N., and Kashlev, M. (1997). RNA polymerase switches between inactivated and activated states by translocating back and forth along the DNA and the RNA. J. Biol. Chem. *272*, 15329–15338.

Krahn, J.M., Beard, W.A., and Wilson, S.H. (2004). Structural insights into DNA polymerase beta deterrents for misincorporation support an induced-fit mechanism for fidelity. Structure *12*, 1823–1832.

Kuchta, R.D., Mizrahi, V., Benkovic, P.A., Johnson, K.A., and Benkovic, S.J. (1987). Kinetic mechanism of DNA polymerase I (Klenow). Biochemistry *26*, 8410–8417.

Kuchta, R.D., Benkovic, P., and Benkovic, S.J. (1988). Kinetic mechanism whereby DNA polymerase I (Klenow) replicates DNA with high fidelity. Biochemistry *27*, 6716–6725.

Kuhn, C.D., Geiger, S.R., Baumli, S., Gartmann, M., Gerber, J., Jennebach, S., Mielke, T., Tschochner, H., Beckmann, R., and Cramer, P. (2007). Functional architecture of RNA polymerase I. Cell *131*, 1260–1272.

Kunkel, T.A., and Bebenek, K. (2000). DNA replication fidelity. Annu. Rev. Biochem. *69*, 497–529.

Kwok, S., Kellogg, D.E., McKinney, N., Spasic, D., Goda, L., Levenson, C., and Sninsky, J.J. (1990). Effects of primer-template mismatches on the polymerase chain reaction: human immunodeficiency virus type 1 model studies. Nucleic Acids Res. *18*, 999–1005.

Lai, M.D., and Beattie, K.L. (1988). Influence of DNA sequence on the nature of mispairing during DNA synthesis. Biochemistry *27*, 1722–1728.

Landick, R. (1997). RNA polymerase slides home: pause and termination site recognition. Cell *88*, 741–744.

Landick, R. (2006). The regulatory roles and mechanism of transcriptional pausing. Biochem. Soc. Trans. *34*, 1062–1066.

Lee, D.N., Phung, L., Stewart, J., and Landick, R. (1990). Transcription pausing by *Escherichia coli* RNA polymerase is modulated by downstream DNA sequences. J. Biol. Chem. *265*, 15145–15153.

McCoy, A.J., Grosse-Kunstleve, R.W., Storoni, L.C., and Read, R.J. (2005). Likelihood-enhanced fast translation functions. Acta Crystallogr. D Biol. Crystallogr. *61*, 458–464.

McCulloch, S.D., and Kunkel, T.A. (2008). The fidelity of DNA synthesis by eukaryotic replicative and translesion synthesis polymerases. Cell Res. *18*, 148–161.

Mendelman, L.V., Boosalis, M.S., Petruska, J., and Goodman, M.F. (1989). Nearest neighbor influences on DNA polymerase insertion fidelity. J. Biol. Chem. *264*, 14415–14423.

Mendelman, L.V., Petruska, J., and Goodman, M.F. (1990). Base mispair extension kinetics. Comparison of DNA polymerase alpha and reverse transcriptase. J. Biol. Chem. *265*, 2338–2346.

Morales, J.C., and Kool, E.T. (2000). Importance of terminal base pair hydrogen-bonding in 3′-end proofreading by the Klenow fragment of DNA polymerase I. Biochemistry *39*, 2626–2632.

Muse, G.W., Gilchrist, D.A., Nechaev, S., Shah, R., Parker, J.S., Grissom, S.F., Zeitlinger, J., and Adelman, K. (2007). RNA polymerase is poised for activation across the genome. Nat. Genet. *39*, 1507–1511.

Naji, S., Bertero, M.G., Spitalny, P., Cramer, P., and Thomm, M. (2008). Structure function analysis of the RNA polymerase cleft loops elucidates initial transcription, DNA unwinding and RNA displacement. Nucleic Acids Res. *36*, 676–687. Published online December 10, 2007. 10.1093/nar/gkm1086.

Nesser, N.K., Peterson, D.O., and Hawley, D.K. (2006). RNA polymerase II subunit Rpb9 is important for transcriptional fidelity in vivo. Proc. Natl. Acad. Sci. USA *103*, 3268–3273.

Neuman, K.C., Abbondanzieri, E.A., Landick, R., Gelles, J., and Block, S.M. (2003). Ubiquitous transcriptional pausing is independent of RNA polymerase backtracking. Cell *115*, 437–447.

Nudler, E., Mustaev, A., Lukhtanov, E., and Goldfarb, A. (1997). The RNA-DNA hybrid maintains the register of transcription by preventing backtracking of RNA polymerase. Cell *89*, 38–41.

Palangat, M., and Landick, R. (2001). Roles of RNA:DNA hybrid stability, RNA structure, and active site conformation in pausing by human RNA polymerase II. J. Mol. Biol. *311*, 265–282.

Palangat, M., Meier, T.I., Keene, R.G., and Landick, R. (1998). Transcriptional pausing at +62 of HIV-1 nascent RNA modulates formation of the TAR RNA structure. Mol. Cell *1*, 1033–1042.

Palangat, M., Hittinger, C.T., and Landick, R. (2004). Downstream DNA selectively affects a paused conformation of human RNA polymerase II. J. Mol. Biol. *341*, 429–442.

Park, N.J., Tsao, D.C., and Martinson, H.G. (2004). The two steps of poly(A)-dependent termination, pausing and release, can be uncoupled by truncation

of the RNA polymerase II carboxyl-terminal repeat domain. Mol. Cell. Biol. *24*, 4092–4103.

Perrino, F.W., and Loeb, L.A. (1989). Differential extension of 3′ mispairs is a major contribution to the high fidelity of calf thymus DNA polymerase-alpha. J. Biol. Chem. *264*, 2898–2905.

Perrino, F.W., Preston, B.D., Sandell, L.L., and Loeb, L.A. (1989). Extension of mismatched 3′ termini of DNA is a major determinant of the infidelity of human immunodeficiency virus type 1 reverse transcriptase. Proc. Natl. Acad. Sci. USA *86*, 8343–8347.

Reeder, T.C., and Hawley, D.K. (1996). Promoter proximal sequences modulate RNA polymerase II elongation by a novel mechanism. Cell *87*, 767–777.

Sawaya, M.R., Prasad, R., Wilson, S.H., Kraut, J., and Pelletier, H. (1997). Crystal structures of human DNA polymerase beta complexed with gapped and nicked DNA: Evidence for an induced fit mechanism. Biochemistry *36*, 11205–11215.

Saxowsky, T.T., and Doetsch, P.W. (2006). RNA polymerase encounters with DNA damage: transcription-coupled repair or transcriptional mutagenesis? Chem. Rev. *106*, 474–488.

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.

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

Toulokhonov, I., Zhang, J., Palangat, M., and Landick, R. (2007). A central role of the RNA polymerase trigger loop in active-site rearrangement during transcriptional pausing. Mol. Cell *27*, 406–419.

Vassylyev, D.G., Vassylyeva, M.N., Perederina, A., Tahirov, T.H., and Artsimovitch, I. (2007a). Structural basis for transcription elongation by bacterial RNA polymerase. Nature *448*, 157–162.

Vassylyev, D.G., Vassylyeva, M.N., Zhang, J., Palangat, M., Artsimovitch, I., and Landick, R. (2007b). Structural basis for substrate loading in bacterial RNA polymerase. Nature *448*, 163–168.

Wang, D., Meier, T.I., Chan, C.I., Feng, G., Lee, D.N., and Landick, R. (1995). Discontinuous movements of DNA and RNA in RNA polymerase accompany formation of a paused transcription complex. Cell *81*, 341–350.

Wang, D., Bushnell, D.A., Westover, K.D., Kaplan, C.D., and Kornberg, R.D. (2006). Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis. Cell *127*, 941–954.

Wang, D., Bushnell, D.A., Huang, X., Westover, K.D., Levitt, M., and Kornberg, R.D. (2009). Structural Basis of Transcription: Backtracked RNA Polymerase II at 3.4 Angstrom Resolution. Science *324*, 1203–1206.

Washington, M.T., Prakash, L., and Prakash, S. (2001). Yeast DNA polymerase eta utilizes an induced-fit mechanism of nucleotide incorporation. Cell *107*, 917–927.

Wong, I., Patel, S.S., and Johnson, K.A. (1991). An induced-fit kinetic mechanism for DNA replication fidelity: direct measurement by single-turnover kinetics. Biochemistry *30*, 526–537.

Zenkin, N., Yuzenkova, Y., and Severinov, K. (2006). Transcript-assisted transcriptional proofreading. Science *313*, 518–520.

Zhang, C., Yan, H., and Burton, Z.F. (2003). Combinatorial control of human RNA polymerase II (RNAP II) pausing and transcript cleavage by transcription factor IIF, hepatitis delta antigen, and stimulatory factor II. J. Biol. Chem. *278*, 50101–50111.