Topology of Yeast RNA Polymerase II Subunits in Transcription Elongation Complexes Studied by Photoaffinity Cross-Linking†

Christine I. Wooddell‡ and Richard R. Burgess*

McArdle Laboratory for Cancer Research, University of Wisconsin−Madison, Madison, Wisconsin 53706

Received June 21, 2000; Revised Manuscript Received August 30, 2000

ABSTRACT: The subunits of *Saccharomyces cerevisiae* RNA polymerase II (RNAP II) in proximity to the DNA during transcription elongation have been identified by photoaffinity cross-linking. In the absence of transcription factors, RNAP II will transcribe a double-stranded DNA fragment containing a 3′-extension of deoxycytidines, a "tailed template". We designed a DNA template allowing the RNAP to transcribe 76 bases before it was stalled by omission of CTP in the transcription reaction. This stall site oriented the RNAP on the DNA template and allowed us to map the RNAP subunits along the DNA. The DNA analogue 5-[*N*-(*p*-azidobenzoyl)-3-aminoallyl]-dUTP (N_3RdUTP) [Bartholomew, B., Kassavetis, G. A., Braun, B. R., and Geiduschek, E. P. (1990) *EMBO J. 9*, 2197−205] was synthesized and enzymatically incorporated into the DNA at specified positions upstream or downstream of the stall site, in either the template or nontemplate strand of the DNA. Radioactive nucleotides were positioned beside the photoactivatable nucleotides, and cross-linking by brief ultraviolet irradiation transferred the radioactive tag from the DNA onto the RNAP subunits. In addition to N_3RdUTP, which has a photoreactive azido group 9 Å from the uridine base, we used the photoaffinity cross-linker $5N_3$dUTP with an azido group directly on the uridine ring to identify the RNAP II subunits closest to the DNA at positions where multiple subunits cross-linked. In cross-linking reactions dependent on transcription, RPB1, RPB2, and RPB5 were cross-linked with N_3RdUTP. With $5N_3$dUTP, only RPB1 and RPB2 were cross-linked. Under certain circumstances, RPB3, RPB4, and RPB7 were cross-linked. From the information obtained in this topological study, we developed a model of yeast RNAP II in a transcription elongation complex.

RNA polymerase II (RNAP II)[1] is the nuclear DNA-dependent RNAP that transcribes pre-mRNA from genes that encode the proteins in a eukaryotic organism. RNAP II transcription is a cyclic process consisting of binding of the RNAP at a promoter to form an initiation complex, initiation of transcription, promoter clearance, transcript elongation, and termination. Regulation of gene expression can occur at each of these steps and for many eukaryotic genes is at the level of elongation (*2*−*5*).

Saccharomyces cerevisiae RNAP II is a large enzyme with a molecular mass of 550 kDa and is composed of 12 subunits ranging in size from 220 to ∼8 kDa (*6*−*9*). All the subunits have been cloned. Five of the smaller subunits, RPB5, RPB6, RPB8, RPB10, and RPB12, are common to yeast RNAP I, II, and III (*10*−*13*). RNAP II structure, function, and amino acid sequences of the subunits are highly conserved among species (*7*, *8*, *14*−*21*). Six of the human subunits, homologues of yeast RPB6−RPB10 and RPB12, can functionally replace their yeast counterparts (*22*−*25*). What we learn from the yeast RNAP II will provide insights into understanding the human enzyme.

The largest two RNAP II subunit have been shown to bind DNA by Southwestern assays (DNA probing of RNAP subunits separated via SDS−PAGE and blotted), a filter binding assay, and antibody inhibition of DNA binding (*26*−*31*). The sequence of the demonstrated DNA-binding region in *Drosophila* RPB1 is similar to the region in *Escherichia coli* DNA polymerase I that is part of the cleft suggested to interact with the double-stranded DNA being synthesized (*29*). DNA and RNA polymerases have structurally similar active sites, and substantial evidence suggests they use a common mechanism for polymerization (*32*−*42*). Horikoshi et al. (*27*) demonstrated that a subassembly of the three largest subunits of mouse RNAP II (homologous to RPB1−RPB3) could be sedimented with DNA in a glycerol gradient. This subassembly was obtained by partial dissociation in urea of the intact enzyme. With a filter binding assay, Kimura et al. (*43*) detected DNA binding by a subassembly of RPB2, RPB3, and RPB11 of the fission yeast, *Schizosaccharomyces pombe*, although this binding was much weaker than that of RPB1. A subassembly of only RPB1−RPB3 and RPB11 from *S. cerevisiae* was found to have weak transcription activity (V. Svetlov and R. Burgess, unpublished results).

† This work was supported by NIH Grants T32-GM07215, T32-CA09135, and CA60896.

* To whom correspondence should be addressed. E-mail: burgess@oncology.wisc.edu. Phone: (608) 263-2635.

‡ Current address: Mirus Corp., 505 S. Rosa Rd., Madison, WI 53719.

[1] Abbreviations: aPCR, asymmetric polymerase chain reaction; CTP, cytidine 5′-triphosphate; dA, 2′-deoxyadenosine; DTT, dithiothreitol; dNTP, 2′-deoxynucleoside 5′-triphosphate; dTTP, thymidine 5′-triphosphate; GTP, guanosine 5′-triphosphate; KOAc, potassium acetate; N_3RdUTP, 5-[*N*-(*p*-azidobenzoyl)-3-aminoallyl]-dUTP; $5N_3$dUTP, 5-azidouridine triphosphate; NTP, nucleoside triphosphate; RNAP, RNA polymerase; RPB1−RPB12, yeast RNA polymerase II subunits; RPC, yeast RNA polymerase III subunit; TE, 10 mM Tris-HCl (pH 7.9) and 0.1 mM EDTA.

10.1021/bi0014249 CCC: $19.00
Published on Web 10/10/2000

The structure of yeast RNAP II at 5 and 3 Å resolution was recently determined by X-ray crystallographic analysis (*44*, *49*), and this structure resembles that of *E. coli* core RNAP (*45*) and *Thermus aquaticus* core RNAP (*42*). Upon crystallization of the *E. coli* holoenzyme (*46*), a dramatic difference in conformation was noted between the form of the RNAP that engages in transcription elongation, core, and the form of the RNAP that initiates transcription, holoenzyme. Yeast RNAP II has also been crystallized in two different conformations resembling the *E. coli* core RNAP and holoenzyme (*44*, *47*, *48*). These conformational differences are attributed to two flexible protein domains: an arm that surrounds a 25 Å DNA-binding channel and a hinged domain that may enclose the RNA−DNA hybrid (*44*, *48*, *49*). Subunits RPB4 and RPB7 fill a cavity in the Δ4/7 enzyme (i.e., enzyme lacking subunits RPB4 and RPB7) near, but not in, the DNA-binding cleft (ref *50* and R. Kornberg, personal communication). Regions of RPB1 and RPB2 are close in three-dimensional space so that they form the active center (*51*−*57*).

To determine which of the 12 yeast RNAP II subunits are close to the DNA while the RNAP is in a transcriptionally active form, we allowed the RNAP to initiate transcription from tailed templates with 3′-extensions of polydeoxycytidines (*58*) and then to transcribe 76 bases before being stalled because of the omission of CTP from the reaction mixture. Photoaffinity cross-linkers were positioned in the DNA with respect to the stall site to map the relative locations of RNAP subunits along the DNA. Elongation complexes initiated on tailed templates have been used to study elongation, termination, and response of the ternary complex to elongation factors or intrinsic DNA sequences (*58*−*67*). Transcription of cross-linker-containing tailed templates allowed us to examine the topology of RNAP II in the absence of any other transcription factors.

Photo-cross-linking of ^{3}H-labeled DNA by transcribing *Drosophila* RNAP I, II, or III labeled only the largest two subunits (*26*). By using a photoaffinity cross-linker with an azido group about 9 Å from the uridine, however, Bartholomew et al. (*68*) were able to detect eight or nine subunits of yeast RNAP III close to the DNA (six in initiation complexes and seven in ternary complexes that had transcribed 17 bases). Azido-containing thymidine derivatives are more reactive toward photolysis than are natural thymidine residues; the former react with proximal protein residues by means of a reactive nitrene group, which is generated upon ultraviolet irradiation (*69*, *70*). Using an azido cross-linker incorporated into the phosphate backbone of oligodeoxyribonucleotides, Kim et al. (*71*) evaluated the human RNAP II subunits that cross-linked to DNA in transcription initiation complexes. RPB1, RPB2, and RPB5 were cross-linked to the DNA.

Photoaffinity cross-linking has been used to study the protein−nucleic acid interface of single- or multisubunit proteins that interact with nucleic acids, as well as to quantify their affinity (*1*, *68*−*70*, *72*−*80*). In this study, we enzymatically incorporated into specific sites in the DNA the photoactivatable thymidine analogue 5-[*N*-(*p*-azidobenzoyl)-3-aminoallyl]-dUTP (N_3RdUTP) developed by Bartholomew et al. (*1*) and then cross-linked the DNA to nearby proteins by brief ultraviolet irradiation. Radioactive nucleotides incorporated next to the cross-linking analogue transferred a radioactive tag to the protein. This tag was then trimmed to a few nucleotides in length by nuclease treatment. The reactive azido group of N_3RdUTP projects about 9 Å from the 5′-position of the uridine base in the major groove and the surrounding area, thus allowing it to probe proteins nearby, but not necessarily in tight association with the DNA (*68*). In addition to N_3RdUTP, we utilized the photoaffinity cross-linker $5N_3$dUTP, which has the azido group directly on the uridine ring, to identify the RNAP II subunits closest to the DNA at positions where multiple subunits cross-linked. We used the two azido cross-linkers to determine (a) which RNAP II subunits are within 9 Å of the DNA along the footprinted region around the transcription stall site, (b) which subunit is closest to the DNA at each evaluated position, and (c) what one can learn about the topology of the RNAP−DNA interface by measuring the relative amount of cross-linking of DNA at each position to RNAP II subunits.

EXPERIMENTAL PROCEDURES

Enzymes and Reagents. N_3RdUTP was synthesized as described by Bartholomew et al. (*1*). $5N_3$dUTP was from Research Products International Corp. (Mt. Prospect, IL). Restriction enzymes and DNA-modifying enzymes were from New England Biolabs, Promega, or Gibco-BRL/Life Technologies unless otherwise indicated. Sequenase 2.0 enzyme and RNase H were from U.S. Biochemical, and heparin was from Calbiochem, glycogen from Boehringer Mannheim, and micrococcal nuclease from Pharmacia. Terminal deoxynucleotidyl transferase (TdT) and all restriction enzymes used in reactions with azido cross-linkers were from Gibco-BRL and in buffers containing 2-mercaptoethanol (2-ME) rather than DTT. Aryl azides are more rapidly reduced by DTT than by 2-ME at higher concentrations of the reducing agent (*81*, *82*), but neither 1 mM 2-ME nor 0.1 mM DTT reduced the azido group during a 50 min exposure (*83*). TdT buffer was 0.1 mM DTT, but did not inactivate the cross-linker during the 20 min tailing reaction. DNase I was RNase-free and FPLC-pure and was from Pharmacia. NTPs and dNTPs were Ultrapure Solutions (Amersham Pharmacia Biotech). Yeast RNA polymerase II was immunoaffinity purified with monoclonal antibody 8WG16 conjugated to Sepharose (*84*, *85*) and then purified further on a Mono-Q column (Pharmacia; K. Foley, K. J. Chambers, and L. Strasheim, unpublished results). The concentration of the RNAP II was typically 1 mg/mL in 50 mM Tris-HCl (pH 7.9), 0.1 mM EDTA, 0.1 mM DTT, 20% glycerol, and ∼0.5 M NaCl.

Oligodeoxyribonucleotides. Oligos were synthesized by the University of Wisconsin Biotechnology Center (Madison, WI) or Life Technologies. The following oligos were used for cloning. M13 Reverse is 5′-CAGGAAACAGCTATGAC-3′. M13 Forward is 5′-CGCCAGGGTTTTCCCAGTCAC-GAC-3′. At*Eco*RI is 5′-AATTCGGGGGGGGGGGGGG-3′. *Eco*RIAb is 5′-CCCCCCCCCCCCCCG-3′. Figure 1 shows cross-linker positions and oligos used to define the upstream and downstream ends of the cross-linker-containing DNAs. The following downstream oligos were used to define the 3′-ends of long nontemplate strand primers generated by asymmetric PCR (aPCR) (*86*). b63 (5′-CACCCAATACTC-CCTACTCATCTCA-3′) was used to place the cross-linker at −20 nontemplate (−20N), b74 (5′-CACCCCCTACT-

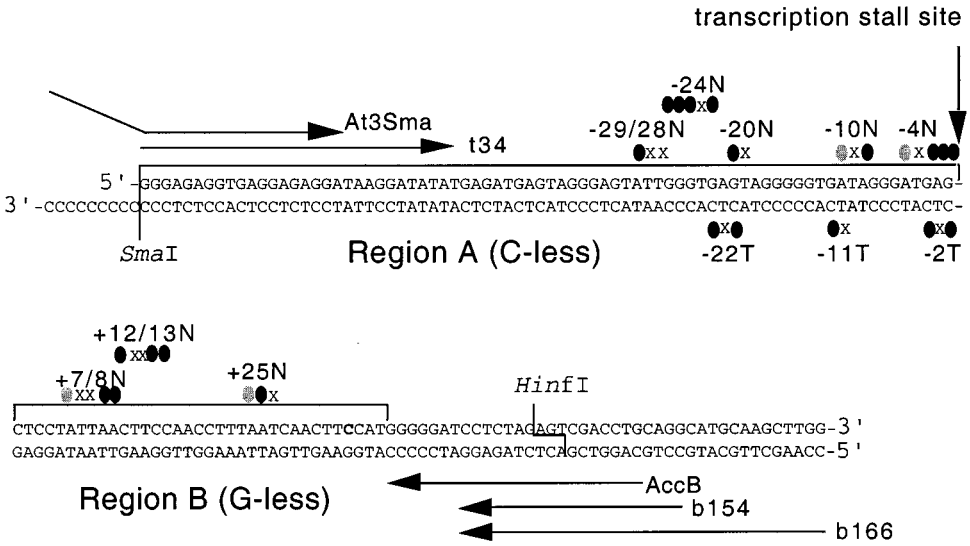

FIGURE 1: Sequence of cross-linking DNA templates and sites of cross-linker incorporation. Tailed templates are generated with upstream primer At3Sma or t34 and downstream primer AccB, b154, or b166. Primer At3Sma includes sequence upstream of the *Sma*I restriction site where the DNA fragment is cut prior to addition of a poly(dC) extension. Yeast RNAP II initiates transcription one to eight bases upstream of the junction of the single-stranded dC tail with the double-stranded DNA using the dC tail as the template. In the presence of ATP, GTP, and UTP, RNAP transcribes 76 bases of the double-stranded DNA prior to reaching the transcription stall site. Region A encodes a C-less region (i.e., the transcript contains no C residues), and region B encodes a G-less region. Sites of cross-linker incorporation are relative to the stall site, with negative numbers indicating the number of bases upstream from the stall site and positive numbers indicating downstream positions. Nontemplate strand cross-linker positions are shown above the DNA sequence, and template strand positions are shown below. X indicates the site of cross-linker incorporation, and the ovals indicate sites of radiolabeled nucleotides next to cross-linkers. The gray ovals indicate positions in the cross-linking DNA constructs that were labeled only on some of the templates, when the asymmetric PCR-generated long primers did not have a non-template-directed 3′-dA added by the *Taq* DNA polymerase.

CACCCAATACT-3′) for −10N, b80 (5′-CCCTATCAC-CCCCTACTCACCC-3′) for −4N, b90 (5′-AGGAGCT-CATCCCTATCACCCCCTA-3′) for +7/+8N, b94 (5′-TAATAGGAGCTCATCCCTATCA-3′) for +12/+13N, and b107 (5′-AAAGGTTGGAAGTTAATAGGAGCTCAT-3′) for +25N. t55 (5′-GTACCCCCCGGGAGAGGTGAG-GAGAGGATAAGGATATATGAGATGAGTAGGGAGT-3′) and for the pretailed template t(−29/−28N) (5′-GGGAGAGGTGAGGAGAGGATAAGGATATATGAG-ATGAGTAGGGAGT-3′) positioned cross-linkers at −29/−28N. t(−24N) (5′-GGGAGAGGTGAGGAGAGGATAAG-GATATATGAGATGAGTAGGGAGTATT-3′) positioned the cross-linker in a pretailed template. Upstream oligos for defining the 3′-ends of template strand primers to be generated by aPCR are t131 for −22 template (−22T) and t121 for −11T. b59 (5′-CAGGTCGACTCTAGAGGATC-CCCCATGGAAGTTGATTAAAGGTTGGAAGTTAAT-3′) was used to place the cross-linker at −2T. Oligos for defining the upstream end of the cross-linking constructs are At3Sma (5′-GTACCCCCCGGGAGAGGTGAGGAGAGGA-3′) and t34 for pretailed constructs (5′-GGGAGAGGTGAG-GAGAGGATAAGGATATATGAGA-3′). Oligos for defining the downstream end of the cross-linking constructs are AccB (5′-CAGGTCGACTCTAGAGGATCCCCC-3′), b154 (5′-GCCTGCAGGTCGACTCTAGAGG-3′), and b166 (5′-CCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGG-3′).

DNA Constructs. The M13 phage construct M13TT18 was described by Wooddell and Burgess (*86*). This phage DNA has a C-less and then a G-less region cloned into polylinker sites. It was used as the template for aPCR to separately generate long primers and complementary DNA strands as described.

Generation of Photoaffinity Tailed Templates. Photoactivatable DNA templates to be transcribed and then cross-linked to the RNA polymerase were made by adding a 3′-poly(dC) tail to the upstream end of a double-stranded DNA fragment containing UV-activatable DNA analogues located at specific sites (Figure 1). A primer and a full-length DNA strand of complementary sequence were generated separately by aPCR using M13TT18 as the template. For aPCR, 50−100-fold more of one oligo than of the other was used (*86*). The primer was then annealed to the full-length DNA strand by combining 4−5 pmol of the long strand, 0.5 pmol more primer than long strand, and 2 μL of 10× buffer A [0.4 M Tris-HCl (pH 7.4), 0.1 M $MgCl_2$, and 0.5 M NaCl] in a volume of ≤20 μL, heating at 90 °C for 10 min, and then setting the mixture in a 70 °C heat block and letting the temperature come down to 37 °C over the course of 1 h. The volume was reduced if necessary so that the following cross-linker incorporation reaction was carried out in 20 μL. DNA analogues, N_3RdUTP or 5N_3dUTP, were incorporated into specific sites in the DNA, next to radiolabeled dNTPs, by extending the annealed primer for 3 min with 20 μM analogue, 2 μM [α-^{32}P]dNTP (3000 Ci/mmol), and 2 units of Sequenase enzyme per picomole of the long strand, in 1 mM 2-ME and 0.1 mg/mL BSA at 37 °C. Both cross-linking analogues were incorporated as efficiently as normal dTTP by Sequenase, a modified T7 DNA polymerase (*83*). An extension reaction was used to form double-stranded DNA. Immediately after the incorporation reaction, an 80 μL solution (pre-equilibrated to 37 °C) was added to result in dATP, dCTP, dGTP, and dTTP each being at a concentration of 0.5 mM, 1× buffer A, 1 mM 2-ME, 0.1 mg/mL BSA, and an additional 1 unit of Sequenase/pmol of DNA. After a 10 min incubation at 37 °C, the enzyme was removed by

applying the reaction mixture to an Amicon Micropure EZ filter set in a Microcon-30 filtration device (Amicon, Beverly, MA) and centrifuged. The retained material was rinsed with H_2O and then with 1× *Sma*I buffer to change buffers. Retentate was recovered, and *Sma*I digestion removed the upstream end from the double-stranded DNA. Then the reaction mixture was heated at 68 °C for 15 min to inactivate the enzyme and applied to a Microcon-30 concentrator to wash out the small *Sma*I digestion fragment and to exchange buffers. After the fluid had been spun out from the *Sma*I digestion, the DNA was washed with 50 μL of 100 mM KOAc, and then three times with 100 μL of H_2O. To improve the recovery of the DNA, 20 μL of H_2O was added to the retentate prior to inverting the filter.

Conditions for generating tailed templates were as described by Dedrick and Chamberlin (*87*). The 400 μL reaction mixtures contained 200 mM potassium cacodylate (pH 7.6), 0.1 mM DTT, 1 mM $CoCl_2$, 100 μM dCTP, 0.1 mg/mL BSA, and 40 units of TdT/pmol of DNA. Samples were incubated at 25 °C for 20 min, followed by addition of EDTA to a final concentration of 15 mM, and heated at 68 °C for 15 min. The samples were then applied to Microcon-30 concentrators that had been treated with 5% Tween 20 (according to the manufacturer's recommendation) to reduce the level of nonspecific binding of single-stranded DNA to the filters, centrifuged, and rinsed with H_2O and then with 1× *Hin*fI restriction buffer (Gibco-BRL). The retentate was digested with *Hin*fI. Samples were then phenol extracted, ethanol precipitated, and resuspended in 0.01% SDS in TE [10 mM Tris-HCl (pH 7.9) and 0.1 mM EDTA]. To gel purify the DNA templates, the resuspended material was run on a 6% polyacrylamide gel, the gel was exposed briefly to film, and gel slices of the appropriate length were cut out. The DNA was electroeluted from the gel slices over the course of 2−3 h at 125 V. Concentrations of recovered DNA were determined by specific activity, depending on how many labeled dNTP sites were next to the cross-linker in each construct. A typical specific activity was 1.5×10^4 dpm/fmol of DNA. Some of the nontemplate strand cross-linking DNA templates were prepared by annealing nontemplate strand primers to a pretailed template strand (*83*). Details are available on request.

Formation of Stalled Transcription Elongation Complexes. Transcription complexes were formed in a final volume of 40 μL after addition of NTPs. In 32.8 μL, 4 pmol of yeast RNAP II (2.4 μg), 50−200 fmol of cross-linker-containing DNA tailed template, and 4 μL of 10× transcription buffer [0.7 M Tris-acetate (pH 8.0), 0.6 M NH_4Cl, 10% glycerol, 60 mM $Mg(OAc)_2$, 50 mM spermidine, and 10 mM 2-ME] were preincubated in siliconized 1.5 mL polypropylene tubes for 10 min at 23 °C. Transcription was initiated by adding 3.2 μL of a solution containing 5 mM ATP and 5 mM GTP and incubating at 30 °C for 10 min. The initiated complexes were extended with 20 μM UTP in the presence of 0.2 mg/mL heparin to allow no more than one RNAP per template DNA; thus, 2 μL of 400 μM UTP and 2 μL of 4 mg/mL heparin were added for a final volume of 40 μL. For some reactions, 1 unit of RNase H was added along with the UTP and heparin to remove RNA that was hybridized to the template strand of the DNA. No-NTP control reactions had H_2O substituted for all the NTP additions.

UV Cross-Linking, Nuclease Digestions, and Protein Electrophoresis. To dissociate nontranscribing RNAP molecules from the DNA, 6.4 μL of 4 M NaCl (0.55 M) was added to the 40 μL transcription reaction mixtures, and they were incubated for 10 min at 23 °C. Samples were irradiated in siliconized 1.5 mL polypropylene tubes for 10 s, 9 cm below a 1 kW mercury vapor lamp, at 23 °C. The high concentration of salt in the UV-irradiated samples was diluted with 350 μL of nuclease mix to digest the template DNA. Nuclease mix consisted of 6 mM $MgCl_2$, 10 mM $CaCl_2$, 20 mM Tris-HCl (pH 7.9), 0.5 mM PMSF, 1 mg/mL leupeptin, 1 mg/mL pepstatin, 1 mM 2-ME, 30 units of DNase I, and 25 units of micrococcal nuclease (Pharmacia). Later experiments contained 1× Complete Mini, EDTA-free, protease inhibitor mix (Boehringer Mannheim) which replaced the PMSF, leupeptin, and pepstatin. Some micrococcal nuclease lots had protease activity; therefore, the protease inhibitors were necessary. Samples were incubated with the nuclease mix for 30 min at 23 °C, then for 15 min at 45 °C, and then for 20 min at 60 °C. Some of the NTP-containing reaction mixtures were treated with an NTP chase of 0.1 mM ATP, UTP, CTP, and GTP during the nuclease digestion to move RNAP molecules that were not cross-linked off the labeled part of the DNA, thus reducing the background level of radioactivity due to undigested DNA. Nuclease digestion reaction mixtures were cooled to room temperature and then precipitated overnight at −20 °C with 6 volumes of cold acetone. Acetone was drawn off the pellets, and promptly thereafter, 2× SDS gel loading buffer [100 mM Tris-HCl (pH 6.8), 200 mM DTT, 4% SDS, 0.2% bromphenol blue, and 20% glycerol (*88*)] was added to resuspend the proteins by vortexing for 45 s, heating at 90 °C for 2 min, letting them sit on ice for a few minutes while processing the remaining samples, vortexing again for 20 s, and heating at 70 °C for 5 min. Samples were then kept on ice until they were loaded onto either 15% SDS−PAGE (*89*) or 10 to 20% Tris-Tricine Novex gels. Gels were silver stained [procedure modified from that of Oakley et al. (*90*)], then soaked in a gel drying solution of 5% glycerol and 30% ethanol, and air-dried between cellophane sheets prior to being put on film. Labels with phosphorescent ink marks were glued to the dried gel and used to align the autoradiogram with the silver-stained gel.

Native Gel Shifts. Template usage in the transcription reactions and the efficiency of dissociation of nontranscribing RNAPs were evaluated by running 1 μL of each reaction mixture after UV irradiation, or DNA alone, mixed with 2.5 μL of 21 mM EDTA and 0.55 M NaCl onto nondenaturing 4 to 15% polyacrylamide gels (Pharmacia PHAST Gel System) for 165 volt hours (400 V and 15 mA) at 23 °C.

Transcription Assays. To evaluate the ability of the RNAP to transcribe each template, transcription reactions were performed in half the volume described above (thus 20 μL), with 1 μL of [α-^{32}P]UTP (10 μCi) added with the nonradioactive UTP and heparin. Transcription was stopped by addition of EDTA to a final concentration of 8 mM and heating at 70 °C for 15 min. The DNA was digested by bringing the volume to 250 μL with 200 μL of H_2O, 25 μL of 10× DNase I buffer [0.1 M Tris-HCl (pH 7.4), 80 mM $MgCl_2$, and 50 mM $CaCl_2$], and 20 units of DNase I and incubating at 37 °C for 20 min. The mixture was then phenol extracted in 1% SDS with 30 μg of yeast tRNA as carrier,

ethanol precipitated, and resuspended in formamide dye to run on a urea denaturing gel.

Analysis of Cross-Linking. Autoradiograms were aligned with the silver-stained protein gels to determine which RNAP subunits were labeled. The PhosphorImager was used to quantify differences in cross-linking reaction mixtures that either did or did not contain NTPs. Linear graphs of the SDS−PAGE lanes of no-NTP cross-linking reactions and their +NTP counterparts were superimposed to determine the level of transcription-specific cross-linking. Within each lane, the labeled polypeptide bands were boxed and quantified to determine the transcription-dependent distribution of labeling among subunits.

Computer Modeling. The yeast RNA polymerase II backbone model containing all of the RNAP II subunits except RPB4 and RPB7 was provided by Cramer et al. (*49*). We modeled an RNAP II transcription elongation complex by combining and positioning the RNAP II model and nucleic acid fragments in MIDAS (*91*, *92*). Final images were prepared with GRASP (*135*). Nucleic acid fragments were built to scale using the actual sequence of the tailed template and RNA in SYBYL version 6.2 with the TRIPOS, Inc., module BIOPOLYMER. Double-stranded DNA fragments that were used were 13 bp upstream (−17 to −29) and 21 bp downstream (+3 to +23). Single-stranded DNA fragments were an 18 nt nontemplate strand in the transcription bubble put together from 8 nt (−6 to +2) and 10 nt (−7 to −16) strands to allow a bend between, an 8 nt template strand upstream of the RNA−DNA hybrid, and a 2 nt template strand downstream of the hybrid (+1 and +2). Single-stranded RNA was 15 bases long. Coordinates for the 8 bp RNA−DNA hybrid were from a structure in the database (not our template sequence). The position of the groove from which the downstream DNA enters the RNAP was determined by Poglitsch et al. (*48*). The location of the active site was determined by Cramer et al. (*49*). The remaining DNA and RNA strands were positioned within the contours of the protein envelope by the best fit that would allow for the helical nature of the nucleic acids and would account for cross-linking data.

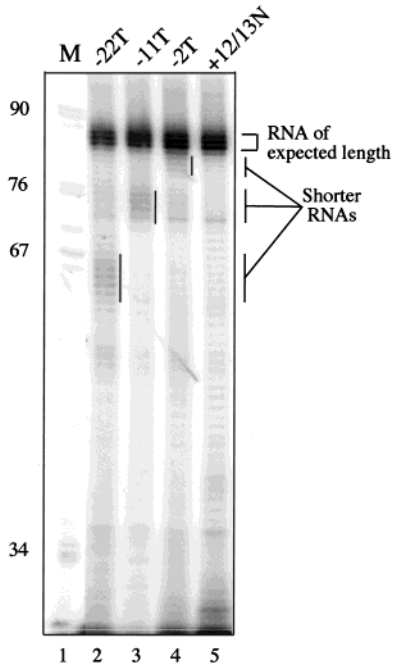

FIGURE 2: Yeast RNAP II transcribes through cross-linker N_3-RdUTP in the template strand. Cross-linker N_3RdUTP was positioned in the template strand of tailed template DNA at positions −22T (lane 2), −11T (lane 3), or −2T (lane 4), and in nontemplate strand position +12/13N (lane 5). RNA from transcription of tailed templates typically appears as a cluster of three to five bands which in this case are primarily 79−81 bases long, indicated as "RNA of expected length". Template +12/13N has the cross-linkers incorporated downstream of the stall site; thus, the RNAP did not transcribe through the cross-linker sites to generate the transcripts shown in lane 5. Transcription through cross-linker sites in the template strand resulted in RNAs that were about 22 bases shorter than the expected cluster when the cross-linker was at −22T (lane 2), 11 bases shorter when the cross-linker was at −11T (lane 3), and just a few bases shorter when the cross-linker was at −2T (lane 4). The percentage of RNAs that were these shorter lengths (indicated by vertical lines beside each lane) was determined by PhosphorImager analysis. In this experiment, the cross-linker templates were radiolabeled next to the cross-linker in the DNA and the RNA was also labeled. The template DNA was digested with DNase I following transcription, resulting in some radioactive background throughout the reaction lanes. M denotes the lane (lane 1) with molecular mass markers.

RESULTS

Experimental Approach. Cross-linking residues were incorporated in the DNA templates at positions from 29/28 bases upstream of the stall site to 25 bases downstream (Figure 1). Once the cross-linking group is incorporated into the DNA, it is no longer a triphosphate; therefore, we will refer to the incorporated long cross-linker as N_3RdU. Each cross-linking template had only one cross-linker incorporated, except −29/28N, +7/8N, and +12/13N which each had two neighboring cross-linker sites in the same DNA. Transcription was initiated on the cross-linker-containing tailed templates by providing ATP and GTP. UTP and heparin were then added, and transcription could proceed 76 bases up to the stall site. Heparin prevents further initiation, such that only the first RNAP, which had already initiated, would occupy each template. Evaluation of the RNA produced in these transcription reactions shows that the ternary complexes (containing RNAP, DNA, and RNA) transcribe up to and then stall at the expected stall site, thus orienting them uniformly at this site (*83*). Stall-length RNAs appear as a cluster of bands due to differences in start site selection (shown as "RNA of expected length" in Figure 2). To determine the relative locations of RNAP II subunits along the DNA in an elongation complex, both upstream and downstream of the stall site, yeast RNAP II must be able to transcribe through cross-linker-containing sites in the DNA templates to reach the stall site. RNAP II transcribes through cross-linker sites in the nontemplate strand of the DNA as well as it does on normal DNA (*83*). Yeast RNAP II was able to transcribe through approximately 90% of the sites in the DNA that had either N_3RdUTP or $5N_3$dUTP incorporated in the template strand (Figure 2). Transcriptions of tailed templates with cross-linker N_3RdU in the template strand, 22, 11, or 2 bases upstream of the stall site (lanes 2−4, respectively), or 12 and 13 bases downstream of the stall site (lane 5) are shown in Figure 2. Essentially all the templates had cross-linker incorporated (data not shown). In the absence of the cross-linker, the RNA pattern appears to be identical to the pattern of transcription of a template with the cross-linker downstream of the position where the

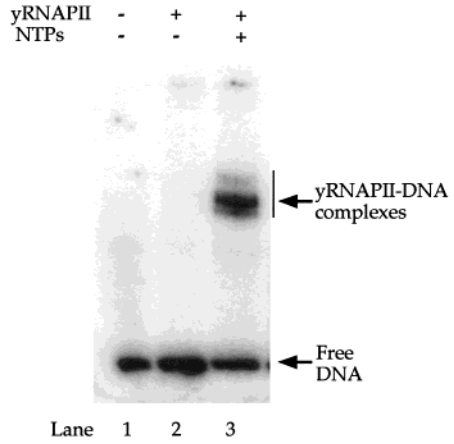

FIGURE 3: Native gel shifts were used to evaluate formation of NTP-dependent complexes. For each cross-linking reaction, transcription complexes were formed on cross-linkable tailed templates, incubated in 0.55 M NaCl for 10 min, and UV irradiated, and then a 1 μL sample from each reaction mixture was run on a 4 to 15% native PHAST gel as described. Lane 1 had DNA template alone with cross-linker N_3RdU at −22T. The sample in lane 2 was from a mock transcription reaction mixture without NTPs and shows that the concentrations of NaCl and heparin were adequate to dissociate nontranscribing RNAP from the DNA. Lane 3 shows the gel-shifted ternary complexes from a transcription reaction, including ATP, GTP, UTP, and 1 unit of RNase H. The mobility of the free DNA is indicated.

RNAP is stalled (data not shown). Templates with the cross-linker in the template strand, in a region that had to be transcribed to reach the stall site, show a pattern of RNAs that are shorter than the standard stall-length cluster (lanes 2−4), and the differences in length (determined by counting bands on an overexposed autoradiogram) correspond to transcripts stopped at the site of cross-linker incorporation. Only about 10% of the transcripts ended at cross-linker incorporation sites.

Native gel shift experiments were used to evaluate the formation of NTP-dependent complexes, confirming that the RNAP was active and indicating what proportion of a preparation of tailed DNA templates was used by the RNAP in each reaction (Figure 3). The native gel shifts that were carried out for every cross-linking reaction required only 1 μL of sample to run on small PHAST gels (Pharmacia) and ran in about 30 min. Templates with very short dC extensions did not support transcription (data not shown). After transcription elongation complexes were formed, they were challenged for 10 min in 0.5−0.6 M NaCl to dissociate nontranscribing RNAPs prior to UV irradiation. Nontranscribing RNAP II binds DNA. In cross-linking reactions with nontranscribing RNAP and cross-linker-containing DNA templates, many of the RNAP II subunits were labeled when the salt concentration was not high enough to disrupt these associations (data not shown). Elongating RNAP complexes are stable at higher salt concentrations. The salt concentration that was used was determined for each RNAP preparation by assessing the NTP dependence of RNAP binding to the DNA. At the appropriate salt concentration, the native gel shift pattern of free DNA was identical to that of reactions without NTPs (compare lanes 1 and 2 in Figure 3). After cross-linking, the samples were diluted to allow DNase I and micrococcal nuclease to digest the DNA, leaving only small radioactive tags on cross-linked subunits. The proteins were acetone precipitated and resuspended in a smaller volume to run on the SDS−PAGE gel. Acetone precipitations gave consistent protein recoveries, allowing for quantitative comparisons of the amounts of each subunit labeled. Following electrophoresis, the gels were silver stained to compare sample recoveries and to identify the location of each subunit in the gels that contained labeled subunits.

Use of RNase H To Reanneal the DNA Template Strands following Transcription. Transcription of tailed templates has been well characterized (*58*, *61*, *64*, *87*, *93*, *94*). The 3′-dC tail is used as the template, and the transcript initiates one to eight bases immediately upstream of the junction of the dC extension with the double-stranded DNA (*61*, *87*). An anomalous feature of transcription from tailed templates which has also been observed with linear and supercoiled DNA templates is that the RNA sometimes becomes extensively hybridized to the template DNA strand (*58*, *61*, *87*, *93*, *95*−*97*). When RNase H is added to transcription reaction mixtures from tailed templates, the RNA is shortened from its 5′-end and the template and nontemplate strands of the DNA reanneal (*61*, *87*, *94*). In transcriptions from tailed templates, 10−60% of the ternary complexes have DNA strands that are reannealed, as detected by the sensitivity of the RNA to RNase A or RNase T1, and the remainder of the ternary complexes have DNA template strands hybridized to RNA, detected by the sensitivity of the RNA to RNase H. The proportion of RNAs detected as being hybridized to DNA or displaced from the DNA depends in part on the sequence of the DNA at the 5′-end, the species from which the RNAP II is purified (*61*, *87*, *93*), and the procedure used to stop the transcription (*61*). When RNase H was present during the transcription, Sluder et al. (*61*) observed that the RNAs were shortened by 5−15 nucleotides from the 5′-end, but the overall amount of transcript was barely affected. These observations led to the conclusion that the presence of RNase H during transcription from tailed templates allows the transcription bubble to reform and the RNA product to be displaced from the template strand by cleaving the RNA that hybridizes to the DNA, thus allowing the two DNA strands to reanneal (*61*, *93*, *94*). We chose to use this property of RNase H to compare cross-linking results from transcription complexes that have the DNA strands reannealed, a more normal situation, with cross-linking results from complexes that have some separated DNA strands. DNase I footprinting was carried out to verify that the DNA strands had reannealed and that the RNAP was appropriately stalled at the expected site (*83*). In the absence of RNase H, NTP-dependent complexes migrated at two different native gel mobilities. The DNA was recovered from DNase I-treated transcription complexes stable to 0.5 M NaCl. The complexes with reannealed DNA strands showed a DNase I footprint from nucleotides −31 to +21 with partial protection to nucleotide +25. Complexes with separated DNA strands showed protection downstream to nucleotide +23, but no upstream boundary could be detected because the DNase I-cleaved nontemplate strand was presumably dissociated from the transcription complex (DNA was labeled on the 5′-end of the nontemplate strand). RNase H-treated complexes showed DNase I protection from nucleotides −26 to +19 and partial protection to nucleotide +24, an indication of DNA strands that had reannealed upstream. Template DNA with the cross-linker at nucleotide +25N did not label any RNAP subunits, but cross-linkers at the upstream boundary (−28/−29N) did

label RPB2 and to a lesser extent RPB1, suggesting that the upstream DNA may wrap around the RNAP while downstream DNA protrudes away from the RNAP. A small number of RNAP molecules positioned elsewhere, for example, having slid backward on the DNA, would probably not be detectable.

Elongation Competence of Transcription Complexes. Sluder et al. (*61*) noted that about half of the ternary complexes formed on tailed templates with *Drosophila melanogaster* RNAP II were not able to transcribe beyond the first 14−17 bases under the most favorable buffer conditions that were used. Transcription of our tailed templates also revealed many short transcripts (*83*). Since some of the population of RNAPs unable to transcribe beyond the first 14−17 bases remains bound at the upstream end of the template, we placed cross-linkers farther from the upstream end of the template, within the region of the DNA that was footprinted by RNAP after having transcribed 76 bases to reach the stall site. Only the RNAPs that transcribed up to the stall site were within reacting distance of the cross-linkers. On the basis of the careful study of exonuclease III footprints of human RNAP II ternary complexes stalled after transcribing between 20 and 130 bases, Samkurashvili and Luse (*98*) concluded that the elongation complex may not reach its fully mature form until it has transcribed at least 40 bases. In our experiments, elongation complexes that stalled after transcribing 76 bases would thus be expected to be of the mature form. Although the ternary complexes in the indicated study were formed in a HeLa cell nuclear extract and had initiated at a promoter, previous studies have shown similar DNase I footprints of ternary complexes initiated from a promoter (*99*) and on tailed templates (*100*). The footprints of yeast RNAP II we obtained in our system (data not shown) are comparable with those of other productive elongation complexes (*98*−*101*) and show no evidence of having slid back (*102*, *103*). Komissarova and Kashlev (*103*) demonstrated that at some DNA sequences the RNAP translocated backward 6−10 bases after being stalled by omission of NTPs. The 3′-end of the RNA had disengaged from the active site in such complexes, causing them to become arrested and unable to resume elongation unless the 3′-end of the RNA was cleaved.

Our ternary complexes with or without RNase H treatment were able to elongate further upon addition of all four NTPs, although approximately half of the ternary complexes that had reached the stall site (76 bases from the start) were not stable to the 0.6 M NaCl incubation and released their stall-length transcripts (*83*). Since the RNA−DNA hybrid is believed to account for the stability of the RNAP on the DNA (*104*), RNAPs in ternary complexes that release the RNA are then most likely dissociated from the DNA. Arrested complexes retain the RNA, although they are unable to resume elongation. Most of our complexes either resumed elongation or dissociated. Sluder et al. (*61*) noted the instability of ternary complexes that had been formed on the tailed template when transcription reactions were stopped with 100 mM NaCl and 10 mM EDTA. We added no EDTA, but used 600 mM NaCl. The 0.6 M NaCl incubation was used to distinguish NTP-dependent cross-linking from nonspecific RNAP−DNA binding, despite its destabilizing effect on ternary complexes.

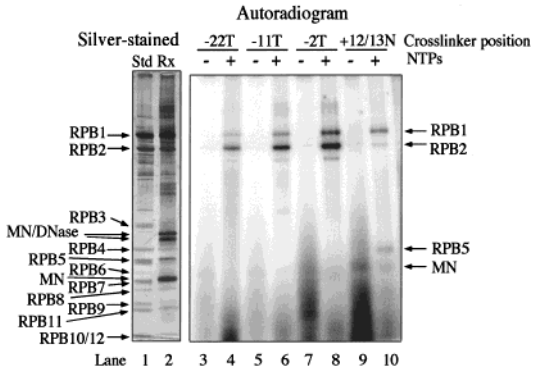

FIGURE 4: Representative photo-cross-linking experiment showing that long cross-linker labels RPB1, RPB2, and RPB5 in a transcription elongation complex. The panel on the left shows a silver-stained 10 to 20% Tris-Tricine SDS/PAGE with yeast RNAP II standards in lane 1 and a cross-linking reaction mixture in lane 2. The cross-linking reactions included DNase I and micrococcal nuclease (MN) that were added after UV irradiation, and then acetone precipitated along with the RNAP subunits. The panel on the right is an autoradiogram of a larger portion of the same gel shown on the left. Cross-linker N_3RdU was incorporated into position −22T (lanes 3 and 4), −11T (lanes 5 and 6), −2T (lanes 7 and 8), or +12/13N (lanes 9 and 10). Transcription reaction mixtures including NTPs are shown in lanes 4, 6, 8, and 10. Control reaction mixtures without NTPs are shown in lanes 3, 5, 7, and 9. The labeled subunits RPB1, RPB2, and RPB5 are indicated. Polypeptides originating from the micrococcal nuclease preparation are labeled MN. This nuclease becomes associated with the label (as described in the text) and can be seen as a band migrating below RPB5 in most lanes of the autoradiogram.

Long Cross-Linker N_3RdU Labeled Primarily RPB1, RPB2, and RPB5. (1) Experimental Design. Figure 4 shows a representative silver-stained SDS−PAGE gel in the left panel. In addition to the yeast RNAP II subunits (shown in the standards, lane 1), cross-link reaction samples run on the gel include the nucleases used to digest the DNA (lane 2). The efficiency of the nuclease digestions varied somewhat, accounting for different degrees of background radioactivity in some lanes. For each cross-linker position, samples from a mock transcription reaction (without any NTPs added, shown as −NTP lanes) were run on the gel beside the samples from reactions that did include NTPs (+NTP lanes). To determine the cross-linking that was dependent on transcription and to quantify the amount of label transferred to RNAP subunits by cross-linking, PhosphorImager analysis was employed. Relative amounts of labeling on each subunit, compared to the background from −NTP lanes, were visualized by overlaying line graphs of the gel lanes (Figure 6). Label sometimes becomes associated with the micrococcal nuclease used to digest the labeled DNA in the cross-linking reactions. This association is independent of cross-linker being in the DNA and independent of UV irradiation (data not shown). The labeling on the nuclease does, however, obstruct detection of labeling on RPB7 in the 10 to 20% Tris-Tricine gels used in many of the experiments.

(2) Measuring the Frequency of DNA Cross-Linking to RPB1, RPB2, and RPB5 at Each Position along the Footprinted Part of the DNA. For each cross-linking DNA template with the photoactivatable analogue in a specific position, the relative amount of cross-linking to each subunit was determined by comparing the number of counts in each band on a cross-linking gel. Each position was evaluated in at least three experiments and reproducibly gave a similar

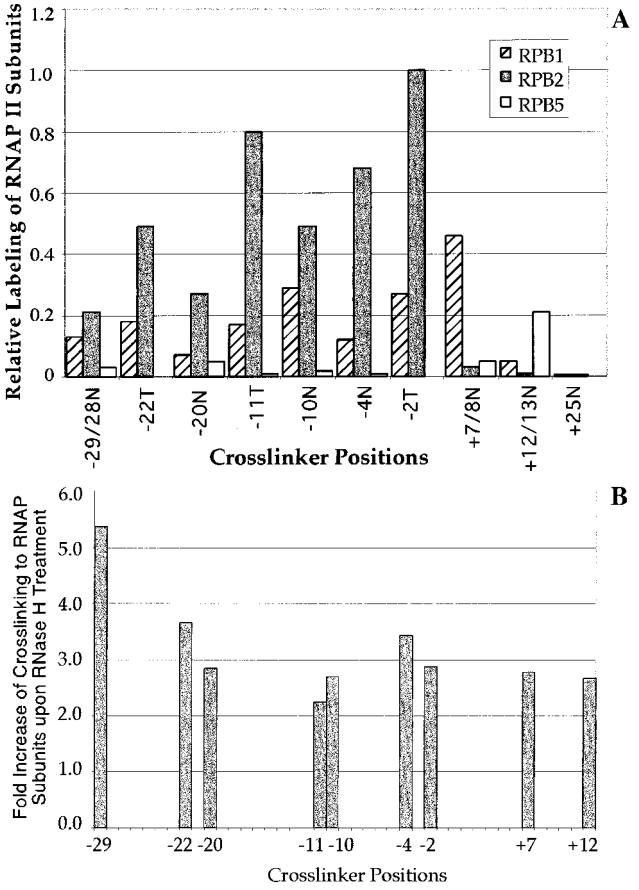

FIGURE 5: Distribution of cross-linking to yeast RNAP II subunits from the long cross-linker in the indicated positions along the DNA template. (A) Transcription elongation complexes were formed in 40 μL reaction mixtures with 2 pmol of RNAP II and varying amounts of radiolabeled cross-linker DNA templates from 66 fmol (for −4N) to 303 fmol (for both −20N and −11T) such that each cross-linking reaction mixture included the same amount of radioactivity. Along with the UTP and heparin, 1 unit of RNase H was added to each transcription reaction mixture. Cross-linking reaction mixtures were processed as described, and the radioactivity associated with each RNAP subunit was quantified. Relative amounts of cross-linking to subunits RPB1, RPB2, and RPB5 are shown for each cross-linker position in the DNA (shown on the *x*-axis). The relative amounts of labeling on each subunit are shown on the *y*-axis. The subunit that cross-linked the most frequently (RPB2 at −2T) was given a value of 1, and the extent of cross-linking to all other subunits is relative to this value. (B) Transcription−cross-linking reactions with or without RNase H were carried out side by side. The increase in the level of cross-linking to RNAP in the presence of RNase H (on the *y*-axis) is shown for each DNA position (on the *x*-axis). The increase value was calculated by quantifying cross-linking to the subunit that was most often cross-linked at each DNA position and then dividing the amount of cross-linking on that subunit in reaction mixtures that included RNase H by the amount of cross-linking to that subunit in the absence of RNase H.

distribution of cross-linking to RPB1, RPB2, and at many positions also RPB5 (see Figure 5). As shown in Figure 4 and the composite graph in Figure 5, RPB1 and RPB2 were cross-linked from every position upstream of the ribonucleotide addition site that was evaluated. In every case, RPB2 was labeled more than RPB1. At position −10N, the amount of cross-linking on RPB1 was more similar to that on RPB2 (Figure 5A). From −24N, the combined labeling on RPB1 and RPB2 was equal to the combined labeling from −29/28N, but from −24N, RPB1 and RPB2 were tagged with equal frequency (data not shown). Downstream of the active site at +7/8N and +12/13N, however, RPB1 was cross-linked most often and only a very small amount of RPB2 was cross-linked. Interestingly, RPB5 was cross-linked both downstream of the active site (positions +7/8N and +12/13N) and upstream of the active site (positions −10N and −20N). In the vicinity of the active site (at both −4N and −2T), however, there was no significant labeling of RPB5.

We examined the resulting effects on the distribution of cross-linking at each DNA position with or without RNase H to reanneal the DNA strands. The distribution of cross-links to RPB1, RPB2, and RPB5 at each individual position was similar in the absence (*83*) or presence (Figure 5A) of RNase H. One notable difference was that with +12/13N, RPB5 was cross-linked much more than RPB1 in the presence of RNase H, whereas the amount of labeling of RPB1 and RPB5 was more similar in the absence of RNase H.

(3) Use of Photo-Cross-Linking To Quantify the Proximity of the RNAP to the DNA at Specific Positions. The proximity of the RNAP to the template DNA was evaluated by comparison of the amount of cross-linking to each RNAP subunit from each DNA position (shown in Figure 5 A). RNAP subunits cross-link to the DNA much more frequently from the upstream DNA positions of nucleotide −11 down to nucleotide −2, and less frequently upstream or downstream of this middle region. Upstream positions −22 and −20, expected to be just upstream of the typical transcription bubble, cross-link at a similar frequency to the downstream position +7/8N. Essentially no cross-linking to position +25N was detected, consistent with this position being just downstream of the edge of the DNase I footprint of our transcription complexes (*83*).

To compare the amount of cross-linking from the different DNA positions, we used two different approaches. The first approach suffered from the imprecise understanding of the frequency with which *Taq* DNA polymerase adds a nontemplate directed dA on the 3′-end of a PCR product. Most of the primers used to generate the tailed template probes were obtained by aPCR. Approximately 75% of the time, the 3′-dA was added by the *Taq* DNA polymerase (*86*). In some cross-linking distribution experiments, we used equal amounts of cross-linker DNA templates in each transcription−cross-linking reaction, quantified the amount of label transferred to each of the subunits, and divided the radioactive counts by the number of radiolabeled dNTPs in that template to normalize each reaction (see Figure 1). The number of radiolabeled dNTPs incorporated into each DNA depended on the sequence. In probe −29/−28N, for example, the one dA just upstream of the two cross-linkers was the labeled base. The primer to make this probe was from a purchased oligonucleotide; thus, all templates were expected to incorporate the dA from the radiolabeled dATP during generation of the probe. In contrast, probe +7/8N had two labeled dAs downstream of the two cross-linkers and one dA just upstream. The partially shaded oval upstream of the cross-linkers in Figure 1 represents the dA that would have been added by *Taq* DNA polymerase during the process of making the long primer strand by aPCR. Since this DNA polymerase adds a nontemplated dA onto the PCR product, about 75% of the time, we estimated the specific activity of probe +7/8N by considering that it has 2.25 labeled dAs

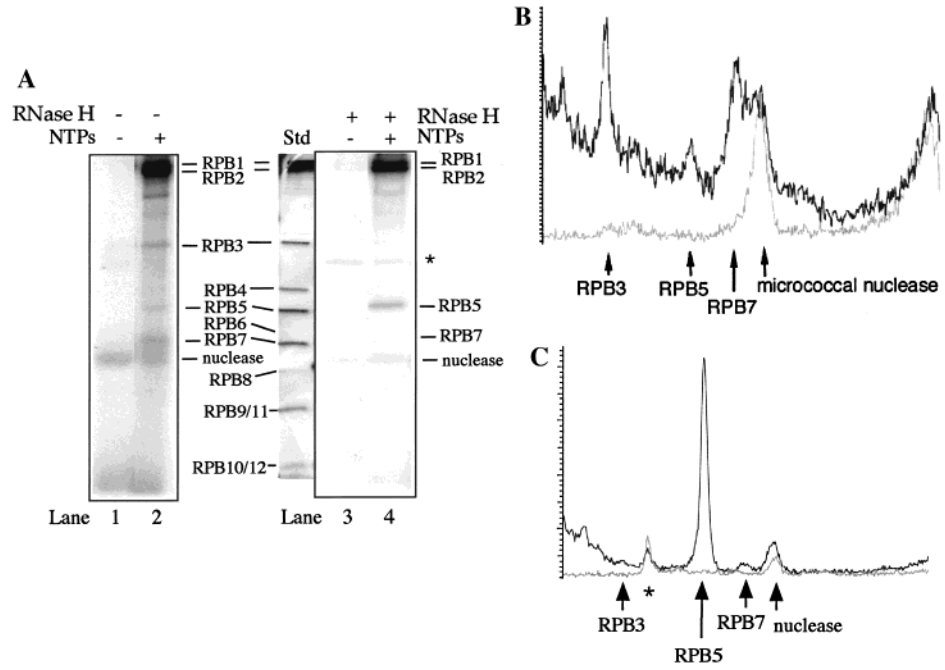

FIGURE 6: At DNA position −29/−28N, the distribution of cross-linking to RNAP II subunits changes in the presence and in the absence of RNase H to reanneal the DNA template strands. (A) Autoradiograms of cross-linking gels showing that the labeling of RNAP II subunits from cross-linker template −29/−28N differs between the transcription reaction mixtures that had no RNase H added (lanes 1 and 2) and reaction mixtures that did have RNase H added (lanes 3 and 4). These cross-linked samples were run on a 15% SDS−PAGE gel and silver stained. Silver-stained yeast RNAP II standards are shown beside the autoradiogram from reaction mixtures that did include the RNase H. The −NTP reactions are shown in lanes 1 and 3; the +NTP reactions are shown in lanes 2 and 4. RPB1 and RPB2 migrate close together on this gel, and both were labeled by cross-linking; the labeling of RPB1 and RPB2 is overexposed compared to the other subunits that became labeled. In the absence of RNase H, RPB1−RPB3, a small amount of RPB5, and RPB7 became cross-linked. When RNase H was included in the transcription, RPB1, RPB2, RPB5, and barely any RPB7 became labeled. The bands on the autoradiogram due to micrococcal nuclease are indicated. A band of artifactual labeling is shown in lanes 3 and 4 (indicated with an asterisk). (B) Superimposed PhosphorImager scans of −NTP and +NTP reaction mixtures in the absence of RNase H (lanes 1 and 2 from part A). RNAP subunits that became labeled in an NTP-dependent manner are indicated. (C) Superimposed scans of −NTP and +NTP reaction mixtures in the presence of RNase H (lanes 3 and 4 from part A) are shown. Those subunits labeled in an NTP-dependent manner are indicated.

per DNA template. Probe −11T had to have had a 3′-dA added by *Taq* to be radiolabeled, because the label was dCTP and only the thymidine analogue cross-linker and labeled dCTP were used in the initial step to incorporate cross-linker into the DNA.

The second approach used to relate the amount of cross-linking from different positions was based on our finding that the amount of RNA produced in transcription reactions with these tailed templates correlated directly with the amount of DNA template in the reaction as long as the DNA concentration was within the linear range from 25 to 100 fmol of template DNA per 20 μL of transcription reaction mixture (*83*). For the set of cross-linking experiments shown in Figure 5, the number of DNA templates used was chosen to give an equal number of radioactive counts per transcription−cross-linking reaction. The amount of label that transferred onto subunits from each cross-linker position could then be directly compared using the assumption that the amount of cross-linking is proportional to the amount of DNA added to the transcription reaction. In this experiment, the DNA templates all had the same specific activity of radiolabeled dNTPs and all of the cross-linking gels were exposed to the PhosphorImager screen at the same time for the same period of exposure. The transcription−cross-linking reactions shown in Figure 5 included RNase H treatment to allow reannealing of the template DNA strand upstream of the transcription bubble. It is important to note that the cross-linking distribution of reactions without RNase H, quantified by either method, had a similar cross-linking distribution (data not shown).

When the distribution graphs of the experiments with or without RNase H were compared, the relative amount of cross-linking from −20N and on downstream to +25N looked similar. In transcriptions that included RNase H, the major difference was that the cross-linker at −29/−28N cross-linked more frequently. There was also somewhat more cross-linking from −22T. RNase H in the transcription reactions resulted in 2−3-fold more cross-linking from each DNA position (Figure 5B), except at the upstream edge of the footprint, position −29/−28N, which cross-linked 5 times as well when RNase H was included. The enhancement of cross-linking was assessed by comparing the amount of label transferred to the subunit that cross-linked the most at each given position. The overall enhancement of cross-linking is probably due to the added stability of ternary complexes with template strands that are reannealed following transcription. As many as half of our ternary complexes (formed in the absence of RNase H) dissociate after being subjected to 0.6 M NaCl for 10 min (*83*). The added enhancement of cross-linking at −29/−28N could be explained by the nontemplate DNA strand coming back together with the template strand, rather than dangling out in space.

Cross-Linking of Subunits RPB3, RPB4, and RPB7. DNA templates with N_3RdU positioned at the upstream edge of the footprinted region on the DNA, at −29/−28N, cross-linked to RPB3 and RPB7 (Figure 6A, lane 2). The cross-

linking to RPB3 seemed reasonable in light of RPB3 being structurally analogous to *E. coli* RNAP subunit α [phylogenetic tree and review by Svetlov et al. (*21*); *43*, *105*−*108*]. The prokaryotic subunit interacts with DNA in the upstream region of some promoters (*109*−*113*), and the finding that RPB3 cross-linked from the upstream edge of the RNAP was consistent with the analogous functions of these eukaryotic and prokaryotic subunits.

Cross-linking of RPB7 to this upstream position was perplexing in light of the localization of RPB4 and RPB7 in what appeared to be the DNA-binding floor of the enzyme (*50*). These data where obtained with two-dimensional crystals. Recently determined higher-resolution structures have revealed that the RPB4−RPB7 subassembly is positioned near the DNA-binding groove, but not in it (R. Kornberg, personal communication). The nontemplate strand of the DNA at position −29 is reannealed with the template strand in a normal transcription complex. The transcription bubble is typically 18 bp with the DNA being opened from nucleotide +2 to −16 or so (*114*). Since tailed template transcription leaves some of the ternary complexes with the template strand of the DNA annealed to the RNA, and not to the nontemplate strand, we used RNase H to re-form the transcription bubble in the transcription reactions shown in Figure 6. When the DNA strands were brought back together, the cross-linker at position −29/−28 no longer reacted with RPB3, the cross-linking to RPB7 was barely detectable, and instead, RPB5 was labeled much more than in the reactions without RNase H (compare lane 2 with lane 4 in Figure 6A). Line graphs showing the relative peak heights of labeling on RPB3, RPB5, and RPB7 are shown in panels B and C of Figure 6. The extent of labeling of RPB5 from −29/−28N in the presence of RNase H was, however, much lower than the extent of labeling of RPB1 and RPB2 (Figure 5A). Other than the labeling of RPB7 from −29/−28N, RPB7 labeling was only detected at a very low level from −4N and −10N (data not shown), but not from any other sites that were evaluated.

Subunits RPB4 and RPB7 form a subassembly that is not stoichiometrically associated with yeast RNAP II that is purified from cells in the log phase of growth (*115*−*117*). This subassembly was present in our yeast RNAP II preparations (see Figures 4 and 6). When yeast RNAP II was incubated with double-stranded DNA or tailed templates and then NaCl was added at a concentration not sufficient to dissociate all of the nonspecific binding, we frequently detected labeling of RPB4 that was not NTP-dependent in a cross-linking reaction with N_3RdU incorporated into the DNA (*83*). Although RPB4 readily cross-linked to DNA that was not in transcription complexes, we never saw it labeled in NTP-dependent ternary complexes.

Use of the Short Cross-Linker 5N_3dU To Detect Which of the RNAP II Subunits Is Nearest to Each Position along the DNA. Since two or three subunits were found to cross-link to the RNAP from each of the DNA positions that were evaluated, we used the short cross-linker with an azido group directly on the uridine ring to determine which of these two or three subunits was nearest the DNA (at each position). At all of the positions upstream of the active site (−29/−28N, −22T, −20N, −11T, −10N, and −4N), RPB2 was predominantly cross-linked, except at −2T (Figure 7A and data not shown). Just two bases upstream of the active site,

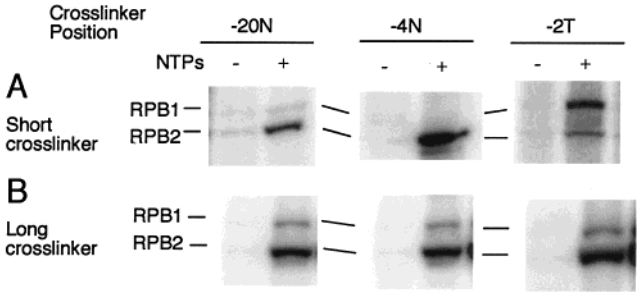

FIGURE 7: Differences between cross-linking with the short cross-linker 5N_3dUTP and with the long cross-linker N_3RdUTP. Cross-linker positions are shown above each column of autoradiogram panels: −20N is on the left, −4N is in the center, and −2T is on the right. Acrylamide gels are 10 to 20% Tris-Tricine. (A) The short cross-linker was incorporated into the indicated positions in the template DNA. For each panel, −NTP reaction mixtures are shown on the left and +NTP reaction mixtures on the right. The positions of RPB1 and RPB2 labeling are shown. From −20N, RPB2 was labeled much more than RPB1. From −4N, RPB2 was labeled almost exclusively. From −2T, RPB1 was labeled more than RPB2. (B) The long cross-linker was incorporated at the same positions shown for the short cross-linker. At each of these positions, RPB2 was labeled much more than RPB1.

the template strand labeled both RPB1 and RPB2, with more of the cross-linking to RPB1. Curiously, the long cross-linker labeled predominantly RPB2 from this position. The active site is known to be composed of both RPB1 and RPB2; thus, the cross-linking of both subunits near the active site by this short cross-linker is not surprising. Bartholomew et al. (*78*) also detected a different cross-linking distribution on yeast RNAP III subunits when using a zero-distance cross-linker (4-thio-dTMP) compared to the long cross-linker N_3RdU. 4-Thio-dTMP cross-linked to the 82 kDa (C82) RNAP III subunit much more efficiently, relative to other RNAP III subunits, than N_3RdU did. N_3RdU reacted strongly with the 34 kDa subunit (C34), but not with the 31 kDa subunit (C31), whereas 4-thio-dTMP reacted with both of these subunits. Differences in cross-linking results with short and long cross-linkers will be discussed below.

With a short cross-linker at −4N, RPB2 was exclusively labeled. At the other upstream positions, RPB2 was predominantly labeled, but some RPB1 was also labeled in each case (see the representative result with −20N in Figure 7A). For comparison, the cross-linking to RPB1 and RPB2 with long cross-linker is shown for each of these templates. Although with the short cross-linker the cross-linking pattern is different for each of these three positions (Figure 7A), the long cross-linker exhibited essentially the same distribution of cross-linking to RPB1 and RPB2 (Figure 7B) with RPB2 labeled more than RPB1.

Downstream of the active site at +7/8N and +12/13N, RPB1 was the major subunit labeled with the long cross-linker (Figure 8A) and the only subunit labeled with the short cross-linker (Figure 8B). With both cross-linkers, the extent of minor labeling of RPB2 was so low that it was barely above the nonspecific extent of binding in the −NTP lane. With the short cross-linker, RPB1 was cross-linked only one-third as often from position +12/13N as from +7/8N. The gels shown in this figure also show that cross-linking is dependent on all three NTPs required to reach the stall site. Reactions with ATP and GTP (lanes 2 and 5 in both panels A and B), allowing the RNAP to initiate but not transcribe

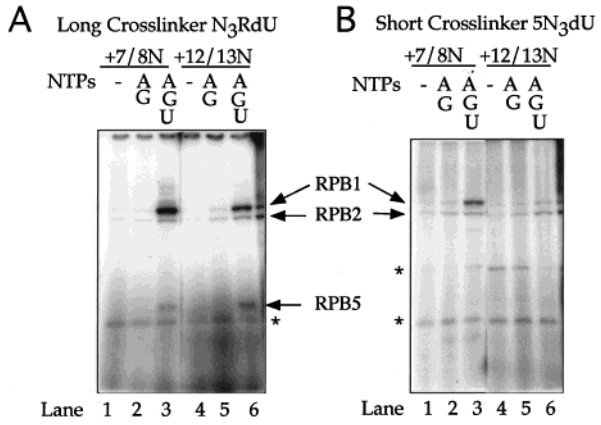

FIGURE 8: Short cross-linker used to determine the proximity of DNA to yeast RNAP II at downstream positions +7/8N and +12/13N. (A) The long cross-linker was incorporated at +7/8N (lanes 1−3) or +12/13N (lanes 4−6). Cross-linking reaction mixtures without NTPs are shown in lanes 1 and 4. Cross-linking reaction mixtures of ternary complexes formed by initiation with ATP and GTP, but without UTP, are shown in lanes 2 and 5. Transcription complexes formed with ATP, GTP, and UTP prior to UV irradiation are shown in lanes 3 and 6. Arrows point to the labeled subunits RPB1, RPB2, and RPB5. (B) The short cross-linker was incorporated at +7/8N (lanes 1−3) or +12/13N (lanes 4−6). Lanes are as described for part A. Artifactual bands in the gel (due to the nuclease labeling) are indicated with asterisks.

beyond the eighth base of the template, do not show any more cross-linking than the −NTP lanes (compare lanes 1 and 2, and 4 and 5).

DISCUSSION

Cross-Linking of Multiple Subunits from a Single DNA Position. Cross-linker N_3RdU has its azido moiety on a 9−10 Å side chain projecting from the 5-position of the pyrimidine ring into the major groove and its surrounding space (*68*). The reactive nitrene formed from photoactivation of an azido group reacts nonspecifically, and it can even be inserted into carbon−hydrogen bonds. Although each azido cross-linker can react only with one other molecule, the cross-linking of multiple subunits from a single DNA position is not surprising. The length of the side chain that connects the azido to the uridine base allows this cross-linker to probe polypeptides within about 10 Å of the DNA and does not require the protein to be closely bound in the major or minor groove. This tether also allows the cross-linker to probe within a space of perhaps as many as 4 bp (*68*). The short cross-linker $5N_3$dU serves to detect closer interactions between the RNA polymerase and the template DNA, since the reactive azido is directly on the uridine ring.

As mentioned above, Bartholomew et al. (*78*) found that a short cross-linker reacted with some subunits not detected by the longer cross-linker in their study of yeast RNAP III transcription initiation complexes and ternary complexes that had transcribed 10 or 17 bases. These authors attribute the differences between cross-linking results with the short versus long cross-linker to the following possibilities: (a) the side chain projecting from C-5 in N_3RdU may prevent RNAP III from being positioned on the DNA as it otherwise would be, (b) the aryl azide cross-linker of N_3RdU may not be as effective at scanning the space close to the DNA as the zero-distance 4-thio-dTMP cross-linker, and (c) RNAPs may be in multiple states even in transcription complexes where they are thought to be precisely positioned. Both of our cross-linkers react rapidly and indiscriminately by means of a nitrene group; cross-linking, therefore, measures proximity, not differences in reactivity. In addition to the reasons suggested by Bartholomew et al. above, the different length cross-linkers may yield different results due to the aromatic structure of the aryl azide of N_3RdU. If this cross-linker preferentially associates with aromatic amino acid residues, even though it cross-links nonspecifically, it may not be positioned randomly among the nearby polypeptides that it could otherwise contact.

To transcribe rapidly, the RNAP must not bind the DNA too tightly. A possible reason multiple subunits are cross-linked from a single position could be that the DNA is flexible outside of a tight active site interaction that serves to clamp the DNA in place and may include movement of the hinged "clamp" domain over the DNA (*46*, *49*). The DNA could move somewhat from side to side in the downstream DNA-binding groove and upstream of the active site. Single-stranded DNA in the transcription bubble is also likely to be flexible. In a crystallized transcribing T7 RNA polymerase initiation complex, Cheetham and Steitz (*118*) found the unhybridized nontemplate strand in the transcription bubble to be disordered in electron density. Another possibility (which does not exclude the others) is that the RNAP was sliding reversibly forward and backward on the DNA (*102*, *104*). Guajardo and Sousa (*41*, *119*) have proposed a model for the mechanism of polymerase translocation that is based on diffusional sliding between the pre- or post-translocated positions in a manner dependent on the relative free energy for each position.

RNAP may also translocate backward along the DNA about 6−10 bp, displacing the RNA from the active site and causing the RNAP to arrest (*103*, *104*, *120*). Some transcription complexes stalled by withholding an NTP become arrested, unable to elongate further when provided with the missing NTP (*104*, *121*−*123*). Nudler et al. (*104*) noticed that the lateral stability of RNAP (tendency to slide back) depended on the strength of the RNA−DNA hybrid in the transcription bubble. The DNA sequence around the stall site in our template does not have runs of Us, but rather is GC-rich. We do not believe sliding back was a significant cause of cross-linking to several different subunits from a given single DNA position for the following reasons. The RNAP in our system was correctly positioned at the stall site with a DNase I footprint comparable to those of other elongation-competent RNAP II complexes (*98*−*100*); many of the ternary complexes could be elongated and were therefore not arrested, and the transcripts that were not elongated had dissociated (*83*).

Orientation of Nucleic Acids in the Transcription Complexes. Cramer et al. (*49*) have determined the X-ray crystal structure of yeast RNAP II to 3 Å resolution and with these data derived a backbone model of the enzyme. From the higher-resolution structural data, these researchers determined that the active site was farther toward the interior of the enzyme than was previously believed (*44*, *48*, *124*). The active site is within the groove initially observed to be too narrow to accommodate double-stranded DNA. Fu et al. (*44*) determined that a hinged domain either could partially enclose this groove and make the groove appear more narrow or could open more widely. Using our cross-linking data,

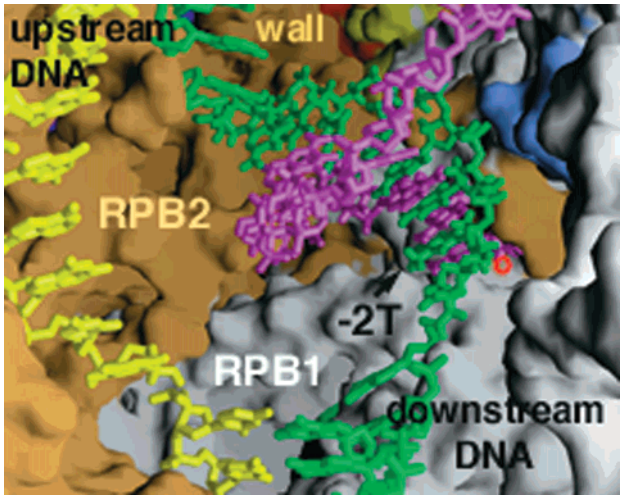

FIGURE 9: Position of the RNA−DNA hybrid in a molecular model of yeast RNAP II in a transcription elongation complex. The hybrid is viewed from above, looking downward into the deep groove partially formed by the flexible domain of RPB1 called the clamp. RPB1 (in white) and RPB2 (tan) form the sides of this groove. A segment of the melted region of the transcription bubble is shown with the template strand in green, the nontemplate strand in yellow, and the RNA in magenta. RPB6 is shown in blue, RPB3 in red, and RBP11 in yellow. The hybrid is positioned in a depression in the surface of RPB1, in front of a domain of RPB2 termed the wall. The wall blocks the trajectory of the downstream DNA; thus, we have modeled the upstream DNA projecting upward from the hybrid. The hybrid is nearly perpendicular to the downstream DNA. The template strand at −2T in the hybrid is close to regions of both RPB1 and RPB2. The approximate location of the active site magnesium is indicated by a red o near the 3′-terminus of the transcript.

we have modeled onto the RNAP structure provided to us by Cramer et al. (*49*) the nucleic acids present in a transcription elongation complex: double-stranded DNA both upstream and downstream of the active site, an 8 bp RNA−DNA hybrid, 18 bases of single-stranded nontemplate DNA and 10 bases of single-stranded template DNA in the transcription bubble, and the RNA transcript 5′ to the hybrid (Figures 9 and 10). The RNA−DNA hybrid in *E. coli* RNAP is 8 bp (*104*), and the transcription bubble is about 18 bp, open from nucleotide −16 to +2. The hinged domain was clearly open in the crystals from which this model of RNAP II was derived, although in a transcribing complex the hinge is likely to partially enclose the hybrid. This domain was referred to as a "clamp" by Cramer et al. (*49*).

The positions along the DNA that we cross-linked are in agreement with the results of DNase I footprinting of the ternary complexes. Our model shows DNA downstream to position +23, and at this position, the DNA protrudes beyond the RNAP, explaining DNase footprinting results and the lack of cross-linking from +25N (Figure 10). The location of the downstream DNA in a yeast RNAP II transcription complex was determined by electron crystallography (*48*). We used the level of total cross-linking of RNAP subunits to each DNA position as a measure of the proximity of the RNAP to the DNA at that position. We detected much more cross-linking from upstream cross-linkers than from the two downstream positions, +7/8N and +12/13N, consistent with upstream DNA being positioned in the deep cleft formed by the hinged domain and the downstream region lying in a shallow groove.

The downstream DNA cannot continue in a straight path through the groove in the RNAP (*49*). First of all, the active site is below the plane of the downstream DNA. Second, a wall composed of RPB2 is behind the active site (looking in toward the active site from the entrance of the downstream DNA; see Figure 10). This wall blocks the path of the DNA. The full extent of the wall is not present in the model shown in Figures 9 and 10 (compare to ref *49*).

RPB1 and RPB2 Are Close Together Near the Active Site. RPB1 and RPB2 are known to bind DNA and form the active site. With the short cross-linker $5N_3$dU placed at −2T, just one base upstream of the last templating base, most of the cross-linking detected was to RPB1, consistent with the finding by Sheng et al. (*56*) that only RPB1 was cross-linked by 4-thio-UMP in the active site. The short cross-linker at −2T is expected to be in the major groove of the cross-linking DNA−RNA hybrid and to detect proteins that are closely associated with the DNA. Although to a lesser extent, RPB2 was also cross-linked by the short cross-linker at −2T. In contrast, with the long cross-linker (N_3RdUTP), RPB2 was labeled much more often than RPB1. Riva et al. (*53*, *54*) cross-linked RPB2 with a dinucleotide that had a cross-linker at the 5′-end, and after limited chain elongation, this probe also cross-linked to RPB1. Indeed, Cramer et al. have detected regions of RPB1 and RPB2 in proximity at the active site in yeast RNAP II (*49*), as was the case in *Taq* RNAP (*42*).

To model the nucleic acids onto the enzyme, we first positioned the RNA−DNA hybrid so that the 3′-ribonucleotide was near the active site and template position −2T was close enough to RPB1 for the short cross-linker to react with it, but also close enough to RPB2 so that the longer cross-linker could react with RPB2 (see Figure 9). A short segment of single-stranded template connects the template strand of the hybrid (shown in an active site depression) with downstream DNA. To position the active site and also to avoid having the hybrid colliding with the wall, we have modeled the hybrid as Cramer et al. have (*49*), nearly perpendicular to the downstream DNA. Such a bend in the template strand would not be unprecedented: the template strand in the active center of *Bacillus* DNA polymerase (*125*) is also deeply bent.

Topography Near the Upstream DNA. The highest frequency of cross-linking was to upstream DNA sites from position −11 to −2. Template and nontemplate positions within this region cross-linked to a similar degree. RPB1 (192 kDa) and RPB2 (139 kDa) comprise about 60% of the bulk of the approximately 550 kDa RNAP II and are implicated in DNA binding, so it is not surprising that one or both of them cross-link at every DNA position that was tested. An X-ray diffraction study suggested that the two largest subunits have similar folds and may bind the DNA between them (*126*). The cross-linking study of a transcription initiation complex with human RNAP II confirmed that the DNA was between RPB1 and RPB2 because the individual subunits cross-linked to opposite faces of the DNA helix (*71*). These researchers detected cross-linking of RPB1 from position −53 to +9 along a single face of the DNA and cross-linking of RPB2 from position −49 to +19, also to a single face of the DNA (except from position −4 to +3). In contrast to the results of Kim et al., we detected more cross-linking to RPB2 than to RPB1 at every position

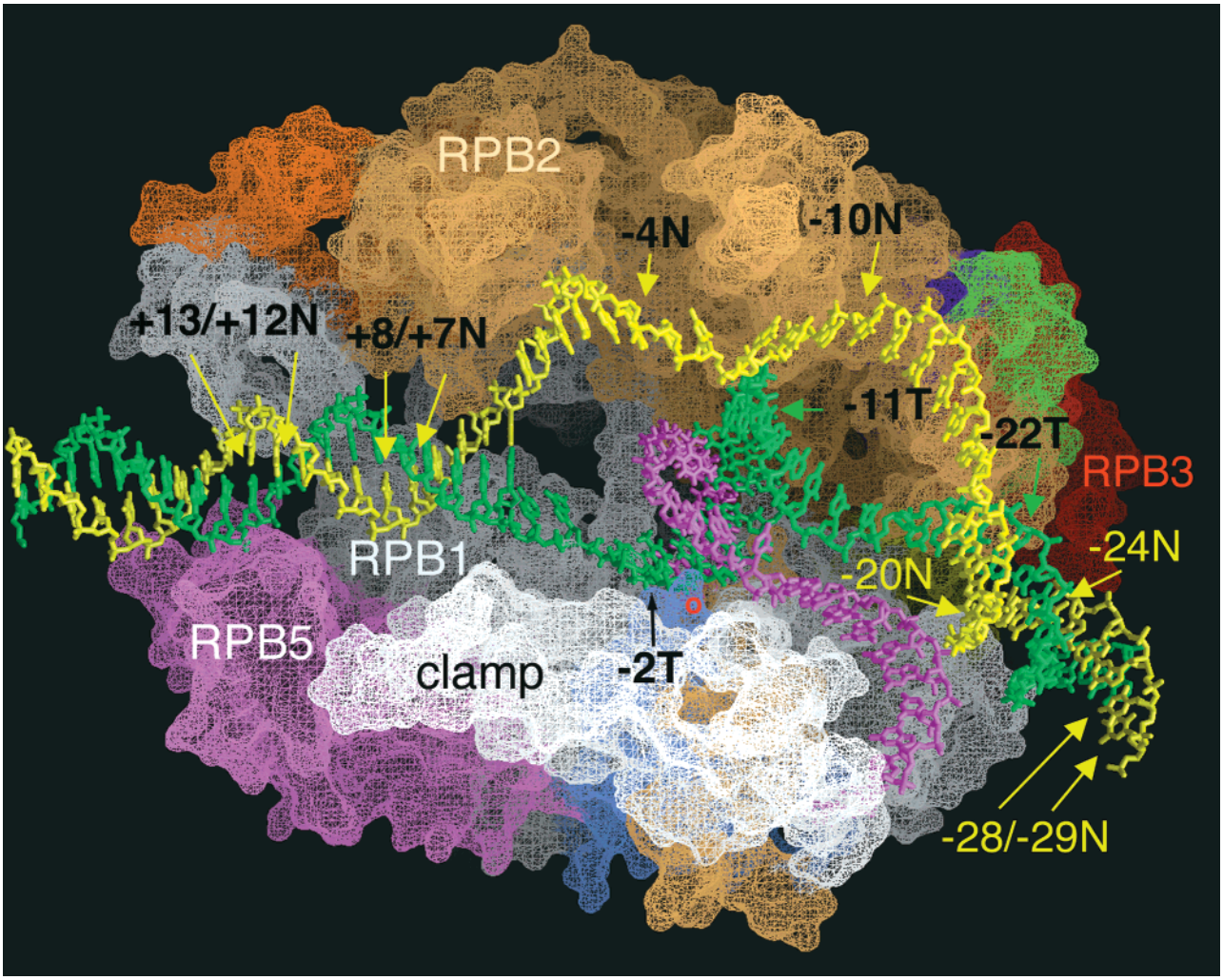

FIGURE 10: Molecular model of yeast RNAP II in a transcription elongation complex with positions of cross-linkers in the DNA. The backbone model for yeast RNA polymerase II was provided by P. Cramer and R. Kornberg (*49*). Positioned onto the RNAP II model are the nucleic acids expected to be present in a transcription elongation complex with 18 bases of open transcription bubble (described in Experimental Procedures). The RNA (in magenta) is shown in an RNA−DNA hybrid of 8 bp with an additional 15 nucleotides extruded in a groove beside the clamp. Sites of cross-linker incorporation in the template strand of the DNA (green) and those in the nontemplate strand (yellow) are indicated. Subunits are colored as in described in ref *49*: RPB1 (white), RPB2 (tan), RPB3 (red), RPB5 (pink), RPB6 (lavender), RPB9 (orange), RPB10 (purple), RPB11 (barely visible, in yellow), and RPB12 (light green). RPB8 is not visible from this view. RPB4 and RPB7 are not present in this model. The active site magnesium is indicated by a red o and is positioned behind and below the RPB6 density shown in this view. RPB2 makes up most of the surface that the upstream DNA traverses. All DNA positions from −10 upstream also label RPB1 at a low level. The RPB1 clamp shown in an open conformation in this model is hypothesized to partially enclose the DNA during transcription. Such a closure would bring RPB1 closer to the upstream DNA in our model.

upstream of nucleotide −2 that we evaluated, with both the long and short cross-linker (our cross-linkers were positioned at −29/−28N, −24N, −22T, −20T, −11T, −10N, and −4N) and detected almost no cross-linking to RPB2 downstream. The differences between our results and those of Kim et al. may be attributed to conformational changes that the RNAP undergoes between preinitiation and becoming a mature elongating polymerase (*47*, *98*). Nontemplate strand position −20 is very nearly on the opposite side of the DNA helix from template strand position −22, yet the relative distribution of cross-linking we detected to RPB1 versus RPB2 was very similar, in the both presence and absence of RNase H to reanneal the DNA strands. When RNase H was used to reanneal the DNA template strands, overall cross-linking at −29/−28 was enhanced 5-fold, and at all other positions, this enhancement was 2−3-fold. The only DNA position likely to be affected by the reannealing would be −29/−28, because the rest of the DNA was either double-stranded or normally single-stranded in a transcription bubble. We have modeled the upstream double-stranded DNA to contact primarily RPB2, but also RPB1. With the long cross-linker at −4N, RPB2 was cross-linked much more than RPB1 and almost exclusively so with the short cross-linker. The nontemplate strand in the transcription bubble was, therefore, modeled close to RPB2 (Figure 10). Inspection of the enzyme structure shows a deep groove to the right side of the clamp as shown in Figure 10. The single-stranded RNA has been modeled in this groove as hypothesized by Cramer et al. (*49*). The groove would be of suitable size to accommodate the upstream double-helical DNA, but we do not believe the helix is positioned in it for the following reasons. The groove is composed of RPB1 to either side, with just a small segment of RPB2 in the center, whereas we detect much more cross-linking to RPB2 from the upstream DNA. Movement of the clamp over the active site would also increase the level of protein contact with both

template and nontemplate DNA strands. As the transcription complex is modeled, cross-linking to RPB1 would have to occur by flexible movement of the single-stranded DNA. An inward shift of the clamp would allow for the lesser but consistent cross-linking of RPB1 from all upstream position. We propose that the upstream DNA bends around the RNAP between RPB1 and RPB2 and toward RPB2 rather than projecting away from the enzyme at −29/−28N, since we detect as much cross-linking to −29/−28N as we do to +12/+13N. We have modeled the DNA in the transcription complex with a bend. We believe the bend is probably even larger than we have shown it, projecting −29/−28N closer to RPB2. The short cross-linker at −29/−28N cross-linked exclusively to RPB2. The overall amount of bending would depend on how much of the upstream DNA wraps against the RNAP. DNA bending has been detected in transcription preinitiation complexes (*71*, *127*, *136*) and in elongation complexes (*128*, *129*). We did not have any cross-linking probes that failed to cross-link to the RNAP upstream; therefore, we could not determine the upstream boundary of the interaction. DNase I footprinting might not be effective to determine if the DNA in an elongation complex wrapped around the RNAP if the DNA and RNAP were not in close contact. The extent of DNA wrapping around the RNAP in tailed template complexes in vitro may differ from the in vivo situation as well. One would be hard-pressed to distinguish DNA wrapping from RNAP back-sliding if cross-linking were detected further upstream (at position −40, for example).

Our data would also support a model in which the upstream DNA was within the groove formed by the clamp and bent back toward the downstream DNA, as long as both template and nontemplate strands in the transcription bubble were closer to RPB2 than RPB1. Such a model would have to take into account the fact that the DNA is protected from nuclease digestion (not protrude too far from the surface of the enzyme) and that the DNA strands would fit within this groove. A greater understanding of the conformation of the upstream DNA would be aided by cross-linking studies that identify residues of the RNAP in contact with the DNA at different positions.

Cross-Linking to RPB5. RPB1, RPB2, and RPB5 are the subunits that cross-link to DNA along the protein−DNA interface of a transcription elongation complex. In our RNAP II ternary complexes, RPB5 was cross-linked at downstream positions +7/8N and +12/13N and also upstream at positions −10N, −20N, and −29/−28N. Bartholomew et al. (*68*) also detected this 27 kDa subunit common to RNAP I, II, and III in photoaffinity cross-linking studies of yeast RNAP III. In ternary complexes that had transcribed 17 bases, this subunit cross-linked only at upstream position −7N, but in binary (nontranscribing) initiation complexes, this subunit cross-linked at downstream positions +6, +11, and +16. Kim et al. (*71*) cross-linked RPB5 from position +5 to +15 on a single face of the DNA. The downstream DNA in our model positions +7/+8N cross-linkers to react primarily with RPB1 and at a very low level with RPB5. Cross-linkers at +12/+13N react primarily with RPB5, consistent with the location of RPB5 determined by Cramer et al. (*49*). Cross-linking of RPB5 from upstream positions is perplexing. The frequency of RPB5 cross-linking from any of these positions is low: −20N > −29/−28N > −10N. The template strand is expected to be more spatially confined in the region of the RNA−DNA hybrid than the single-stranded nontemplate strand, especially during transcription from templates that are not topologically constrained, such as tailed templates. Flexibility of the nontemplate strand to swing back and forth could account for the cross-linking to RPB5. The cross-linker at −4N may not react with RPB5 either because it is more confined or because the single-stranded DNA at that position in the transcription bubble is not long enough to reach back toward RPB5. Alternatively or additionally, in a small proportion of transcription complexes the nontemplate strand may be in a groove that loops around the clamp back toward the downstream DNA. Such a situation would also explain the low level of cross-linking of RPB7 from −29/−28N. The RNA may normally be positioned in such a groove to hold the clamp closed during transcription elongation [as Cramer et al. (*49*) have proposed]. In the absence of RNase H, −29/−28N cross-links to RPB7 at a low level. Under these conditions, some of the template strands are hybridized to RNA. RPB7 is positioned near RPB5 (*49*).

Cross-Linking to Other Subunits. We noticed that RPB4 was cross-linked to our DNA probes sometimes when the salt concentration was not high enough, but we never saw it cross-linked in a transcription-dependent manner. The part of the gel where RPB4 migrates is well-separated from other polypeptides so, if it had been cross-linked in transcription reactions, we would very likely have seen it. This result is quite intriguing because RPB4 and RPB7 form a subassembly that is not stoichiometrically present in cells that are isolated from the log phase of growth, but is stoichiometrically present in cells grown to stationary phase. The RPB4−RPB7 subassembly has a function in promoter-directed initiation (*116*), but is not required for nonspecific transcription. RPB7 is an essential gene, but RPB4 is only essential under some conditions (*117*, *130*). Previously, RPB7 was believed to associate with the RNAP only in the presence of RPB4, but this view is changing because overexpression of RPB7 can relieve the effects of RPB4 deletion (*131*). RPB4 and RPB7 were present in all of our yeast RNAP II preparations; thus, detection of RPB7 but not RPB4 in ternary complexes leads us to speculate whether RPB4 could have a function related to that of the *E. coli* sigma factor. Some sequence similarity between RPB4 and *E. coli* sigma 70 has been reported (*132*). Transcription elongation is a function of the *E. coli* core RNAP. Association of one of several sigma factors with core RNAP allows the RNAP to initiate transcription, but then sigma is released after transcription of about 8−10 bases. Is it possible that the removal of RPB4 from an initiating complex may be required for the conformational change termed the "mature" elongation form (*98*)?

Comparison with Yeast RNAP III Transcription Complexes. In transcription initiation complexes that were photoaffinity cross-linked by Bartholomew et al. (*68*), the counterparts to RNAP II subunits RPB1 and RPB2 and the common 27 kDa subunit (RPB5) were cross-linked to the DNA. RNAP III ternary complexes that had transcribed 17 bases exhibited cross-linking results similar to ours with N_3-RdU at the following DNA positions (*68*). Only the largest RNAP III subunit (RPC1) cross-linked 12 bases downstream; the second-largest subunit (RPC2) cross-linked more than RPC1 from 2 bases and from 19 or 20 bases upstream, and RPC1 and RPC2 cross-linked approximately the same

amount from 7 bases upstream (similar to our results at −10N). The biggest difference between the cross-linking results presented here with RNAP II and the results of Bartholomew et al. with RNAP III is that there are additional subunits that cross-link at DNA positions 20 or 19, 12 or 11, and 7 bases upstream of the active site in RNAP III complexes: C82, C40 (or 37), C34, and C31. Subunits of 82, 53, and 34 kDa were cross-linked in binary complexes (that had not initiated transcription). In ternary complexes, subunits of 82, 40 or 37, 34, and 31 kDa were cross-linked in addition to the largest two subunits and RPB5. Curiously, a subassembly of 82, 37, and 34 kDa subunits sometimes is not copurified with RNAP III (*133*). The 37 kDa subunit has been reported to be missing from some preparations of the fully active wild-type enzyme (*134*). A mutation of the largest subunit of RNAP III was found to result in loss of the 82, 34, and 31 kDa subunits (*134*), indicating a close association of these subunits. The RNAP III lacking these subunits showed a thermosensitive phenotype, much like the RPB4 mutants. Another of the cross-linked RNAP III subunits, RPC53, was found to have a function in the cell cycle. Thus, four of the RNAP III subunits that were cross-linked may have a role more in regulation than in transcription elongation alone, as RPB4 and possibly RPB7 do. It would be of great interest to see which yeast RNAP III subunits cross-link in elongation complexes that have transcribed well beyond the initiation site.

ACKNOWLEDGMENT

For purifying yeast RNAP II, we thank Lee Strasheim, Kit Foley, Kaylene Chambers, and Dr. Monika de Arruda. Dr. Nancy Thompson and Kit Foley provided antibodies for verifying the identities of RNAP II subunits. We thank Roger Kornberg, Patrick Cramer, Jianhua Fu, Grant Jensen, Avi Gnatt, and David Bushnell (Stanford University, Stanford, CA) for sharing with us structural data, protein models, and electron density coordinates for yeast RNAP II. Dr. Jean-Yves Sgro provided the expertise for all of the computer modeling. We thank him for hours of dedicated effort.

REFERENCES

1. Bartholomew, B., Kassavetis, G. A., Braun, B. R., and Geiduschek, E. P. (1990) *EMBO J. 9*, 2197−205.
2. Spencer, C. A., and Groudine, M. (1990) *Oncogene 5*, 777−85.
3. Reines, D., Conaway, J. W., and Conaway, R. C. (1996) *Trends Biochem. Sci. 21*, 351−5.
4. Uptain, S. M., Kane, C. M., and Chamberlin, M. J. (1997) *Annu. Rev. Biochem. 66*, 117−72.
5. Shilatifard, A. (1998) *FASEB J. 12*, 1437−46.
6. Sentenac, A. (1985) *CRC Crit. Rev. Biochem. 18*, 31−90.
7. Young, R. A. (1991) *Annu. Rev. Biochem. 60*, 689−715.
8. Thuriaux, P., and Sentenac, A. (1992) in *The Molecular and Cellular Biology of the Yeast Saccharomyces: Gene Expression*, pp 1−49, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
9. Archambault, J., and Friesen, J. D. (1993) *Microbiol. Rev. 57*, 703−24.
10. Woychik, N. A., Liao, S. M., Kolodziej, P. A., and Young, R. A. (1990) *Genes Dev. 4*, 313−23.
11. Woychik, N. A., and Young, R. A. (1990) *J. Biol. Chem. 265*, 17816−9.
12. Treich, I., Carles, C., Riva, M., and Sentenac, A. (1992) *Gene Expression 2*, 31−7.
13. Lalo, D., Carles, C., Sentenac, A., and Thuriaux, P. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 5524−8.
14. Allison, L. A., Moyle, M., Shales, M., and Ingles, C. J. (1985) *Cell 42*, 599−610.
15. Jokerst, R. S., Weeks, J. R., Zehring, W. A., and Greenleaf, A. L. (1989) *Mol. Gen. Genet. 215*, 266−75.
16. Sweetser, D., Nonet, M., and Young, R. A. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 1192−6.
17. Falkenburg, D., Dworniczak, B., Faust, D. M., and Bautz, E. K. (1987) *J. Mol. Biol. 195*, 929−37.
18. Woychik, N. A., and Young, R. A. (1990) *Trends Biochem. Sci. 15*, 347−51.
19. Woychik, N., and Young, R. (1994) in *Transcription: Mechanisms and Regulation* (Conaway, R. C., and Conaway, J. W., Eds.) pp 227−40, Raven Press, New York.
20. Langer, D., Hain, J., Thuriaux, P., and Zillig, W. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 5768−72.
21. Svetlov, V., Nolan, K., and Burgess, R. R. (1998) *J. Biol. Chem. 273*, 10827−30.
22. McKune, K., and Woychik, N. A. (1994) *Mol. Cell. Biol. 14*, 4155−9.
23. Khazak, V., Sadhale, P. P., Woychik, N. A., Brent, R., and Golemis, E. A. (1995) *Mol. Biol. Cell 6*, 759−75.
24. McKune, K., Moore, P. A., Hull, M. W., and Woychik, N. A. (1995) *Mol. Cell. Biol. 15*, 6895−900.
25. Shpakovski, G. V., Acker, J., Wintzerith, M., Lacroix, J. F., Thuriaux, P., and Vigneron, M. (1995) *Mol. Cell. Biol. 15*, 4702−10.
26. Gundelfinger, E. D. (1983) *FEBS Lett. 157*, 133−8.
27. Horikoshi, M., Tamura, H., Sekimizu, K., Obinata, M., and Natori, S. (1983) *J. Biochem. 94*, 1761−7.
28. Chuang, R. Y., and Chuang, L. F. (1987) *Biochem. Biophys. Res. Commun. 145*, 73−80.
29. Kontermann, R. E., Kobor, M., and Bautz, E. K. (1993) *Protein Sci. 2*, 223−30.
30. Carroll, S. B., and Stollar, B. D. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 7233−7.
31. Carroll, S. B., and Stollar, B. D. (1983) *J. Mol. Biol. 170*, 777−90.
32. Delarue, M., Poch, O., Tordo, N., Moras, D., and Argos, P. (1990) *Protein Eng. 3*, 461−7.
33. Bonner, G., Patra, D., Lafer, E. M., and Sousa, R. (1992) *EMBO J. 11*, 3767−75.
34. Beese, L. S., Derbyshire, V., and Steitz, T. A. (1993) *Science 260*, 352−5.
35. Jacobo-Molina, A., Ding, J., Nanni, R. G., Clark, A. D., Jr., Lu, X., Tantillo, C., Williams, R. L., Kamer, G., Ferris, A. L., Clark, P., et al. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 6320−4.
36. Sawaya, M. R., Pelletier, H., Kumar, A., Wilson, S. H., and Kraut, J. (1994) *Science 264*, 1930−5.
37. Pelletier, H., Sawaya, M. R., Kumar, A., Wilson, S. H., and Kraut, J. (1994) *Science 264*, 1891−903.
38. Joyce, C. M., and Steitz, T. A. (1994) *Annu. Rev. Biochem. 63*, 777−822.
39. Kim, W. J., Burke, L. P., and Mortin, M. A. (1994) *J. Mol. Biol. 244*, 13−22.
40. Zaychikov, E., Martin, E., Denissova, L., Kozlov, M., Markovtsov, V., Kashlev, M., Heumann, H., Nikiforov, V., Goldfarb, A., and Mustaev, A. (1996) *Science 273*, 107−9.
41. Guajardo, R., and Sousa, R. (1997) *J. Mol. Biol. 265*, 8−19.
42. Zhang, G., Campbell, E. A., Minakhin, L., Richter, C., Severinov, K., and Darst, S. A. (1999) *Cell 98*, 811−24.
43. Kimura, M., Ishiguro, A., and Ishihama, A. (1997) *J. Biol. Chem. 272*, 25851−5.
44. Fu, J., Gnatt, A. L., Bushnell, D. A., Jensen, G. J., Thompson, N. E., Burgess, R. R., David, P. R., and Kornberg, R. D. (1999) *Cell 98*, 799−810.
45. Darst, S. A., Kubalek, E. W., and Kornberg, R. D. (1989) *Nature 340*, 730−2.
46. Polyakov, A., Severinova, E., and Darst, S. A. (1995) *Cell 83*, 365−73.
47. Asturias, F. J., Meredith, G. D., Poglitsch, C. L., and Kornberg, R. D. (1997) *J. Mol. Biol. 272*, 536−40.

48. Poglitsch, C. L., Meredith, G. D., Gnatt, A. L., Jensen, G. J., Chang, W. H., Fu, J., and Kornberg, R. D. (1999) *Cell 98*, 791−8.
49. Cramer, P., Bushnell, D. A., Fu, J., Gnatt, A. L., Maier-Davis, B., Thompson, N. E., Burgess, R. R., Edwards, A. M., David, P. R., and Kornberg, R. D. (2000) *Science 288*, 640−9.
50. Jensen, G. J., Meredith, G., Bushnell, D. A., and Kornberg, R. D. (1998) *EMBO J. 17*, 2353−8.
51. Grachev, M. A., Hartmann, G. R., Maximova, T. G., Mustaev, A. A., Schaffner, A. R., Sieber, H., and Zaychikov, E. F. (1986) *FEBS Lett. 200*, 287−90.
52. Grachev, M. A., Kolocheva, T. I., Lukhtanov, E. A., and Mustaev, A. A. (1987) *Eur. J. Biochem. 163*, 113−21.
53. Riva, M., Schaffner, A. R., Sentenac, A., Hartmann, G. R., Mustaev, A. A., Zaychikov, E. F., and Grachev, M. A. (1987) *J. Biol. Chem. 262*, 14377−80.
54. Riva, M., Carles, C., Sentenac, A., Grachev, M. A., Mustaev, A. A., and Zaychikov, E. F. (1990) *J. Biol. Chem. 265*, 16498−503.
55. Treich, I., Carles, C., Sentenac, A., and Riva, M. (1992) *Nucleic Acids Res. 20*, 4721−5.
56. Sheng, N., Mougey, E. B., Kelly, S., and Dennis, D. (1993) *Biochemistry 32*, 2248−53.
57. Powell, W., and Reines, D. (1996) *J. Biol. Chem. 271*, 6866−73.
58. Kadesch, T. R., and Chamberlin, M. J. (1982) *J. Biol. Chem. 257*, 5286−95.
59. Coulter, D. E., and Greenleaf, A. L. (1985) *J. Biol. Chem. 260*, 13190−8.
60. Dedrick, R. L., Kane, C. M., and Chamberlin, M. J. (1987) *J. Biol. Chem. 262*, 9098−108.
61. Sluder, A. E., Price, D. H., and Greenleaf, A. L. (1988) *J. Biol. Chem. 263*, 9917−25.
62. Reinberg, D., and Roeder, R. G. (1987) *J. Biol. Chem. 262*, 3331−7.
63. Reines, D., Wells, D., Chamberlin, M. J., and Kane, C. M. (1987) *J. Mol. Biol. 196*, 299−312.
64. Sluder, A. E., Greenleaf, A. L., and Price, D. H. (1989) *J. Biol. Chem. 264*, 8963−9.
65. SivaRaman, L., Reines, D., and Kane, C. M. (1990) *J. Biol. Chem. 265*, 14554−60.
66. Wiest, D. K., Wang, D., and Hawley, D. K. (1992) *J. Biol. Chem. 267*, 7733−44.
67. Bradsher, J. N., Tan, S., McLaury, H. J., Conaway, J. W., and Conaway, R. C. (1993) *J. Biol. Chem. 268*, 25594−603.
68. Bartholomew, B., Durkovich, D., Kassavetis, G. A., and Geiduschek, E. P. (1993) *Mol. Cell. Biol. 13*, 942−52.
69. Williams, K. R., and Konigsberg, W. H. (1991) *Methods Enzymol. 208*, 516−39.
70. Meisenheimer, K. M., and Koch, T. H. (1997) *Crit. Rev. Biochem. Mol. Biol. 32*, 101−40.
71. Kim, T. K., Lagrange, T., Wang, Y. H., Griffith, J. D., Reinberg, D., and Ebright, R. H. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 12268−73.
72. Hanna, M. M. (1989) *Methods Enzymol. 180*, 383−409.
73. Pinol-Roma, S., Adam, S. A., Choi, Y. D., and Dreyfuss, G. (1989) *Methods Enzymol. 180*, 410−8.
74. Hockensmith, J. W., Kubasek, W. L., Vorachek, W. R., Evertsz, E. M., and von Hippel, P. H. (1991) *Methods Enzymol. 208*, 211−36.
75. Buckle, M., Geiselmann, J., Kolb, A., and Buc, H. (1991) *Nucleic Acids Res. 19*, 833−40.
76. Pashev, I. G., Dimitrov, S. I., and Angelov, D. (1991) *Trends Biochem. Sci. 16*, 323−6.
77. Bartholomew, B., Kassavetis, G. A., and Geiduschek, E. P. (1991) *Mol. Cell. Biol. 11*, 5181−9.
78. Bartholomew, B., Braun, B. R., Kassavetis, G. A., and Geiduschek, E. P. (1994) *J. Biol. Chem. 269*, 18090−5.
79. Braun, B. R., Bartholomew, B., Kassavetis, G. A., and Geiduschek, E. P. (1992) *J. Mol. Biol. 228*, 1063−77.
80. Kassavetis, G. A., Joazeiro, C. A., Pisano, M., Geiduschek, E. P., Colbert, T., Hahn, S., and Blanco, J. A. (1992) *Cell 71*, 1055−64.
81. Staros, J. V., Bayley, H., Standring, D. N., and Knowles, J. R. (1978) *Biochem. Biophys. Res. Commun. 80*, 568−72.
82. Czarnecki, J., Geahlen, R., and Haley, B. (1979) *Methods Enzymol. 56*, 642−53.
83. Wooddell, C. I. (1999) Ph.D. Thesis, University of Wisconsin, Madison, WI.
84. Thompson, N. E., Aronson, D. B., and Burgess, R. R. (1990) *J. Biol. Chem. 265*, 7069−77.
85. Edwards, A. M., Darst, S. A., Feaver, W. J., Thompson, N. E., Burgess, R. R., and Kornberg, R. D. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 2122−6.
86. Wooddell, C. I., and Burgess, R. R. (1996) *Genome Res. 6*, 886−92.
87. Dedrick, R. L., and Chamberlin, M. J. (1985) *Biochemistry 24*, 2245−53.
88. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning, A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
89. Laemmli, U. K. (1970) *Nature 227*, 680−5.
90. Oakley, B. R., Kirsch, D. R., and Morris, N. R. (1980) *Anal. Biochem. 105*, 361−3.
91. Ferrin, T. E., Huang, C. C., Jarvis, L. E., and Langridge, R. (1988) *J. Mol. Graphics 6*, 13−27, 36−7.
92. Ferrin, T. E., Couch, G. S., Huang, C. C., Pettersen, E. F., and Langridge, R. (1991) *J. Mol. Graphics 9*, 27−32, 37−8.
93. Kane, C. M., and Chamberlin, M. J. (1985) *Biochemistry 24*, 2254−62.
94. Edwards, A. M., and Kane, C. M. (1996) *Methods Enzymol. 274*, 419−36.
95. Sekeris, C. E., Schmid, W., and Roewekamp, W. (1972) *FEBS Lett. 24*, 27−31.
96. Lescure, B., Dauguet, C., and Yaniv, M. (1978) *J. Mol. Biol. 124*, 87−96.
97. Kane, C. M. (1988) *Biochemistry 27*, 3187−96.
98. Samkurashvili, I., and Luse, D. S. (1998) *Mol. Cell. Biol. 18*, 5343−54.
99. Linn, S. C., and Luse, D. S. (1991) *Mol. Cell. Biol. 11*, 1508−22.
100. Rice, G. A., Chamberlin, M. J., and Kane, C. M. (1993) *Nucleic Acids Res. 21*, 113−8.
101. Selby, C. P., Drapkin, R., Reinberg, D., and Sancar, A. (1997)- *Nucleic Acids Res. 25*, 787−93.
102. Reeder, T. C., and Hawley, D. K. (1996) *Cell 87*, 767−77.
103. Komissarova, N., and Kashlev, M. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 1755−60.
104. Nudler, E., Mustaev, A., Lukhtanov, E., and Goldfarb, A. (1997) *Cell 89*, 33−41.
105. Pati, U. K. (1994) *Gene 145*, 289−92.
106. Ulmasov, T., Larkin, R. M., and Guilfoyle, T. J. (1996) *J. Biol. Chem. 271*, 5085−94.
107. Larkin, R. M., and Guilfoyle, T. J. (1997) *J. Biol. Chem. 272*, 12824−30.
108. Korobko, I. V., Yamamoto, K., Nogi, Y., and Muramatsu, M. (1997) *Gene 185*, 1−4.
109. Hayward, R. S., Igarashi, K., and Ishihama, A. (1991) *J. Mol. Biol. 221*, 23−9.
110. Igarashi, K., Hanamura, A., Makino, K., Aiba, H., Aiba, H., Mizuno, T., Nakata, A., and Ishihama, A. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 8958−62.
111. Igarashi, K., and Ishihama, A. (1991) *Cell 65*, 1015−22.
112. Thomas, M. S., and Glass, R. E. (1991) *Mol. Microbiol. 5*, 2719−25.
113. Ross, W., Gosink, K. K., Salomon, J., Igarashi, K., Zou, C., Ishihama, A., Severinov, K., and Gourse, R. L. (1993) *Science 262*, 1407−13.
114. Chamberlin, M. (1995) *Harvey Lect. 88*, 1−21.
115. Huet, J., Dezelee, S., Iborra, F., Buhler, J. M., Sentenac, A., and Fromageot, P. (1976) *Biochimie 58*, 71−80.
116. Edwards, A. M., Kane, C. M., Young, R. A., and Kornberg, R. D. (1991) *J. Biol. Chem. 266*, 71−5.
117. Choder, M., and Young, R. A. (1993) *Mol. Cell. Biol. 13*, 6984−91.
118. Cheetham, G. M. T., and Steitz, T. A. (1999) *Science 286*, 2305−9.

119. Guajardo, R., Lopez, P., Dreyfus, M., and Sousa, R. (1998) *J. Mol. Biol. 281*, 777−92.
120. Gu, W., Wind, M., and Reines, D. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 6935−40.
121. Arndt, K. M., and Chamberlin, M. J. (1990) *J. Mol. Biol. 213*, 79−108.
122. Krummel, B., and Chamberlin, M. J. (1992) *J. Mol. Biol. 225*, 221−37.
123. Nudler, E., Goldfarb, A., and Kashlev, M. (1994) *Science 265*, 793−6.
124. Darst, S. A., Edwards, A. M., Kubalek, E. W., and Kornberg, R. D. (1991) *Cell 66*, 121−8.
125. Kiefer, J. R., Mao, C., Braman, J. C., and Beese, L. S. (1998) *Nature 391*, 304−7.
126. Fu, J., Gerstein, M., David, P. R., Gnatt, A. L., Bushnell, D. A., Edwards, A. M., and Kornberg, R. D. (1998) *J. Mol. Biol. 280*, 317−22.
127. Forget, D., Robert, F., Grondin, G., Burton, Z. F., Greenblatt, J., and Coulombe, B. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 7150−5.
128. ten Heggeler-Bordier, B., Wahli, W., Adrian, M., Stasiak, A., and Dubochet, J. (1992) *EMBO J. 11*, 667−72.
129. Rees, W. A., Keller, R. W., Vesenka, J. P., Yang, G., and Bustamante, C. (1993) *Science 260*, 1646−9.
130. Woychik, N. A., and Young, R. A. (1989) *Mol. Cell. Biol. 9*, 2854−9.
131. Sheffer, A., Varon, M., and Choder, M. (1999) *Mol. Cell. Biol. 19*, 2672−80.
132. McCracken, S., and Greenblatt, J. (1991) *Science 253*, 900−2.
133. Valenzuela, P., Hager, G. L., Weinberg, F., and Rutter, W. J. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 1024−8.
134. Werner, M., Hermann-Le Denmat, S., Treich, I., Sentenac, A., and Thuriaux, P. (1992) *Mol. Cell. Biol. 12*, 1087−95.
135. Nicholls, A., Sharp, K., and Honig, B. (1991) *Proteins: Struct., Funct., Genet. 11*, 281.
136. Robert, F., Douziech, M., Forget, D., Egly, J.-M., Greenblatt, J., Burton, Z. F., and Coulombe, B. (1998) *Mol. Cell 2*, 341−51.

BI0014249