Proc. Natl. Acad. Sci. USA
Vol. 95, pp. 9111–9116, August 1998
Biochemistry

RNA-binding site in T7 RNA polymerase

SRIN SASTRY* AND BARBARA M. ROSS

Laboratory of Molecular Genetics, Box 174, The Rockefeller University, 1230 York Avenue, New York, NY 10021

Edited by Michael J. Chamberlin, University of California, Berkeley, CA, and approved June 1, 1998 (received for review February 3, 1998)

ABSTRACT **Recent models of RNA polymerase transcription complexes have invoked the idea that enzyme-nascent RNA contacts contribute to the stability of the complexes. Although much progress on this topic has been made with the multisubunit *Escherichia coli* RNA polymerase, there is a paucity of information regarding the structure of single-subunit phage RNA polymerase transcription complexes. Here, we photo-cross-linked the RNA in a T7 RNA polymerase transcription complex and mapped a major contact site between amino acid residues 144 and 168 and probably a minor contact between residues 1 and 93. These regions of the polymerase are proposed to interact with the emerging RNA during transcription because the 5′ end of the RNA was cross-linked. The contacts are both ionic and nonionic (hydrophobic). The specific inhibitor of T7 transcription, T7 lysozyme, does not compete with T7 RNA polymerase for RNA cross-linking, implying that the RNA does not bind the lysozyme. However, lysozyme may act indirectly via a conformational change in the polymerase. In the current model, the DNA template lies in the polymerase cleft and the fingers subdomain may contact or maintain a template bubble, and a region in the N terminus forms a partly solvent-accessible binding channel for the emerging RNA.**

Gene regulation occurs primarily through mechanisms that modify the behavior of DNA-dependent RNA polymerases (RNAPs) (1, 2). The essential structural framework of all transcribing RNAPs has been conserved (1). This includes template DNA-binding sites in the enzyme, an RNA-DNA hybrid in the single-stranded "bubble" of the DNA template, a catalytic center for RNA synthesis, and binding sites for the nascent RNA transcript (1). Recently, the stability and movement of RNAP has been a topic of intense research and much controversy. Early models suggested that the RNA-DNA hybrid and the transcription bubble were important determinants for transcription complex stability (3, 4). On the other hand, the inchworming model (5) postulated that RNA- and DNA-polymerase-binding sites were pivotal for stability (1, 5). The latest models described RNAP as a sliding clamp (3) in various stages of active, translocated, and arrested states on the template (6, 7), with a definite role for the RNA-DNA hybrid in sliding (8).

Although the models described above were derived largely from studies with *Escherichia coli* RNAP, there is a lack of information regarding the movement of T7 RNAP (9) on the DNA template and the architectural elements holding the transcribing complexes together. This is because T7 RNAP transcription complexes are notoriously unstable for biochemical work. T7 RNAP is particularly attractive for structural studies because, unlike bacterial and eukaryotic RNAPs, it is monomeric (molecular mass = 98.8 kDa) and transcribes DNA without accessory factors. The T7 RNAP crystallographic model (10) is shaped like a partly open right hand with a palm-thumb-fingers configuration, which is common among all nucleic acid polymerases (11). Mutagenesis revealed the catalytically important residues (e.g., see refs. 12–17). Two conserved aspartates D537 and D812 are in the active center, a specificity loop (amino acids 742–773) interacts with the promoter and an N-terminal region contains an RNA-binding site. Earlier, we demonstrated that the promoter DNA was bound in the T7 RNAP cleft and a segment of the fingers subdomain interacted with single-stranded DNA (18, 19). Because RNA-binding sites are an important architectural framework of transcription complexes (e.g., as discussed in ref. 1), mapping these sites is crucial for a structural understanding of T7 RNAP transcription. Here, we directly demonstrate that the 5′ end of the emerging nascent RNA interacts with regions that partly form α-helices (88–168 or more specifically 144–168 and amino acids 1–93) in the N-terminal region of T7 RNAP. In our model, the 5′ end of the nascent RNA is held by these regions in a channel and the fingers contact the transcription bubble that is open for RNA synthesis while the thumb wraps around the upstream part of the DNA template to stabilize the complex.

MATERIALS AND METHODS

T7 RNAP was prepared according to published procedures (20, 21). T7 lysozyme was a gift from the laboratory of F. W. Studier (Brookhaven National Labs, Upton, NY). Plasmid pBS (Stratagene), which contained a T7 RNAP promoter, was prepared using the standard procedures (22).

Synthesis of GTP-γ-Benzophenone (BP) (23). Briefly, 25 mg of GTP (Sigma) was dissolved in 1 ml of 100 mM Mes buffer (pH 5.2; Sigma) followed by the addition of 477 mg of 1-cyclohexyl-3,2-(morpholinoethyl)carbodiimde metho-*p*-toulenesulfonate (a coupling agent; Aldrich). To this whitish solution was added 200 mg of benzyltributylammonium chloride (a phase transfer catalyst; Aldrich). The mixture was fully dissolved by gently shaking at room temperature (25°C) for 2.5 h. To this was added 158 mg of 4-aminobenzophenone (Aldrich) in 1 ml of dimethylforamide (Aldrich). The mixture was stirred for 5 h at room temperature, after which 1 ml of water was added. Portions of this solution were either subjected to TLC or loaded on a 10-cm long × 2-cm i.d. DEAE-Sepharose (Pharmacia) column or on an anion exchange HPLC column. GTP-γ-BP was eluted with a linear gradient of 0–0.5 M triethylammonium acetate or triethylammonium bicarbonate. The GTP-γ-BP-containing fraction (20–30% yield) was identified by its absorption spectrum (Fig. 1*A*) and with polyethyleneimine-cellulose TLC (R_f = 0.65). The material was concentrated by lyophilization and redissolved in MeOH/H_2O and stored at −20°C. GTP-γ-BP is also available from Research Products International.

Transcription and Cross-Linking Reactions. Standard transcription reactions (60 μl) contained 2 μg of plasmid pBS, 110 μM GTP-BP (and/or cold GTP or 1 mM NTP), 0.1 μM [α-^{32}P]NTP (specific activity, 3,000 Ci/mmol), and 4 μg of T7 RNAP in transcription buffer [30 mM Hepes (pH 7.8), 100 mM K^+ glutamate, 15 mM $Mg(OAc)_2$, 1 mM DTT, and 0.05%

This paper was submitted directly (Track II) to the *Proceedings* office. Abbreviations: RNAP, RNA polymerase; BP, benzophenone; MALDI-TOF, matrix-assisted laser desorption ionization time of flight.

*To whom reprint requests should be addressed. e-mail: sastrys@rockvax.rockefeller.edu.

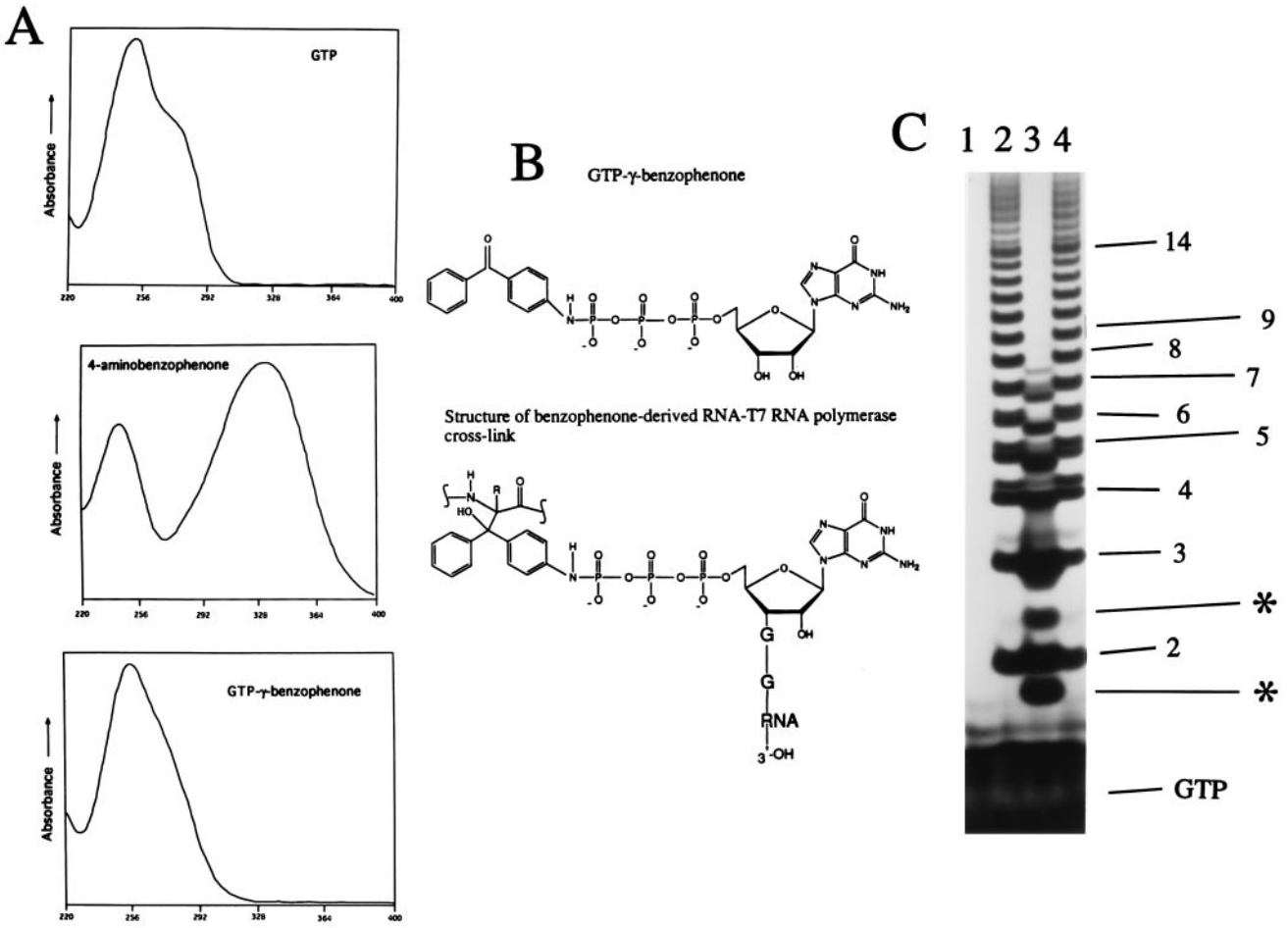

FIG. 1. (*A*) UV-VIS absorption spectra of GTP, 4-aminobenzophenone and GTP-γ-BP. (*B*) Structure of GTP-γ-BP and the presumed structure of 5′-BP-GTP-RNA-T7 RNAP cross-links. RNA cross-link with a generic polypeptide backbone unit is depicted. (*C*) Transcription with cold GTP-BP. Lane 1, no polymerase; lanes 2 and 4, normal G-ladder; and lane 3, G-ladder with cold GTP-γ-BP. Asterisks show examples of BP-RNAs. In lanes 1, 2, and 4, the total concentration of unmodified cold GTP plus [α-^{32}P]GTP was 280 μM. In lane 3, to better visualize the GTP-γ-BP-initiated transcripts, the reaction was biased with a higher concentration of cold GTP-γ-BP (250 μM) compared with the total concentration (30 μM) of unmodified cold GTP plus [α-^{32}P]GTP. [α-^{32}P]GTP was the only radio nucleotide in these reactions.

Tween 20). The mixture was incubated at 37°C for 10 min, distributed into microtiter wells, and then irradiated for 15 min in a Stratalinker box at a distance of 4 cm from the UV lamps (maximum spectral irradiance at 254 nm). The samples were made 50 mM EDTA plus 1% SDS, heated at 95°C for 5–10 min, and run either on 10 or 12% acrylamide-SDS Bio-Rad minigels or on larger gels (12-cm long × 16-cm wide). The cross-links were detected by autoradiography.

Large Scale Cross-Linking and Purification. The reactions were scaled up for 10 mg of T7 RNAP using cold GTP-BP plus GTP plus [α-^{32}P]CTP (which produced the transcript GGGCG). The cross-links were precipitated with saturated $(NH_4)_2SO_4$ and redissolved in water. Proteolyzed fragments were run on Tris-Gly or Tris-Tricine gels (see legends to figures) and sized using Bio-Rad protein standards (catalogue no. 161-0305/0309) and the proteolytic fragments from unmodified T7 RNAP alone. The size determinations were within 200–500 Da. Cross-links in SDS gel pieces were digested with V8, and the proteolyzed fragments were either passively or electro eluted. The RNA cross-linked peptides were adsorbed onto strong quaternary amine anion exchange filters (as recommended by the manufacturer; Sartorius). The filters were thoroughly washed with 10 mM KCl-water; the cross-linked fragments were eluted with 1 M KCl, dialyzed extensively against water, and lyophilized. Additional purification was done as in ref. 10. MALDI-TOF mass spectrometry was done using α-cyano-4-hydroxycinnamic acid as the matrix (24, 25) in a VOYAGER-RP workstation. (PerSeptive Biosystems, Framingham, MA). The instrument was calibrated using external protein standards (from PerSeptive Biosystems) and an internal standard (lysozyme). Mass determinations were accurate to ≈0.01% in the 2–15 kDa linear mass range.

RESULTS

Cross-Linking 5′ End of the Nascent RNA. We synthesized a GTP derivative linked to a photoreactive BP moiety, which was attached to the γ-phosphorous (Fig. 1*B*). Upon excitation with 254 nm of UV irradiation, the BP (Fig. 1*B*) cross-links nonspecifically via hydrogen abstraction from the C-α of the nearest polypeptide backbone (26, 27). The cross-link (Fig. 1*B*) is most probably a benzpinacol type of compound (27). Fig. 1*C* is a typical oligo(rG) synthesis reaction in the presence of GTP-γ-BP. To specifically visualize GTP-γ-BP-initiated transcripts, an excess of GTP-γ-BP over GTP was used in this reaction (see Fig. 1*C* legend). The GTP-γ-BP-initiated transcripts appear to have faster electrophoretic mobility compared with the wild-type transcripts. This may be due to the shift of the net charge toward greater electronegativity because of the π-electron cloud in the benzene rings plus the carbonyl oxygen in BP (Fig. 1*B*). The BP-containing G-ladder was abbreviated to ≈8 nt compared with the normal G-ladder (>14 nt; Fig. 1*C*; compare lane 3 with lanes 2 and 4) because the bulky 5′-BP may sterically and/or electrostatically hinder the "filling-up" of a putative RNA-binding site or channel. This effect is unlikely to be due to a probable higher K_m for the

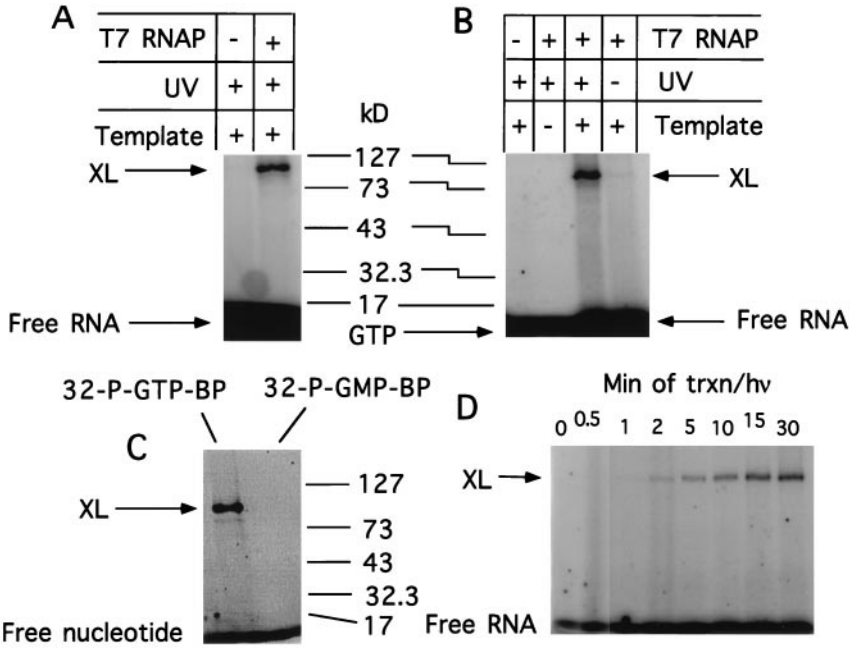

FIG. 2. Cross-linking of T7 RNAP to RNA. (*A*) Requirement of T7 RNAP for cross-linking. (*B*) Requirement of template for cross-linking. (*C*) Cross-linking only with labeled GTP-BP. (*D*) Time course of cross-linking. The numbers at the top are minutes after T7 RNAP addition and start of UV irradiation. In *A*, *B*, and *D*, reactions contained cold GTP-BP + cold GTP + [α-^{32}P]GTP, whereas *C* contained either [γ-^{32}P]GTP-BP + cold GTP or [α-^{32}P]GMP-BP + cold GTP. XL, cross-links.

GTP-BP, because 2- and 3-mer were robustly made (Fig. 1*C*, lane 3). Moreover, with [α-^{32}P]GTP-γ-BP alone, a similar short ladder was seen, suggesting that [α-^{32}P]GTP-γ-BP was indeed utilized to some extent by the enzyme (not shown).

Irradiation with ≈254 nm of UV light resulted in ^{32}P-labeled RNA-T7 RNAP cross-links (Fig. 2*A*). No cross-link was seen in the absence of T7 RNAP nor without DNA template nor with nonspecific protein (BSA), indicating that cross-linking occurred to the nascent transcript in the ternary complex (Fig. 2 *A* and *B*). That cross-linking was to the 5′ end of the RNA but not to the initiating nucleotide (in the active center) is apparent because only [γ-^{32}P]GTP-BP (Fig. 2*C*) or cold GTP-γ-BP plus [α-^{32}P]GTP (Fig. 2 *A*, *B*, and *D*) but not [^{32}P]GMP-BP (Fig. 2*C*) or cold GTP-γ-BP plus [γ-^{32}P]GTP gave cross-links (Fig. 2*C*). Cross-linking increased linearly during RNA synthesis (Fig. 2*D*). Based on the relative phosphor counts in the cross-links versus free RNA, about 2–5% of the RNA was cross-linked.

Cross-Linked T7 RNAP Is Active. First, T7 RNAP was cross-linked (i.e., 5′-binding site was "filled") to cold GTP-γ-BP and was then supplied with either [α-^{32}P]GTP or [α-^{32}P]CTP. With [α-^{32}P]GTP, poly(rG) was synthesized. This was seen in RNA that was cross-linked to T7 RNAP. The radiolabeled cross-linked T7 RNAP was first isolated from SDS gels by electroelution and then treated with 1% trifluoroacetic acid or 10 mM HCl (pH 2) at 37°C. This released the cross-linked RNA because the P-N bond in the BP-GTP is acid labile (see Fig. 1*B*). The released ^{32}P-labeled RNA was subsequently run on an 8 M urea-acrylamide denaturing gel and detected by autoradiography. With the addition of [α-^{32}P]CTP, a 5-mer (GGGCG) was synthesized. Fig. 3 *A* and *B*, respectively, show that with the addition of [α-^{32}P]GTP or [α-^{32}P]CTP the cross-linked 5′-BP-GTP RNA was extended, indicating that the active site is physically separate from the 5′ RNA-binding site or channel. The 5′ cross-linked site may be mobile enough to allow RNA extension while a 1- to 3-bp RNA-DNA hybrid was maintained during oligo(rG) synthesis. Alternatively, if the cross-linked site is rigid, the poly(G) chain may be flexible enough to be packed into the RNA channel to some extent. 5′-BP-RNA in a binary complex was also cross-linked (Fig. 3*B*, lane 1)—albeit with somewhat lower efficiency (three to four times) than in ternary complex—implying that the 5′ RNA-binding site is accessible in the absence of transcription.

Effect of T7 Lysozyme. T7 lysozyme is a specific inhibitor of T7 RNAP transcription that forms a 1:1 complex with T7 RNAP (28, 29). T7 lysozyme was proposed to act by inducing a conformational change in the polymerase (29). When T7 lysozyme was present before T7 RNAP addition, cross-linking was reduced, but not completely eliminated (Fig. 4). The decreased cross-linking coincided with reduced RNA synthesis. When lysozyme was added after transcription initiation, its effect on cross-linking was much weaker (not shown). Here, we may distinguish a direct competitive effect versus an indirect (allosteric) effect by lysozyme. A direct effect was ruled out because any lysozyme-RNA cross-link would have been seen between the T7 RNAP-RNA cross-link and the free RNA in the gel because T7 lysozyme is much smaller (17 kDa) than T7

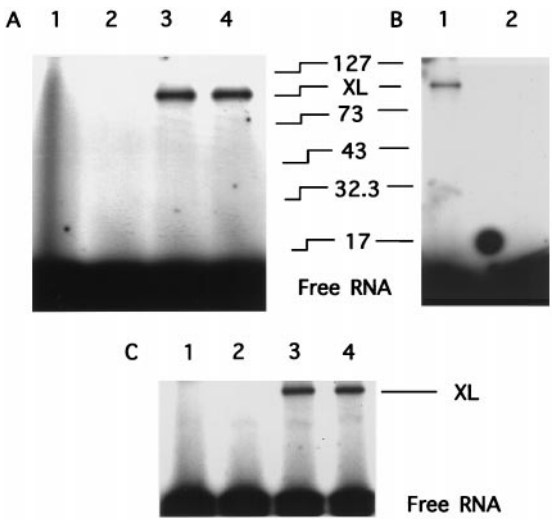

FIG. 3. Cross-linked T7 RNAP is active. (*A*) Lane 1, cold GTP + [α-^{32}P]GTP, no UV irradiation; lane 2, cold GTP + [α-^{32}P]GTP plus UV irradiation; lane 3, cold GTP-BP + [α-^{32}P]GTP + UV irradiation; lane 4, cold GTP-BP + UV irradiation + [α-^{32}P]GTP. (*B*) Cross-linking of the RNA in a binary complex. Isolated 5′-BP containing ^{32}P-labeled RNA was mixed with T7 RNAP in transcription buffer without NTPs and incubated for 15 min at 37°C. The reaction was then split in two halves. One was UV irradiated (lane 1) and the other was not irradiated (lane 2). Note that *A* (10% acrylamide gel) is slightly offset relative to *B* (12% acrylamide). The numbers in the middle show the relative positions of the protein markers in kDa. (*C*) Lane 1, cold GTP + [α-^{32}P]CTP, no UV irradiation; lane 2, cold GTP + [α-^{32}P]CTP plus UV irradiation; lane 3, cold GTP-BP + cold GTP + [α-^{32}P]CTP + UV irradiation; lane 4, cold GTP-BP + cold GTP + UV irradiation + [α-^{32}P]CTP.

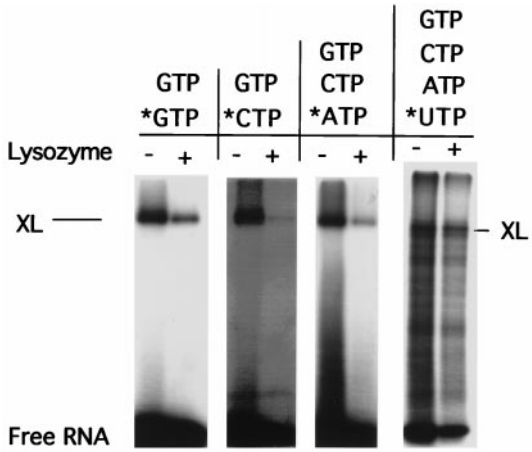

FIG. 4. Effect of T7 lysozyme on cross-linking. A transcription mixture was split in two halves. One received a 4-fold molar excess of T7 lysozyme over T7 RNAP. The other received a compensating volume of lysozyme storage buffer. Then T7 RNAP was added to both tubes and the reactions were incubated at 37°C for 10 min followed by UV irradiation. Asterisk indicates the [^{32}P]NTP present in the reactions. All reactions contained only cold GTP-BP.

RNAP (Fig. 4 and data not shown). Therefore, we suggest that T7 lysozyme may act indirectly by binding elsewhere and altering the conformation in such a way as to destabilize the RNA-binding site.

Ionic and Hydrophobic Interactions. We tested the effects of KCl (for probing ionic contacts) and acrylamide (for probing hydrophobic interactions) on the cross-linking (Fig. 5*A*). Both reagents reduced the cross-linking yields by about 4- to 5-fold, suggesting that ionic and hydrophobic interactions are involved in RNA binding. KCl was only slightly more effective than acrylamide (Fig. 5*B*).

N-Terminal Segment Contacts the 5′ End of the Nascent RNA. SDS/PAGE and mass spectrometry of proteolyzed cross-links were used to identify the RNA-binding site(s). First, to roughly locate the cross-linking site, we used a photochemical redox cleaving agent (thallium acetate) and a chemical cleaving agent (hydroxyl amine). Tl (III) plus 254 nm of UV irradiation generates hydroxyl radicals (30), which can cleave the peptide backbone in the solvent-exposed surfaces (much like OH radical footprinting). T7 RNAP is known to be especially susceptible to proteolytic cleavage near K179 or K180, resulting in a nicked polymerase (80 kDa + 20 kDa) (31–33). Cleavage with Tl (III) + UV irradiation gave a unique fragment of about 25.5 kDa (Fig. 6*A*). Since no ^{32}P-labeled ≈80,000 fragment was seen, we tentatively concluded that the 20-kDa fragment (20 kDa plus the mass of the attached RNA = ≈25.5 kDa) contained the ^{32}P-labeled RNA-binding site. Cleavage with NH_2OH (specific to Asn-Gly bonds; there are only two Asn-Gly bonds in T7 RNAP) gave fragments of ≈65.3 kDa and ≈32 kDa (Fig. 6*B*). These masses were closest to two fragments from amino acids 1–588 and 1–289, respectively, in the computer-generated proteolytic map of T7 RNAP. These data indicated that the cross-linking site was probably between residues 1–289, consistent with an earlier result with Tl (III) + UV irradiation. Partial digestion with clostripain (endoproteinase Arg-C; Fig. 6*C*) yielded the following fragments: 63.5 kDa, 52 kDa, 37.6 kDa, 29.6 kDa, 24.6 kDa, and 16.5 kDa. These determinations were based on calibration using marker proteins (Bio-Rad) within this size range and polypeptides from the cleavage of unmodified T7 RNAP with clostripain. The sizes of the radiolabeled fragments in the gel were deduced from a semilogarithmic plot of the mobility distance (cm) versus molecular masses (kDa) of all standard proteins and T7 RNAP polypeptides. Examination of all possible clostripain fragments of T7 RNAP, which also fitted the Tl (III) + UV irradiation and the NH_2OH cleavage data, revealed that the 63.5-kDa fragment was closest to the fragments containing either amino acids 144–720 or 144–722; the 52-kDa fragment was closest to the 85–557 fragment; the 37.6 was closest to either amino acids 144–478 or amino acids 53–386 or amino acids 53–389 or amino acids 51–386 fragments; the 29.6-kDa fragment was closest to either amino acids 32–298 or amino acids 144–412; the 24.6-kDa fragment was closest to either amino acids 85–307 or amino acids 1–215 or amino acids 144–357; and the 16.5-kDa fragment was closest to either amino acids 144–291 or amino acids 85–231. Since the common N-terminal end of the longest (63.3 kDa) and the shortest (16.5 kDa) fragments (and several others in between) in

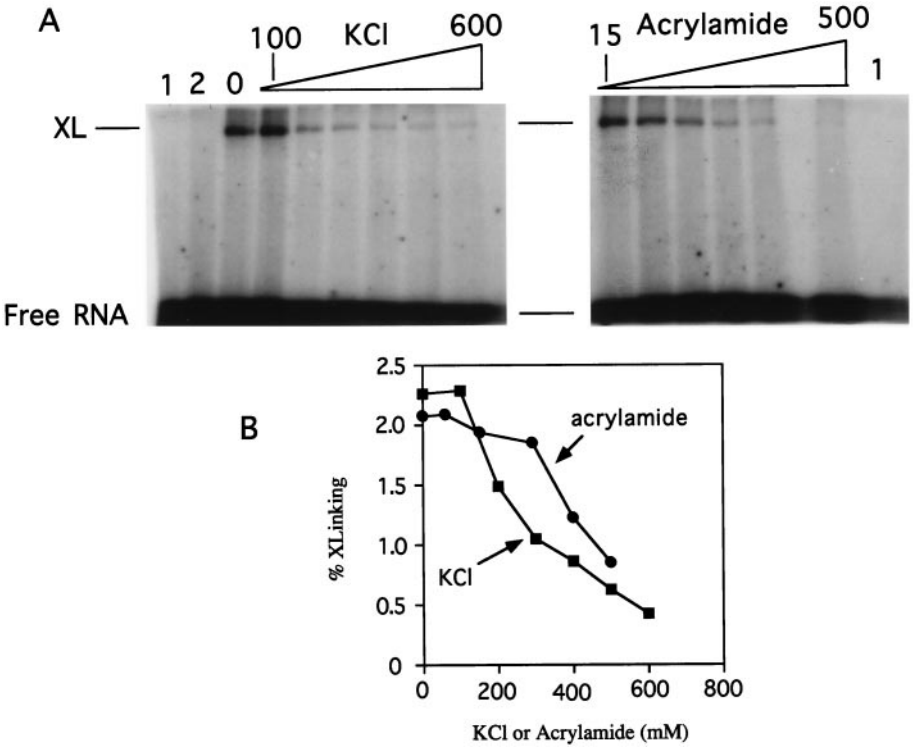

FIG. 5. Effect of KCl or acrylamide. Transcription was carried out for 10 min and then KCl or acrylamide was added before irradiation. (*A*) Lanes 1, no template; lane 2, no UV irradiation. Lanes that are not labeled contained (from left to right) either 200, 300, 400, and 500 mM KCl or 60, 150, 300, and 400 mM acrylamide. The graph (*B*) shows quantitation of the extent of cross-linking.

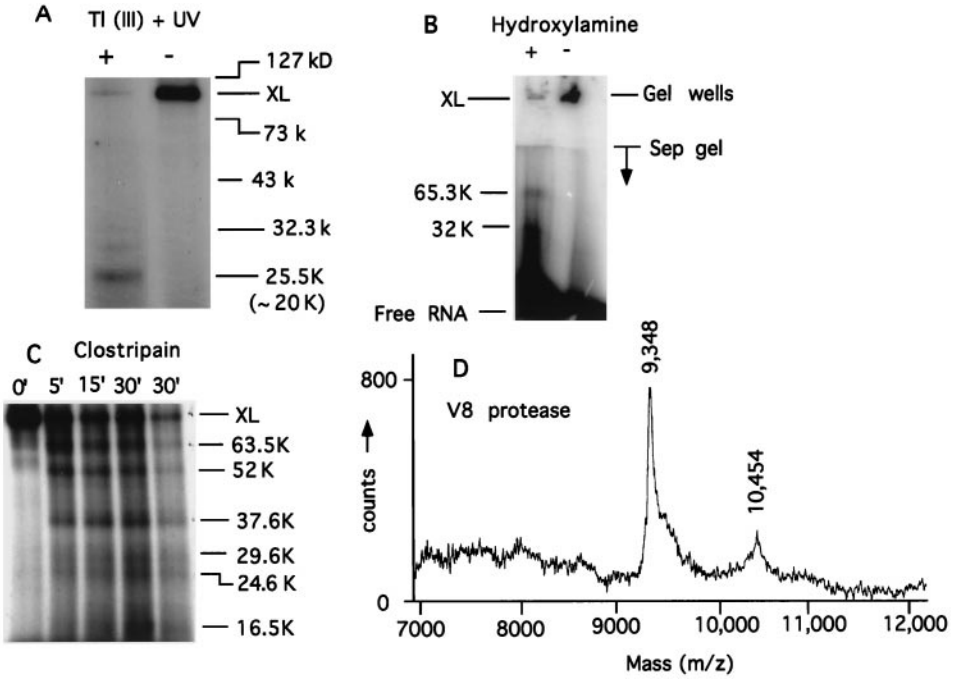

FIG. 6. Mapping the RNA-binding site. (*A*) Ten percent acrylamide-SDS Tris-Gly gel. Cleavage with thallium acetate + 254 nm of UV irradiation). Tl (III) OAc (100 mM) was added to cross-linked T7 RNAP and irradiated with 254 nm of UV light for 5 min. The size range of the markers is indicated while the actual (minus the mass of BP-RNA) mass of the proteolytic fragment is given in parentheses. (*B*) Cleavage with hydroxylamine. Ten percent acrylamide-SDS Tris-Gly gel. Purified cross-link was dissolved in 0.1 M K_2CO_3 (pH 10) and the cleavage was done as in ref. 37. To compensate for the lability of the P-N bond in the cross-link at pH 10, 37°C (where the NH_2 OH-cleavage reaction occurred), we had to use 6- to 8-fold cross-linked T7 RNAP in the + lane to see the cleaved fragments. In lane − the cross-link ran near the gel well. The actual (i.e., minus the mass of BP-RNA) mass of the proteolytic fragments is given. Protein markers (not shown) were the same as in *A*. (*C*) Ten percent acrylamide-SDS Tris-Tricine gel. Cleavage with clostripain was carried out as specified by the manufacturer (Promega) at cross-link to protease mass ratio of 100:1. The actual (minus the mass of BP-RNA) mass of the proteolytic fragments is given. To better visualize the bands, less material was loaded in the last lane (30′) compared with the others. Protein markers (not shown) were the same as in *A*. (*D*) Cleavage with V8 protease in gel pieces was carried out according to the manufacturer (Promega) at cross-link to protease mass ratio of 10:1. The plot shows a trace of the mass spectrogram in the relevant mass range.

this gel was amino acid 144, it was likely that the cross-link(s) was in the N terminus (approximately amino acids 144–289). This conclusion also agreed with the mass spectrometry data (see below).

Because mapping using SDS/PAGE is not very precise, we used mass spectrometry (MALDI-TOF). The cross-linked T7 RNAP was extensively digested with V8 protease (specific to E and D residues). The RNA-bound peptides were specifically purified by ion exchange chromatography (see *Methods* and refs. 10 and 11). Controls indicated that non-cross-linked peptides were not retained. The UV absorption spectrum (not shown) of the RNA cross-linked peptide showed a very broad transition around 260–274 nm, which is a signature of a protein-nucleic acid conjugate. To weigh the naked cross-linked peptide (+BP) using MALDI-TOF, the RNA was freed from the peptides by treatment with 1% trifluoroacetic acid at 37°C, which cleaved the acid-labile P-N bond that joins the γ-phosphorus in 5′-GTP to the BP in the RNA (Fig. 1*B*). The mass spectrum (Fig. 6*D*) showed a major peptide ($M + H^+ m/z$ = 9,348 − 197 (BP) = 9,151) that precisely corresponded to amino acids 88–168 (actual mass = 9,153) in the computer-generated V8 peptides data bank. The other closest fragments were 23 atomic mass units (amino acids 64–147) or 16 atomic mass units (amino acids 570–653) away from the determined mass. These fragments were not considered because: (*i*) our mass determinations are accurate to ±2 atomic mass units, and (*ii*) the data from amino acid composition supported our conclusion (see below). There was a minor peptide (m/z = 10, 454) which corresponded exactly with amino acids 1–91. There were no other major peptides between m/z 2–120 kDa (not shown). To confirm the identity of the major peptide, we determined the amino acid composition. Following is the determined composition (average no. of residues per mol; the composition from the sequence is in parentheses: A, 12.3 (12); C, 0 (1); Glx, 12.7 (Q + E = 12); F, 4.4 (5); G, 5.2 (3); H, 1.1 (1); I, 6.5 (6); K, 7.7 (8); L, 4.2 (4); Asx, 5.6 (N + D = 5); P, 2.1 (2); R, 5.8 (6); S, 2.1 (2); T, 6.8 (7); V, 4.9 (5); W, 0 (1); and Y, 1.3 (1). The determined amino acid composition shows a single H, consistent with the amino acid sequence of the 88–168 fragment. In addition, the other amino acids also closely match with a better than 95% confidence. Some amino acids (e.g., C and W) are difficult to quantitate because they are unstable or destroyed, and Gly is overrepresented in protein-nucleic acid conjugates (34). In summary, based on the above results, the major 5′ RNA-binding site was between amino acids 88 and 168, more specifically between amino acids 144 and 168, and a minor contact was between amino acids 1 and 91.

DISCUSSION

We have, for the first time, directly mapped nascent RNA-binding sites in T7 RNAP. Earlier work with a "nicked" polymerase and site-specific mutagenesis showed that the interactions with nascent RNA probably occurred in the N-terminal side of the polymerase (12, 14, 32). More recent site-directed mutagenesis indicated that E148 was involved in RNA binding (16). Our results are in agreement with these works. The positions of the RNA-binding sites are shown in the crystal structure of T7 RNAP (Fig. 7*A*) and are presumably partly in helix A and in helices H, G, E, and F. Some of these regions are ill-defined in the electron density map (10). The amino acids from 142–150 are highly conserved among mitochondrial and phage polymerases (35) and may have some homology to the region 2.4 of the *E. coli* RNAP σ factor. Mutagenesis showed the importance of this region in catalysis and promoter binding (36). Fig. 7*B* is a model of the emerging RNA in a ternary complex. Our earlier cross-linking work (18, 19) showed that the DNA template was in the cleft and that a part of the fingers subdomain contacted single-stranded DNA. Here, the template (black) is shown in the cleft (C) while the fingers (F) contact or maintain a transcription bubble (perhaps by contacting the nontemplate strand) while the thumb (T) wraps around the upstream DNA to stabilize the complex (Fig. 7*B*). The polymerase is shown riding on the DNA while the 5′ end of the emerging RNA (Fig. 7*B*, yellow) is held in a channel formed perhaps by the juxtaposition of helices (Fig. 7*A*; cyan, major 144–163; yellow, minor 7–91). For convenience, the cross-liking sites are termed as major and minor based on the

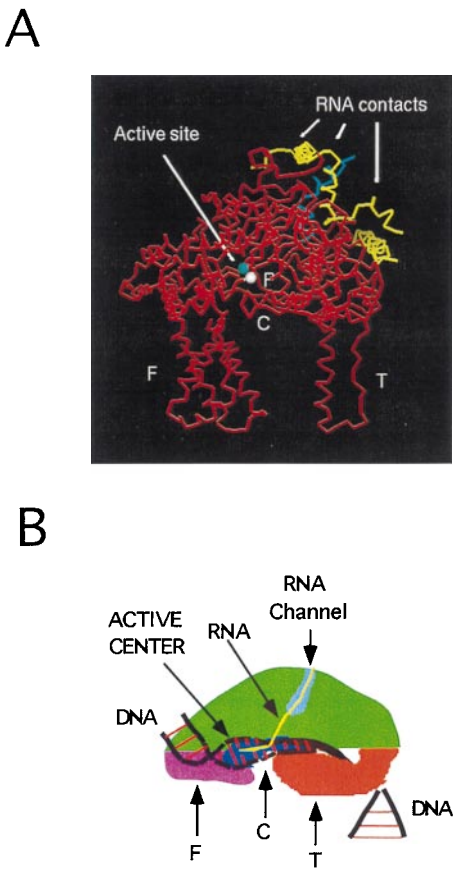

FIG. 7. (*A*) α-Carbon backbone (red) representation of the crystal structure of T7 RNAP (10). The coordinates in Brookhaven PDB file (4RNP) were displayed using an INSIGHTII program and a Silicon Graphics workstation. The major cross-link site is in cyan and the minor site is in yellow. The white and blue spheres are D537 and D812, respectively, in the active site. (*B*) A model of transcription complex based on the shape of T7 RNAP. C, cleft; P, palm is in blue; T, thumb is in orange; F, fingers are in violet; the DNA is in black while the base pairs are in red; RNA is in yellow. The proposed RNA channel is shown in light blue.

peak intensities in the mass spectrum. However, since MALDI-TOF is not quantitative, it is possible that the two sites may be equivalent in the same RNA channel. The distance between the colored helices varies between 16 and 22 Å. The channel may be partly "open" for interaction with factors that may directly or indirectly (through allosteric conformational changes) influence RNA–polymerase interaction (e.g., T7 lysozyme; Fig. 4). The channel may lie on the outer surface of the polymerase (Fig. 7*B*, blue-green). The sensitivity to OH radical [Tl (III) + UV] and solvent accessibility of very small molecules (Fig. 5) support this proposal. The 5′ end of the RNA in the channel may be stabilized by ionic and/or hydrophobic interactions (Fig. 5). Because the bulky BP group impedes the snaking of the growing RNA through this channel during the poly(rG) synthesis, these RNAs are shorter than normal G-ladders (Fig. 1*C*). Normally, longer RNAs may be more stably bound than the shorter ones. Once bound, the 5′-binding site(s) may be flexible to allow for the increasing RNA length (Fig. 3). The longer RNAs finding their way through the channel may trigger the conformational change that produces the abortive to processive polymerase transition. Our results do not exclude alternative interpretations. For example, there may be more than one RNA-binding site, as proposed for the *E. coli* RNAP (1), one or more in the abortive mode and others in the processive (or slippage) mode, where the ternary complexes are more stable. This may imply that the major site is the abortive RNA-binding site (since we cross-linked 5–8-mer RNAs) and the minor site is the other site.

It is unlikely that the site we have identified is involved in interactions with the RNA-DNA hybrid or the RNA 3′ end because only the RNAs containing 5′-BP-GTP are cross-linked (Fig. 2), and the cross-linked residue can be extended (Fig. 3). The distance from the active center (e.g., D537, D812; spheres in Fig. 7*A*) to the different residues in the highlighted helices in Fig. 7*A* is ≈32–50 Å, suggesting that different lengths of nascent RNA may be accommodated in the channel. The distance from the carbonyl of BP to the γ-phosphorus of GTP is about 7 Å. This added flexibility of the probe may influence the positions of the cross-links. We have not identified the exact amino acid that is cross-linked partly because of difficulties in obtaining sufficient quantities of cross-links. Nevertheless, the binding site has been localized to a very small T7 RNAP segment, which is in general agreement with other studies.

We thank Dr. B. T. Chait for suggestions/discussions regarding mass spectrometry, the protein/DNA technology center at The Rockefeller University for help with mass spectrometry and amino acid analysis, and Professor Joshua Lederberg for his general interest and support. This work was supported in part by a Hewlett-Packard Company grant.

1. Uptain, S. M., Kane, C. & Chamberlin, M. (1997) *Annu. Rev. Biochem.* **66,** 117–172.
2. Burley, S. K. & Roeder, R. G. (1996) *Annu. Rev. Biochem.* **65,** 769–799.
3. Gamper, H. B. & Hearst, J. E. (1982) *Cell* **29,** 81–90.
4. Yager, T. D. & von Hippel, P. H. (1991) *Biochemistry* **30,** 1097–1118.
5. Chamberlin, M. J. (1993) *Harvey Lect.* **88,** 1–21.
6. Landick, R. (1997) *Cell* **88,** 741–744.
7. Komissarova, N. & Kashlev, M., (1997) *J. Biol. Chem.* **272,** 15329–15338.
8. Nudler, E., Mustaev, A., Lukhtanov, E. & Goldfarb, A., (1997) *Cell* **89,** 33–41.
9. Chamberlin, M. J. & Ryan, T. (1982) *The Enzymes* **15,** 87–108.
10. Sousa, R., Chung, Y. J., Rose, J. P. & Wang, B.-C. (1993) *Nature (London)* **364,** 593–599.
11. Joyce, C. M. & Steitz, T. A. (1995) *J. Bacteriol.* **177,** 6321–6329.
12. Sousa, R., Patra, D. & Lafer, E. M. (1992) *J. Mol. Biol.* **224,** 319–334.
13. Sousa, R. & Padilla, R. (1995) *EMBO J.* **14,** 4609–4621.
14. Patra, D., Lafer, E. M. & Sousa, R. (1992) *J. Mol. Biol.* **224,** 307–318.
15. Macdonald, L. E., Durbin, R. K., Dunn, J. J. & McAllister, W. T. (1994) *J. Mol. Biol.* **238,** 145–158.
16. He, B., Rong, M., Durbin, R. & McAllister, W. T. (1997) *J. Mol. Biol.* **265,** 275–288.
17. Lyakhov, D. L., He, B., Zhang, X., Studier, F. W., Dunn, J. J. & McAllister, W. T. (1997) *J. Mol. Biol.* **269,** 28–40.
18. Sastry, S. S., Spielmann H. P, Hoang Q. S., Phillips, A. M., Sancar, A. & Hearst, J. E. (1993) *Biochemistry* **32,** 5526–5538.
19. Sastry, S. S. (1996) *Biochemistry* **35,** 13519–13530.
20. Grodberg, J. & Dunn, J. J. (1988) *J. Bacteriol.* **170,** 1245–1253.
21. Zawadzki, V. & Gross, H. J. (1991) *Nucleic Acids Res.* **19,** 1948.
22. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning. A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY).
23. Rajagopalan, K., Chavan, A. J., Haley, B. E. & Watt, D. S. (1993) *J. Biol. Chem.* **268,** 14230–14238.
24. Beavis, R. C. & Chait, B. T. (1996) *Methods Enzymol.* **270,** 519–551.
25. Beavis, R. C. & Chait, B. T. (1990) *Anal. Chem.* **62,** 1836–1840.
26. Prestwich, G. D., Dorman, G., Elliott, J. T., Marecak, D. M. & Chaudhary, A. (1997) *Photochem. Photobiol.* **65,** 222–234.
27. Dorman, G. & Prestwich, G. D. (1994) *Biochemistry* **33,** 3661–3673.
28. Moffatt, B. & Studier, W. F. (1987) *Cell* **49,** 221–227.
29. Zhang, X. & Studier, F. W. (1997) *J. Mol. Biol.* **269,** 10–27.
30. Roundhill, D. M. (1994) *Photochemistry and Photophysics of Metal Complexes* (Plenum, New York), p. 178.
31. Tabor, S. & Richardson, C. C. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 1074–1078.
32. Muller, D. K., Martin, C. T. & Coleman, J. E. (1988) *Biochemistry* **27,** 5763–5771.
33. Ikeda, R. A. & Richardson, C. C. (1987) *J. Biol. Chem.* **262,** 3790–3799.
34. Blackman, S. (1978) *Amino Acid Determination: Methods and Techniques* (Marcel Dekker, New York), pp. 8–35.
35. McAllister, W. T. (1993) *Cell. Mol. Biol.* **39,** 385–391.
36. Gross, L., Chen, W. J. & McAllister, W. T. (1992) *J. Mol. Biol.* **228,** 488–505.
37. Riva, M., Carles, C., Sentenac, A., Grachev, M. A., Mustaev, A. A. & Zaychikov, E. F. (1990) *J. Biol. Chem.* **265,** 16498–16503.