Studies of Contacts between T7 RNA Polymerase and Its Promoter Reveal Features in Common with Multisubunit RNA Polymerases[†]

Christophe Place,*[,‡,§] Jacqueline Oddos,[‡] Henri Buc,[‡] William T. McAllister,[§] and Malcolm Buckle[‡]

Unité de Physico-chimie des Macromolécules Biologiques, CNRS:URA 1773, Institut Pasteur, 25, rue du Dr Roux, 75724 Paris Cedex 15, France, and Department of Microbiology and Immunology, Morse Institute for Molecular Genetics, State University of New York Health Science Center at Brooklyn, 450 Clarkson Avenue, Brooklyn, New York 11203-2098

Received November 11, 1998; Revised Manuscript Received January 21, 1999

ABSTRACT: We have used UV-laser mediated cross-linking, DNase I footprinting and $KMnO_4$ reactivity to probe the interaction between T7 RNA polymerase (RNAP) and a consensus promoter during the early stages of transcription. In a binary complex formed in the absence of substrate on a supercoiled plasmid, direct contacts were observed on the template (T) strand at positions −17, −5, and +3 and on the nontemplate (NT) strand at position −8. These contacts lie within the DNase I cleavage footprint from positions −21 to +11 on the T strand and from positions −17 to +16 on the NT strand and straddle sites of enhanced reactivity of thymines to $KMnO_4$ at position −3 on the T strand and position −2 on the NT strand. Use of supercoiled plasmid templates has allowed the mapping of contacts in the initiation region of the promoter in the binary complex for the first time. Upon addition of GTP, T7 RNAP enters a reiterative mode of synthesis, producing a ladder of poly(G) products. Under these conditions the downstream contact on the T strand switched from position +3 to +4 and +5 while the contact at position −17 was maintained. Under conditions in which the synthesis of transcription products is limited to 6−7 nucleotides, only the contact at position −17 on the T strand was preserved. A comparison of these results with the interaction of *Escherichia coli* RNA polymerase at the *lac* promoter reveals strong similarities in the manner in which these polymerases recognize their promoters.

Although the DNA-dependent RNA polymerase (RNAP)[1] encoded by bacteriophage T7 consists of a single subunit, it is able to carry out all of the steps in the transcription cycle as the more complex multisubunit RNAPs found in prokaryotic and eukaryotic cells (for review see ref *1*). Moreover, many of the events in the transcription cycle are remarkably similar for these two classes of enzyme. In the initial phase, both enzymes must recognize and bind to the promoter and melt open the duplex DNA template at the initiation site. As for the bacterial RNAP, transcription by T7 RNAP involves two separate phases. During the first phase, the enzyme forms an initiation complex that in the presence of NTPs undergoes repeated cycles of abortive transcription in which short RNA products are continuously synthesized and released. A transition occurs after the synthesis of 7−12 nucleotides (nt) of nascent RNA, resulting in the formation of a stable, highly processive elongation complex.

Due to its structural simplicity and the availability of crystallographic data (*2*, *3*), T7 RNAP represents an attractive subject for structure−function studies. A key point of interest is how the polymerase interacts with its 17 bp promoter during recognition and initiation.

A variety of methods have been used to probe T7 RNAP−DNA interactions during promoter binding on linear templates. Due to a low affinity, previous investigators were unable to obtain a DNase I footprint of T7 RNAP on its promoter in the absence of substrate (*4*, *5*). However, footprinting reactions carried out with methidiumpropyl-Fe(II) under more favorable ionic conditions showed protection of the promoter region from residues −21 to −3 (*6*−*8*). Hydroxyl radical footprinting with Fe(II)−EDTA revealed a well-defined pattern of protection predominantly on one face of the DNA helix (*9*). Curiously, the initiation region of the promoter was not protected in the binary complex in these studies, even though this region of the promoter must presumably be melted open prior to or during initiation. In the presence of the first nucleotides, the stability of the RNAP−promoter complex is increased and a DNase I footprint could be detected (*4*, *5*, *10*). Under these conditions the MPE−EDTA−Fe^{2+} footprint extended in the downstream direction, while upstream contacts were maintained (*7*). These studies indicated that T7 RNAP maintains a continued association with the upstream region of the promoter during abortive initiation and that a conformational change in the ternary complex occurs during the synthesis of small RNAs. Alterations in the sensitivity of the RNAP to protease digestion after this transition support this concept (*11*).

In the binary complex, the open and closed forms are present in rapid equilibrium, and on linear templates the

[†] Supported by a grant from the Ligue National contre le Cancer (to C.P.), by an ACC-SV5 grant from the Ministère de la Recherche et de l'Enseignement Supérieur (to H.B.), and by NIH Grant GM38147 (to W.T.M.).

* Corresponding author. Telephone: (33) 01 45 68 85 20. Fax: (33) 01 40 61 30 60. E-mail: cplace@pasteur.fr.

[‡] Unité de Physico-chimie des Macromolécules Biologiques.

[§] Department of Microbiology and Immunology.

[1] Abbreviations: DNase I, deoxyribonuclease I; EDTA, ethylenediaminetetraacetic acid; RNAP, RNA polymerase; NTPs, nucleoside triphosphates; DTT, dithiothreitol; Tris, tris(hydroxymethyl)aminomethane; Nd:YAG, $Y_3Al_5O_{12}$ garnet doped with Nd^{3+}.

10.1021/bi982689e CCC: $18.00
Published on Web 03/30/1999

closed form is thermodynamically favored (*12*, *13*). This may account for failure on the part of previous investigators to observe RNAP−promoter contacts in the initiation region on linear templates. The rapid formation of an unstable open complex does not usually correspond to a rate limiting step, which under optimal conditions occurs later in the pathway (*12*, *13*). The stability of the binary complex is affected by the topological state of the DNA and is greatly enhanced on templates in which collapse of the melted region is inhibited (e.g., on supercoiled templates or on partially single stranded promoters) (*12*−*16*). Taking advantage of these observations, we here characterize binding and initiation complexes formed by T7 RNAP on supercoiled plasmid templates.

In our studies, we have applied techniques that allow not only identification but also quantification of such interactions. The first method, UV laser cross-linking of the polymerase to the DNA, probes intimate contacts between the DNA and the polymerase (within van der Waals radii) (*17*−*20*). Previous photochemical cross-linking studies using psoralen to map contacts between T7 RNAP and single-stranded DNA or a T7 promoter were general in nature and lacked precision due to the long arm of psoralen (8 Å) and the occurrence of multiple reactive positions in the promoter (*21*, *22*). We also used DNase I protection and differential reactivity to potassium permanganate ($KMnO_4$). The former reports the occupancy of the DNA by the polymerase, allowing easy titration of a given promoter site. The latter detects exposure of the 5′−6′ double bonds of thymine residues that may result from helical strain, for example, as a result of partial strand separation.

These methods have revealed contacts of the RNAP with the initiation region during promoter binding and changes in the nature of these contacts during the early stages of initiation. A comparison of these results with those obtained for *Escherichia coli* RNAP using similar methods reveals important similarities in the manner in which these quite different enzymes carry out the early stages of promoter binding and initiation.

EXPERIMENTAL PROCEDURES

T7 RNAP Preparation. T7 RNA polymerase was overexpressed from cultures of *E. coli* BL21 carrying plasmids pDL19 or pBH161 (*23*). These plasmids encode RNAPs with a hexa-His tag at the N-terminus, allowing a single-step purification by absorption on a Ni^{2+} column and elution in Tris-HCl (20 mM), pH 7.9, NaCl (50 mM), imidazole (200 mM), and Tween 20 (0.05%) (*23*). Polymerase was stored at −20 °C in 50% glycerol. The concentration of the purified enzyme was determined using a molar extinction coefficient at 280 nm of $1.4 \times 10^5\ M^{-1}\ cm^{-1}$ (*24*). The specific activity was generally of the order of 310 000 units/mg in transcription buffer containing Tris-HCl (40 mM), pH 8, NaCl (10 mM), $MgCl_2$ (6 mM), spermidine (2 mM), DTT (4 mM), and Tween 20 (0.05%), which is close to the theoretical maximum for a preparation in which all enzyme molecules are active (*23*, *25*).

Plasmid Preparation. Plasmid DNA pPK3 (Karasavas and McAllister, in preparation) was prepared by alkaline lysis from *E. coli* HB101 and purified on an anion-exchange cartridge (Nucleobond). Where appropriate, pPK3 was linearized by digestion at an unique *Afl*III site with the corresponding nuclease (New England Biolabs).

Laser Irradiation. Plasmid DNA pPK3 (5 nM) was incubated with 0−500 nM T7 RNA polymerase in transcription buffer containing Tris-HCl (20 mM), pH 7.9, NaCl (10 mM), $MgCl_2$ (5 mM), and Tween 20 (0.05%) for 15 min at 37 °C. When present, nucleoside triphosphates (NTPs) were added to a final concentration of 0.4 mM each. Samples to be irradiated (10 μL) were placed in a thermostatically controlled Eppendorf tube. High-intensity laser light with a fundamental wavelength of 1064 nm was generated by a Nd:YAG laser. Two frequency doubling crystals generated 512 and 266 nm polarized light. The latter wavelength was selected using a series of dichroic mirrors and directed onto the sample. Irradiation was performed with a single 5 ns energy pulse of 20 mJ ($3.5 \times 10^{10}\ W \cdot m^{-2}$).

Primer Extension. Five microliters of the irradiated sample and 0.4 μL of 17 or 19 nt primers (500 nM) were heated to 95 °C for 5 min in a primer extension buffer containing Tris-HCl (50 mM), pH 8.0, $MgCl_2$ (10 mM), and DTT (1 mM) and abruptly cooled on ice for 2 min. The primers (5′-AAATAGGCTTATCACGAGG-3′ for the T strand and 5′-GCGTTGGCCGATTCATT-3′ for the NT strand) hybridized to the plasmid 58 bases upstream and 63 bases downstream, respectively, from the starting site of transcription.

Extension was performed after addition of 10^{-3} unit of Klenow fragment (1 unit/mL, Amersham Pharmacia Biotech) and a complete set of deoxyribonucleoside triphosphates (4 mM; Amersham Pharmacia Biotech) at 48 °C for 15 min. The Klenow fragment was inactivated by heating for a further 10 min at 65 °C, and the DNA was precipitated with ethanol, washed and dried; the pellet was resuspended in 5 μL of loading buffer: formamide (80%), EDTA (40 mM), xylene cyanol (0.1%), and bromophenol (0.1%). Samples were separated on 8% polyacrylamide denaturing gels containing urea (8 M) and visualized by autoradiography. Quantification was carried out by Phosphorimager densitometry (Molecular Dynamics).

In the presence of nucleoside triphosphates, extraneous bands appear which obscure the extension pattern of the template strand at the start site of transcription. This artifact is due to the hybridization of the abundantly synthesized abortive products and may be avoided by treatment with RNase H (Boehringer Mannheim) prior to primer extension. The synthesis of abundant transcription products in the presence of all four ribonucleoside triphosphates provided high concentrations of RNA strands which in turn served as templates for primer extension, thus producing a high density of labeled bands downstream of the start of transcription. The corresponding primer extension pattern were consequently impossible to interpret on the NT stand.

DNase I Footprinting. T7 RNAP (100 nM) and the plasmid pPK3 (5 nM) were incubated in footprinting buffer containing Tris-HCl (40 mM), pH 8.0, NaCl (10 mM), $MgCl_2$ (6 mM), and Tween 20 (0.05%) with or without nucleoside triphosphates (0.5 mM) for 15 min at 37 °C prior to the addition of DNase I (0.5 mg/mL). The reaction was allowed to proceed for 20 s and stopped with an equal volume of phenol and EDTA to a final concentration of 2.5 mM. The aqueous phase was isolated and the DNA was precipitated with ethanol in the presence of glycogen (20 μg) and sodium acetate (0.3 M). The pellet was washed with 70% ethanol and resuspended in buffer. Primer extension was carried out as described above, and samples were loaded

onto an 8% acrylamide denaturing gel and analyzed as above.

$KMnO_4$ Footprinting. T7 RNAP (100 nM) and the plasmid pPK3 (5 nM) were incubated for 15 min at 37 °C in footprinting buffer with or without nucleoside triphosphates (0.5 mM) prior to the addition of $KMnO_4$ (8 mM). When used, the chain terminator 3′-deoxy-CTP (0.05 mM) was added instead of CTP. The reaction was allowed to proceed for 1 min at 37 °C and was quenched by the addition of $^1/_{10}$ volume of β-mercaptoethanol (14.7 M; Sigma). The sample was filtered on a Sephadex G50 column (Boehringer Mannheim), dried, and redissolved in the primer extension buffer. Primer extension, gel electrophoresis, and analysis were performed as described above.

Transcription. Transcription was carried out in footprinting buffer containing T7 RNAP (100 nM) and plasmid pPK3 (5 nM) at 37 °C with different mixtures of nucleoside triphosphates (0.5 mM) containing 2 mCi of [α-^{32}P]GTP (66 nM) and was stopped after 15 min by the addition of an equal volume of loading buffer: formamide (80%), EDTA (40 mM), xylene cyanol (0.1%), and bromophenol (0.1%).

RESULTS

Photoreactivity of Polymerase−Promoter Complexes. In the present study, a plasmid carrying a T7 promoter was incubated with T7 RNA polymerase under defined conditions and exposed to a 5 ns pulse of high-energy radiation (266 nm, 20 mJ containing 10^{16} photons per pulse). The photoreactivity of bases leads to intramolecular alterations of the DNA (for example to the formation of pyrimidine dimers) and the formation of covalent bonds between the protein and excited bases (*20*, *26*, *27*). However, due to rapid decay of the excited state, cross-linking occurs only if the residues in the protein are within distances corresponding to van der Waals radii of the activated base. To identify positions on the DNA at which cross-linking has occurred, DNA−protein complexes were denatured under conditions that retain the cross-link, and each strand of the DNA was used as a template for primer extension by the Klenow fragment of DNA polymerase. Previous results indicate that the primers are extended to a position immediately before the site of the cross-link (*26*). To facilitate discussion, we identify the resulting bands by the inferred position of the modified base, rather than by the position at which DNA polymerase terminates.

In the experiments that follow, two types of signals may be observed. In the first, cross-linking of the protein results in the appearance of a new band that is not observed in the absence of RNAP. These bands are attributed to direct contacts between the enzyme and the DNA, resulting in the formation of a covalent cross-link which blocks subsequent primer extension. In the second case, the presence of the RNAP leads to an increase or decrease in the inherent photoreactivity of a base (already visible in the absence of the RNAP). These changes may result either from the formation of a cross-link (resulting in an increase in signal intensity) or from changes in photoreactivity due to a close approach to the base (resulting in either an increase or a decrease in signal intensity).

UV laser cross-linking was carried out using the supercoiled plasmid pPK3 as a template. The consensus promoter in this plasmid has been placed in a sequence context that

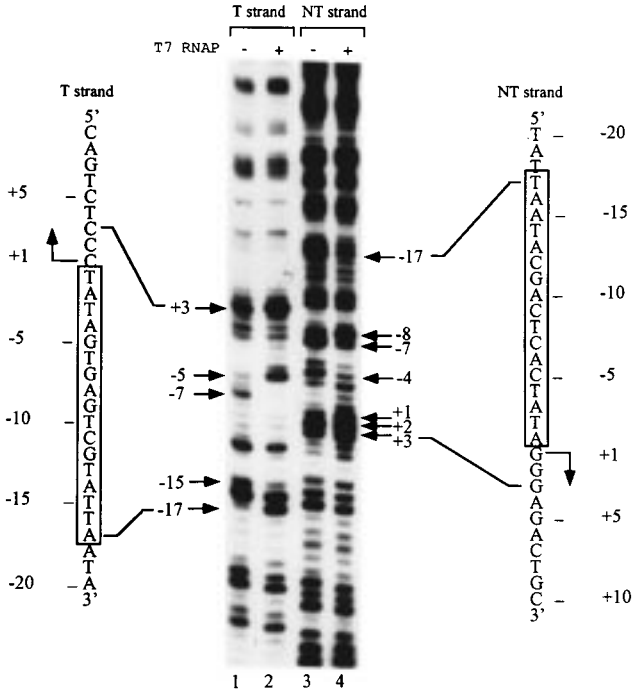

FIGURE 1: Primer extension pattern on irradiated DNA in plasmids and in complexes with RNA polymerase. Complexes formed between T7 RNA polymerase (100 nM) and supercoiled templates (5 nM) were irradiated using a UV laser as described in Experimental Procedures. Primer extension was carried out, again as described in Experimental Procedures, and the products were separated by electrophoresis on 8% (w/v) polyacrylamide denaturing gels and identified by comparison to a sequencing reaction using a dideoxy chain terminator (data not shown). The numbers indicate the modified base that results in termination of primer extension. Lanes: 1 and 2, extension of primer annealed to the template strand; 3 and 4, extension of primer annealed to the nontemplate strand; 1 and 3, irradiated DNA; 2 and 4, irradiated DNA in the presence of T7 RNA polymerase. Bases are numbered according to their position with respect to the start site (broken arrow). The corresponding sequence is included adjacent to the gel with the consensus promoter boxed.

allows transcription to be extended to defined positions in the presence of appropriate mixtures of substrates, and the stability of these complexes has been determined (Karasavas and McAllister, in preparation). Binding of T7 RNAP in the absence of substrates resulted in an increased intensity of the bands at A−17, G−5, and C+3 and a decreased reactivity at G−7 and T−15 on the template strand (Figure 1). The intrinsic photoreactivity of the bands at positions −17 and −5 (in the absence of RNAP) is very low. Their enhancement is thus likely to result from direct contacts between the enzyme and the DNA. The decreased reactivity at psotions −15 and −7 is attributed to the close approach of the polymerase. The spacing of the contacts at positions −17/−15 and −7/−5 suggests that they may be on the same face of the DNA helix. The increased intensity of the band at position +3 is also likely to result from a direct contact at that position (see below).

The observed changes on the NT strand occurred at bases already possessing an intrinsic photoreactivity. Pyrimidine dimer formation took place between T−18 and T−17, T−8 and C−7, and C−5 and T−4. Those signals decreased in the presence of RNAP while the reactivity resulting from dimer formation between C−9 and T−8 increased. There was also a general increase in intensity at positions +1, +2,

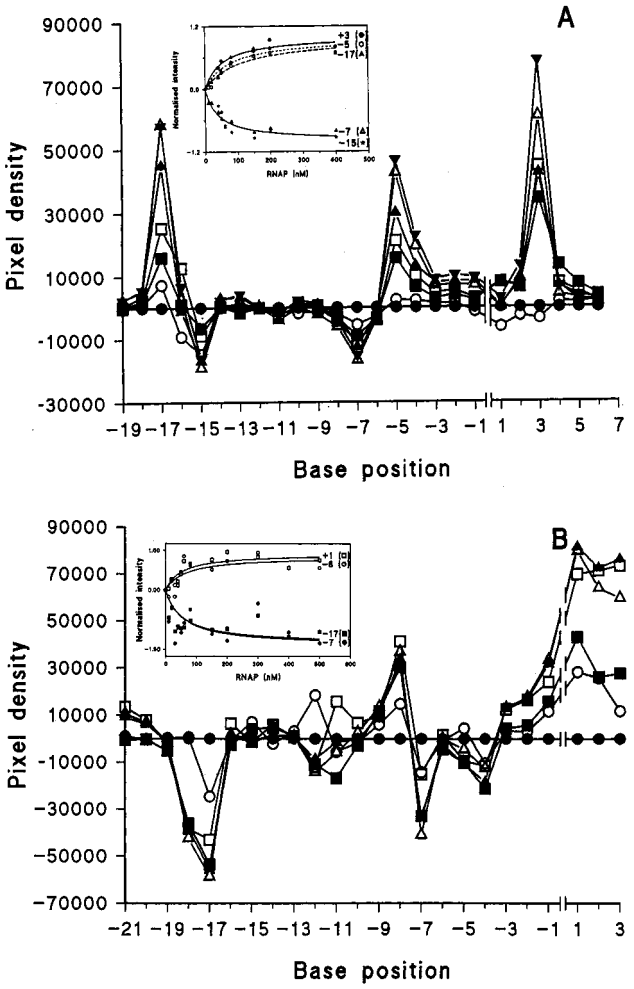

FIGURE 2: Phosphorimager quantification of the intensity of the bands of the type shown in Figure 1 in the presence of increasing amounts of T7 RNA polymerase. The relative pixel density of certain bands remains unaffected by the presence of RNA polymerase. Chosen reference bands were at positions −13 on the template strand and +10 on the nontemplate strand. The pixel density of this position (e.g., position −13) for all lanes was compared to that of the naked DNA, and the resulting ratio was used to multiply other positions in each lane. Pixel density on the ordinates refers to the difference between these values at any given RNA polymerase concentration and that for the DNA alone. (A) Template strand: (●) 0 nM RNA polymerase; (○) 10 nM RNA polymerase; (■) 40 nM RNA polymerase; (□) 50 nM RNA polymerase; (▲) 80 nM RNA polymerase; (△) 150 nM RNA polymerase; (▼) 200 nM RNA polymerase. (B) Nontemplate strand: (●) 0 nM RNA polymerase; (○) 20 nM RNA polymerase; (■) 50 nM RNA polymerase; (□) 80 nM RNA polymerase; (▲) 150 nM RNA polymerase; (△) 500 nM RNA polymerase. The inserts show the variation of the normalized intensity of selected bands with respect to increasing concentrations of RNA polymerase. Normalization was carried out by relating the given pixel density to the fitted asymptote assuming a hyperbolic relationship between intensity and RNA polymerase concentration.

and +3; the bases at these positions are guanines, which are known to have an inherent photoreactivity (*28*). A small increase in photoreactivity was also observed in the bases flanking this region.

The intensity of all these signals on both strands was dependent on RNA polymerase concentration over the range 0−500 nM (Figure 2). The data of Figure 2 were quantified using Phosphorimager analysis (Molecular Dynamics), and the relative pixel density was plotted in the inserts as a function of increasing RNA polymerase concentration. For all the positions shown, an apparent K_d of 50 ± 10 nM was estimated. Taken together, the results of the UV irradiation experiments suggest that the major van der Waals contacts made by the polymerase occur on the template strand at bases −17, −5, and +3. The increased signals at T−8 and G+1, G+2, and G+3 on the nontemplate strand may also represent contacts at these positions.

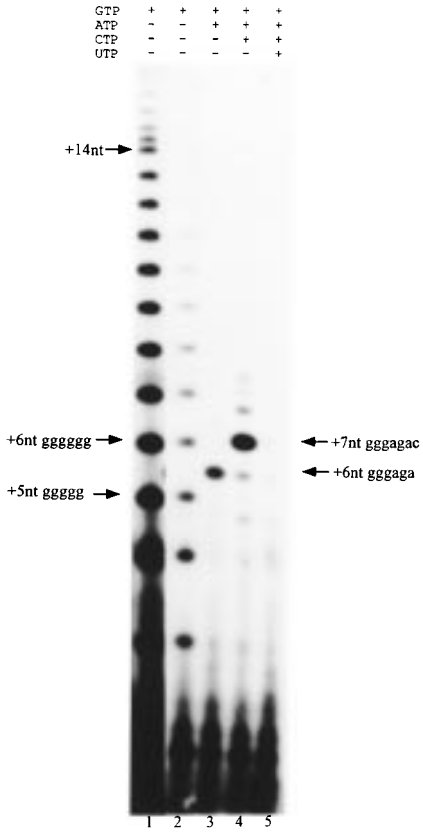

FIGURE 3: Synthesis of short RNAs by T7 RNAP at a consensus promoter in a supercoiled plasmid under limiting conditions of NTPs. The sequence at this promoter starts with GGGAGACU. The products were labeled with [α-^{32}P]GTP and resolved by electrophoresis on 20% (w/v) polyacrylamide denaturing gels. Lanes: 1 and 2, transcription in the presence of GTP alone (lane 1 is an overexposed view of lane 2); 3, GTP and ATP; 4, GTP, ATP, and CTP; 5, all four NTPs. Note that under these conditions high molecular mass major products do not enter the gel (not shown). The lengths of the transcripts and their base composition are indicated.

Photoreactivity of Transcribing Complexes. The sequence that lies just downstream from the +1 start site in pPK3 (GGGAGACT...) allows transcription to defined positions in the presence of limiting mixtures of ribonucleoside triphosphates. In the presence of GTP alone, the polymerase may extend to position +3, but due to slippage of the nascent RNA on the three C residues in the T strand from positions +1 to +3, the enzyme abundantly produces a ladder of poly(rG) products 2−14 nt in length (*29*) (Figure 3). In the presence of GTP and ATP the major product is the expected 6 nt RNA, and in the presence of GTP, ATP, and CTP the major product is 7 nt. (The faster migration of the latter products relative to the 6 and 7 nt poly(rG) products is presumably due to the influence of base composition on migration in this gel system (*30*)].

In the presence of GTP alone, the position −17 contact on the T strand observed in the binary complex is maintained,

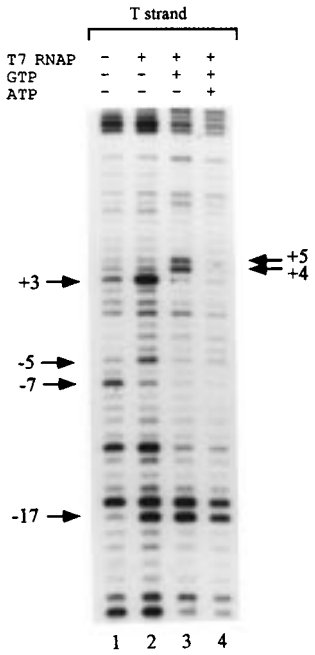

FIGURE 4: Primer extension pattern on the template strand of irradiated plasmid DNA alone or in complexes with RNA polymerase under limiting conditions of NTPs. Lanes 1, DNA alone; 2, DNA and RNA polymerase; 3, DNA, RNA polymerase, and GTP (0.4 mM); 4, DNA, RNA polymerase, GTP (0.4 mM), and ATP (0.4 mM). Primer extension products were resolved by electrophoresis on 8% (w/v) polyacrylamide denaturing gels. Bands of interest are marked by arrows and numbered with respect to the start site of transcription.

but the contacts at positions −5 and +3 are lost and two new signals appear at T+4 and C+5 (Figure 4). A generalized decrease in intensity for most other bands is also observed under these conditions due to absorption of the laser light by free nucleotides. On the NT strand there were no observed changes in photoreactivity in the presence of GTP (data not shown).

Upon addition of GTP and ATP, the position −17 contact on the T strand was maintained, but all contacts in the initiation region from positions +3 to +5 were lost (Figure 4). The addition of CTP did not alter the pattern (data not shown). No changes were observed on the NT strand (data not shown). Thus, a distinct upstream promoter contact is maintained during the early stages of abortive initiation, in agreement with prior observations using hydroxyl radical footprinting (*7*, *9*).

Probing of Polymerase−Promoter Contacts by DNase I Footprinting. Previous attempts to perform DNase footprinting of T7 RNAP on linear DNA in the absence of substrate were unsuccessful, possibly due to the weak binding of polymerase under these conditions (*4*, *12*, *13*, *16*). However, the use of a supercoiled DNA template enhances the stability of RNAP−promoter complexes (*14*). To visualize potential DNase I-induced nicks on a supercoiled template, we again used primer extension, as DNA polymerase stops elongation at a position that is opposite to the 5′-nucleotide at the cleavage site.

We identify the DNase I cutting positions by indicating the base which is 5′ to the cleaved phosphodiester bond.

In the absence of substrate, we observed protection from positions −21 to +11 on the T strand and enhanced cleavage at position −10 (Figure 5).

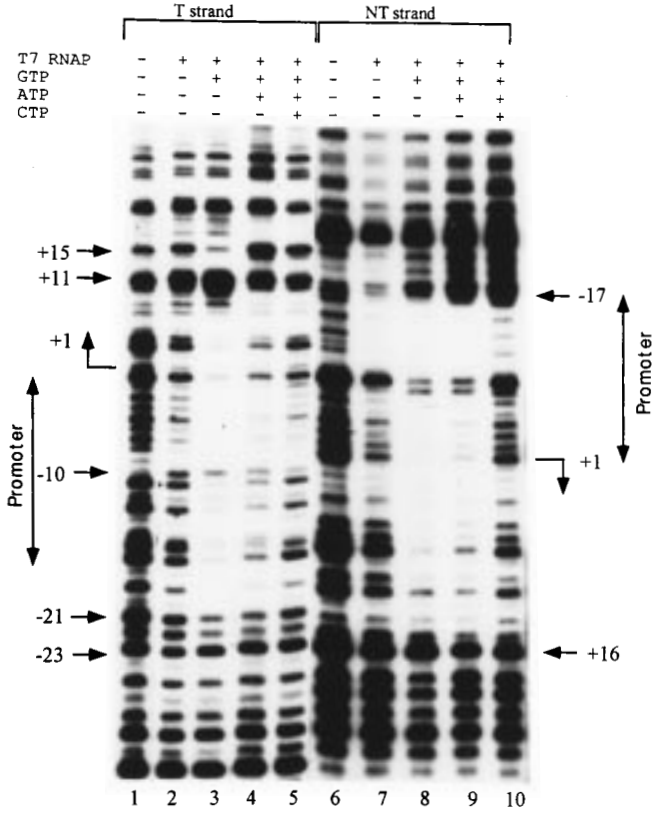

FIGURE 5: DNase I footprinting experiments on supercoiled DNA in the presence of T7 RNA polymerase and initiating nucleotides. Free DNA and DNA in complexes with RNA polymerase were digested by DNase I as described in Experimental Procedures and cleavage sites were revealed by primer extension. The patterns of digestion of the template strand are shown in lanes 1−5 and the nontemplate strand in lanes 6−10. Lanes: 1 and 6, DNA alone; 2 and 7, DNA and T7 RNA polymerase; 3 and 8, DNA, T7 RNA polymerase and GTP; 4 and 9, DNA, T7 RNA polymerase, GTP, and ATP; 5 and 10, DNA, T7 RNA polymerase, GTP, ATP, and CTP. The double arrow indicates the position of the promoter and the broken arrow the position and direction of the starting site. Important reactive bands are numbered according to the position of the corresponding bases with respect to the starting site.

In agreement with previous observations, a stronger footprint was obtained in the presence of GTP (Figure 5). Under these conditions, however, the position +15 band on the T strand vanished while cleavage at position +11 was considerably enhanced, and upstream protection extended to position −23. When GTP and ATP were added, the footprint from −23 to +11 was maintained. However, the intensity of the +11 band decreased and cleavage at position +15 was restored (Figure 5). In the presence of GTP, ATP, and CTP, protection was less pronounced on the promoter, but the overall pattern of digestion was similar to that seen for GTP + ATP. Protection of the promoter was abolished when a complete mixture of nucleotides was added, consistent with polymerase moving away from the promoter under these conditions.

On the NT strand, in the absence of substrate, protection was centered on the starting site and covered the DNA from positions −17 to +16. The weak intensity of the bands upstream of position −17 is due to more aggressive digestion of this particular sample with DNase I and was not observed in other experiments (data not shown). Protection was stronger when GTP or GTP + ATP were added (Figure 5). Less protection of the promoter region was observed in the

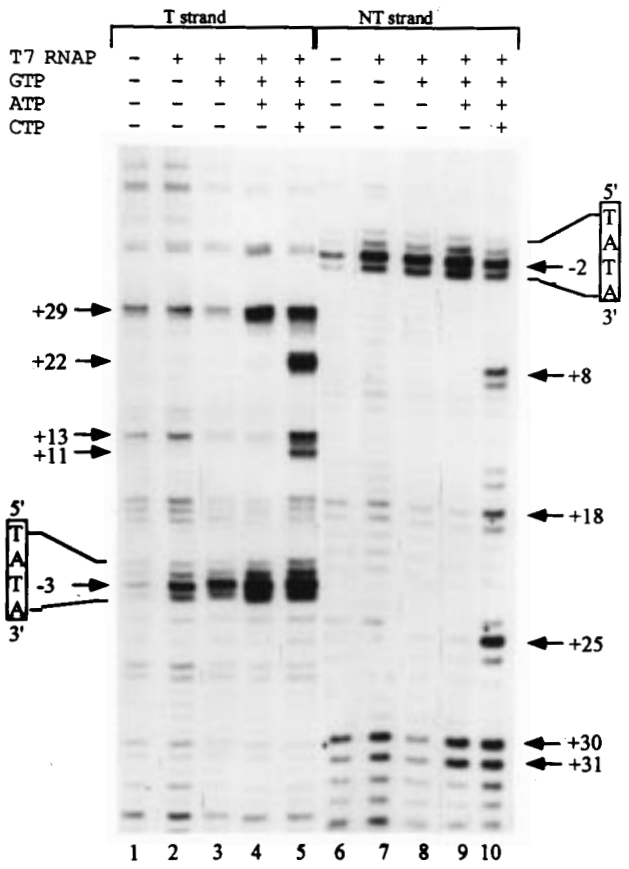

FIGURE 6: $KMnO_4$ attack on supercoiled DNA in the presence of T7 RNA polymerase and initiating nucleotides. Free DNA and DNA in complexes with RNA polymerase were exposed to $KMnO_4$ as described in Experimental Procedures, and cleavage sites were revealed by primer extension. Lanes: 1−5, template strand; 6−10, nontemplate strand; 1 and 6, DNA alone; 2 and 7, DNA and T7 RNA polymerase; 3 and 8, DNA, T7 RNA polymerase, and GTP; 4 and 9, DNA, T7 RNA polymerase, GTP, and ATP; 5 and 10, DNA, T7 RNA polymerase, GTP, ATP, and CTP. $KMnO_4$ hyperreactive sites are marked by arrows and numbered according to their position relative to the start site of transcription. The TATA box region is indicated next to the gel.

presence of GTP, ATP, and CTP. The intensity of the bands from positions −17 to −25 was considerably and reproductively increased in the presence of GTP + ATP or GTP, ATP, and CTP. In the presence of all four nucleotides, the pattern was obscured downstream of the starting site (data not shown) due to an artifact of primer extension (see Experimental Procedures).

Probing of T7 RNAP−Promoter Complexes with Potassium Permanganate. $KMnO_4$ has been used to probe single-stranded regions of duplex DNA and/or regions in which access to the 5′−6′ double bond in the pyrimidine ring of thymine is in some way enhanced (*31*). Primer extension allows visualization of these modifications, as Klenow terminates at the base either opposite the modified thymine or one base prior to the modification (*31*). Incubation of supercoiled templates with $KMnO_4$ without RNAP showed residual reactivity unique to thymine residues (Figure 6). In the absence of substrate, RNAP enhanced modification within the promoter region at positions −3 on the T strand and −2 on the NT strand. This pattern was observed in the presence of GTP, GTP, and ATP or GTP, ATP, and CTP (Figure 6) but disappeared in the presence of all four ribonucleoside triphosphates (data not shown).

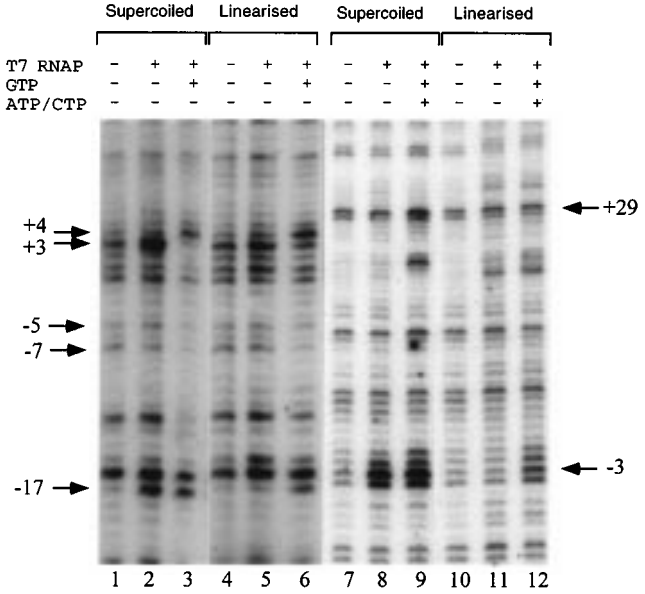

FIGURE 7: Comparison of the interaction of T7 RNA polymerase with supercoiled or linearized plasmids containing the promoter region. Primer extension was performed on the template strand. Lanes: 1−6, UV laser irradiated sample; 7−12, $KMnO_4$ attack; 1−3 and 7−9, supercoiled plasmid; 4−6 and 10−12, linearized plasmid; 1, 4, 7, and 10, DNA alone; 2, 5, 8, and 11, DNA and T7 RNA polymerase; 3 and 6, DNA, T7 RNA polymerase, and GTP: 9 and 12, DNA, T7 RNA polymerase, GTP, ATP, and CTP. The arrows indicate the position relative to the start site of transcription of important reactive sites.

In addition to changes in $KMnO_4$ sensitivity within the promoter region, other differences were also observed in the downstream region. Thus, in the presence of GTP and ATP, changes were observed at positions +29 on the T strand and +30/+31 on the NT strand, and when GTP, ATP, and CTP were present, hyperreactivity was observed at thymines +11, +13, +22, and +29 on the T strand and at +8, +18, and +25 on the NT strand. However, when the experiments were performed using 3′-dCTP as a chain terminator, these hyperreactivities were no longer present, while the thymines in the TATA sequence remained hyperreactive (data not shown). We conclude that the spurious downstream $KMnO_4$ hyperreactivities were likely to result from complexes that had extended beyond positions +7 due to substrate contamination or to base misincorporation.

Interaction of T7 RNA Polymerase with Linear Templates. The pPK3 plasmid was linearized, and the experiments were repeated. In the absence of nucleotides no specific signals due to the presence of the protein were observed by DNase I (not shown), UV laser, or $KMnO_4$ footprinting (Figure 7), probably reflecting the poor affinity of the polymerase for this long linear template. Nevertheless, in the presence of GTP, the photofootprinting pattern already observed with supercoiled plasmid was restored (Figure 7) and all the characteristics of an abortive complex synthesizing small poly(G) transcripts were present. The permanganate hyperreactivity at positions −3 present in an initiation complex on a supercoiled template (Figure 7, lanes 8 and 9) appears only in the presence of GTP or GTP/ATP/CTP (Figure 7, lane 12) on a linear template. These findings indicate that supercoiling thermodynamically stabilizes the open complex as does the addition of the first nucleotides.

DISCUSSION

In binary complexes formed on supercoiled DNA UV laser cross-linking reveals major contacts on the template strand at positions −17, −5, and +3 and on the nontemplate strand at position +8 (Figure 8). These four contacts are thus signatures of the binary complex and lie within the DNase I footprint which delimits the global occupancy of the promoter as being from positions −21 to +11. We assume that these signals represent the major populated species under the conditions used.

Even though T7 RNAP and *E. coli* RNAP are structurally dissimilar and recognize different promoter sequences, they are functionally related. It is thus of interest to compare the results reported here with those obtained with the bacterial enzyme using similar methods (Figure 8) (*26*, *31*−*34*). We note that all direct contacts between T7 RNAP and its promoter occur within the consensus region from positions −17 to +6 (Figure 8). This is paralleled in the situation with *E. coli* RNA polymerase at the *lac* UV5 promoter, where contacts are found mainly within the two consensus hexamers at positions −35 and −10 (Figure 8), irrespective of the fact that although *E. coli* RNA polymerase interacts across four DNA helix turns, T7 RNA polymerase interacts across two helix turns.

Transposition of these data to a three-dimensional representation provides further insight into the organization of these two nucleoprotein complexes. The respective promoters are depicted as canonical B-form DNA from two viewpoints orientated at 180° with respect to each other in Figure 8. This choice of orientation is based upon the results of experiments using ethylation interference and protection toward cleavage by Fe−EDTA (*9*, *32*, *35*). Both polymerases predominately view a single face of their respective promoters (untwisting of the putative open region around positions −5 to +2 for T7 and −10 to +4 for *lac* perturbs the local DNA structure such that the orientation of the contacts within this interval and downstream is ambiguous). If, however, we limit our analysis to that part of the DNA which we know is duplex in the final open complexes, then it is striking that the contacts at positions −17 and −8 for T7 RNAP and position −34 for *E. coli* RNAP are located in the major groove facing the polymerase. In neither case are contacts observed on the opposite face.

On a straight double-stranded helix, the distance separating the extreme border of the footprints, as revealed by DNase I, appears to be greater than the size of the T7 RNA polymerase. Such was also the case for *E. coli* RNAP. In this instance it was shown that the DNA sequence was bent and wrapped around the polymerase (*36-41*; cf. also ref *42* for RNA pol II). In agreement with a hypothesis made by Sousa (*2*), we postulate that the same type of distortion is also induced by T7 RNA polymerase binding on the following grounds: (a) recent observations by Martin (personal communication) using gel shift assays have provided physical evidence for bending of the promoter upon binding of the polymerase. (b) DNase I cleavage is enhanced whenever there is widening of the corresponding minor groove (*43*), and we observe a DNase I hypersensitive site at position −10. A strong bend here would place the corresponding minor groove at the outside, such that the enzyme could contact at positions −17 and −8 through the major grooves on the opposite face. A similar argument could be made for the DNase I hypersensitivity sites observed between positions −25 and −27 when *E. coli* RNA polymerase binds at the *lac* UV5 promoter. (c) An analogy can also be drawn with the binding of cyclic AMP receptor protein (CRP) to its target DNA, as CRP contacts DNA in two consecutive major grooves and bends the DNA toward the protein, thus considerably widening the opposite minor groove. The curvature is due to a kink of 40° between the adjacent bases TG in the consensus sequence (*44*). The hyperreactive DNase I sites for T7 and *E. coli* polymerases are also located at TG steps, suggesting that a similar distortion might exist at these positions in the respective promoters.

The upstream promoter contact at position −17 lies at the extreme edge of the region defined by the consensus sequence. This base is moderately conserved among T7 promoters, and substitution of other nucleotides at this position does not greatly diminish promoter strength. Thus, we identify a contact at position −17 that, although not involved in promoter specificity, is a characteristic of the final binary complex. This contact is located in the upstream major groove facing the polymerase (Figure 8). It is accompanied by a change in the geometry of the major groove as revealed by a decrease in thymine dimer formation between bases −17 and −18 on the NT strand. In the *lac* UV5 promoter, a strong UV cross-link is made at position −34 (*26*). This contact is also located in a major groove facing the polymerase and is adjacent to decreased thymine dimer formation between T−35 and T−36 (Figure 8), again indicating a major groove oriented intrusion. In this respect, the nature of the contact at position −34 is similar to the contact at position −17 in the T7 promoter, possibly reflecting a common function. We suggest that the polymerases use these major groove contacts as important anchorage points, working in synergy with downstream contacts in or around the putative locally melted regions in initiation complexes.

All previous studies of binary complexes formed by T7 RNAP with its promoter failed to detect RNAP−promoter contacts downstream of position −3, even though this region must be melted open prior to or during initiation. The results in this study reveal these contacts for the first time and demonstrate changes in the nature of the contacts during the early stages of initiation. Sousa and co-workers demonstrated that formation of an open complex, while rapid, is thermodynamically disfavored on linear templates (*13*).

As we and others had shown that the stability of binary complexes is greater on templates in which the collapse of the open complex is less favored (i.e., on partially single-stranded promoters or on supercoiled templates), we chose to study the properties of RNAP−promoter complexes on supercoiled plasmid templates. However, we have also performed similar analyses on linear templates. At concentrations of RNAP that give saturating signals at positions −17 and +3 supercoiled templates, no corresponding signature was obtained on linear templates, nor were thymines reactive toward $KMnO_4$ in the putative open region (Figure 7).

In the binary complex on a supercoiled template, T7 RNAP makes contacts at positions −5 and +3 on the T strand. These contacts straddle the putative melted region defined by endonuclease sensitivity (*45*). In contrast, at the *lac* UV5 promoter, *E. coli* RNAP makes contacts predomi-

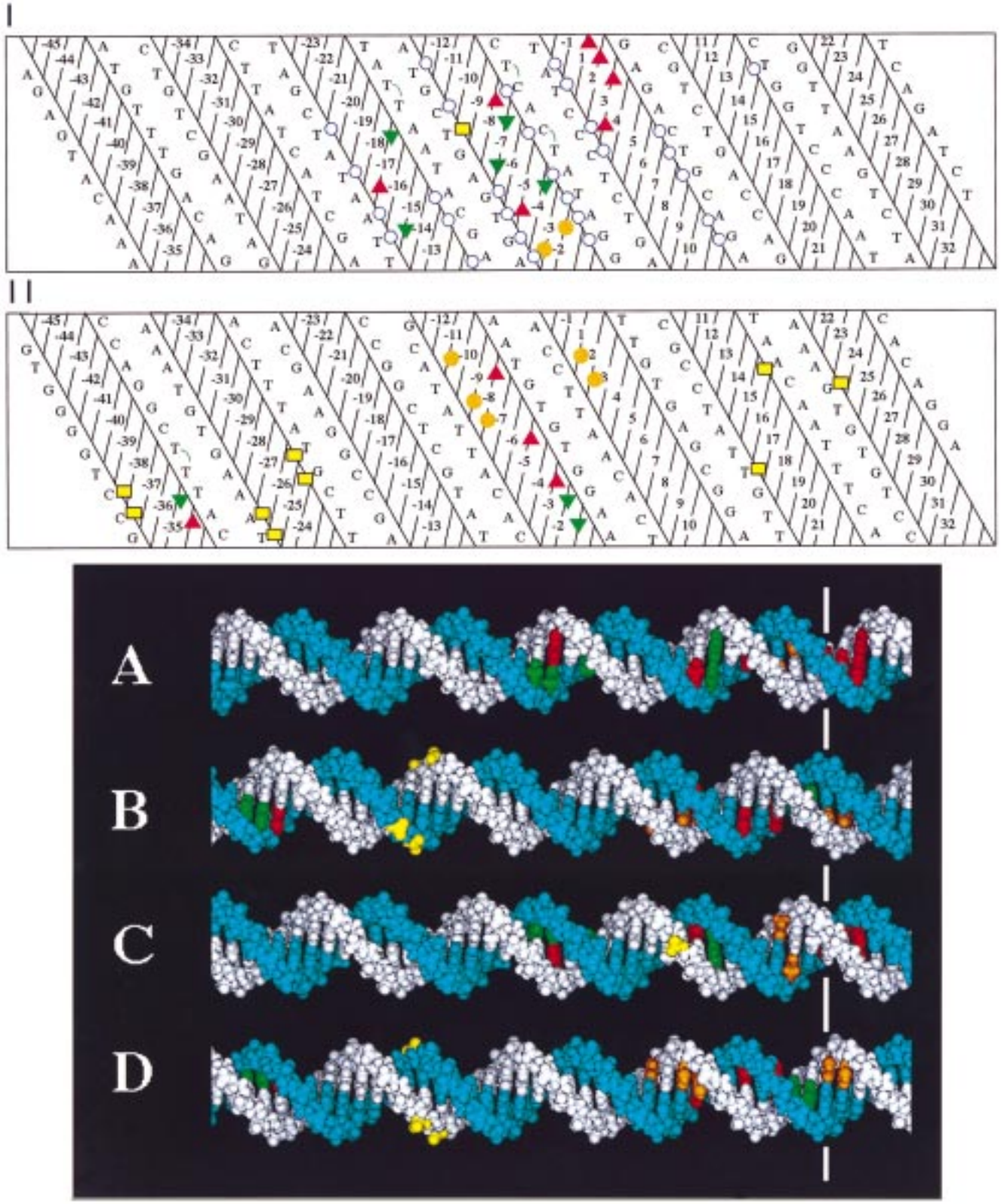

FIGURE 8: (Upper panel) Summary of contacts between the T7 (I) and *E. coli* (II) RNA polymerases and their respective T7 and *lac* UV5 promoters on a two-dimensional projection of the DNA helix. Bases are numbered with respect to the start point of transcription; the strand on the left is the template. Symbols: (▲) increased photoreactivity; (▼) decreased photoreactivity; ( )) decreased pyrimidine dimer formation; (●) $KMnO_4$ hyperreactivity; (□) DNase I hypersensitivity; (○) DNase I protection. (Lower panel) Three-dimensional representation of the contacts between the T7 (A, C) and *E. coli* (B, D) RNA polymerases and their respective T7 consensus and *lac*UV5 promoters. The DNA was drawn in a B-form using the Insight II software on a Silicon Graphics station. The template strand is white and the nontemplate strand is blue. Sequences shown are the same as those presented in the upper panel from positions −37 to +5. Panels A and B represent that side of the helix of the promoter facing the polymerase; panels C and D represent the same sequence rotated by 180°. The vertical line indicates the +1 start site of transcription. Color code for the bases: red, increased photoreactivity; green, decreased photoreactivity and decreased pyrimidine dimer formation; orange, $KMnO_4$ hyperreactivity. Color code for the phosphates: yellow, DNase I hypersensitivity.

nantly within the melted region and on the NT strand (ref *19* and references therein). It is noteworthy that $KMnO_4$ sensitivity is also confined to this region (i.e., position −2 on the NT strand and position −3 on the T strand for T7 RNAP and positions +1, +2, −8, −9, and −11 on the T strand for *E. coli* RNAP) (Figure 8). This reactivity is located at the TATA box and close to the starting site. There is clearly a substantial difference in the relative size of the two putative open regions. The manner in which the two enzymes nucleate and maintain the melted regions may therefore involve different mechanisms.

Specific recognition of the promoter by T7 RNAP involves interactions between amino acids in the specificity loop (residues 742−773) and the base pairs from positions −11 to −8 (*45*, *47*). On the basis of the functional groups involved, it has been proposed that T7 promoter recognition is mediated largely through direct or water mediated hydrogen bonds (*48*, *49*). Using the UV *cross*-linking technique, we however observed only the contact at position −8 on the nontemplate strand. Why might this be so? First, this could be due to physicochemical parameters, as UV laser cross-linking may require more intimate contacts than are associated with these interactions. Furthermore, while there is no predicted preference for the reactivity of individual amino acids, the photoreactivity of individual bases varies (T ≫ C > A > G) (*50*, *51*). A more intriguing possibility is that the interactions that are involved in recognition are short-lived and subsequently lead to a more stable binary complex. Rapid mixing methods coupled with UV laser cross-linking may allow a determination of the temporal order of events during binary complex formation and thus clarify this matter.

Of particular interest is the manner in which these contacts change during the subsequent stages of transcription. In the steps leading to a transcriptionally competent complex between T7 RNAP and its promoter, it has been demonstrated that the formation of the first phosphodiester bond or a conformational change just prior to its formation is the rate-limiting step (*16*). In this work, we have characterized species that are expected to form both before and after this event. In the absence of substrate we observe contacts at positions −5 and +3, but in the presence of GTP these contacts disappear and new contacts at position +4/+5 appear. In the presence of GTP and ATP or GTP, ATP, and CTP, the latter contacts subsequently disappear (Figure 4). However, during all stages of abortive initiation (when the polymerase is continuously synthesizing and releasing nascent RNA), the upstream contact at position −17 is maintained. This result is consistent with our DNase I data (Figure 5) and with hydroxyl radical footprinting experiments in which the upstream promoter contacts were observed to be maintained while the downstream boundary was extended during initiation (*7*). A previous suggestion that the extended footprint might be due to a transient passage of the RNAP over this region during the course of the reaction (as opposed to the formation of a static complex) (*9*) is excluded by the present results, as we observe maintenance of the upstream contact at position −17 even in the absence of downstream contacts on the promoter.

ACKNOWLEDGMENT

We thank M. Dreyfus and C. Martin for helpful discussions and G. Legat for technical assistance.

REFERENCES

1. McAllister, W. T. (1997) *Nucleic Acids and Molecular Biology* (Eckstein, F., and Lilley, D. M. J., Eds.) Vol. 11, pp 15−26, Springer-Verlag, Berlin and Heidelberg.
2. Sousa, R., Chung, Y. J., Rose, J. P., and Wang, B. C. (1993) *Nature 364*, 593−599.
3. Jeruzalmi, D., and Steitz, T. A. (1998) *EMBO J. 17*, 4101−4113.
4. Basu, S., and Maitra, U. (1986) *J. Mol. Biol. 190*, 425−437.
5. Shi, Y. B., Gamper, H., and Hearst, J. E. (1988) *J. Biol. Chem. 263*, 527−534.
6. Gunderson, S. I., Chapman, K. A., and Burgess, R. R. (1987) *Biochemistry 26*, 1539−1546.
7. Ikeda, R. A., and Richardson, C. C. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 3614−3618.
8. Chapman, K. A., Gunderson, S. I., Anello, M., Wells, R. D., and Burgess, R. R. (1988) *Nucleic Acids Res. 16*, 4511−4524.
9. Muller, D. K., Martin, C. T., and Coleman, J. E. (1989) *Biochemistry 28*, 3306−3313.
10. Mookhtiar, K. A., Peluso, P. S., Muller, D. K., Dunn, J. J., and Coleman, J. E. (1991) *Biochemistry 30*, 6305−6313.
11. Sousa, R., Patra, D., and Lafer, E. M. (1992) *J. Mol. Biol. 224*, 219−334.
12. Jia, Y., Kumar, A., and Patel, S. S. (1996) *J. Biol. Chem. 271*, 30451−30458.
13. Villemain, J., Guajardo, R., and Sousa, R. (1997) *J. Mol. Biol. 273*, 958−977.
14. Diaz, G. A., Rong, M., McAllister, W. T., and Durbin, R. K. (1996) *Biochemistry 35*, 10837−10843.
15. Ujvari, A., and Martin, C. T. (1996) *Biochemistry 35*, 14574−14582.
16. Jia, Y., and Patel, S. S. (1997) *Biochemistry 36*, 4223−4232.
17. Pashev, I. G., Dimitrov, S. I., and Angelov, D. (1991) *Trends Biol. Sci. 16*, 323−326.
18. Hockensmith, J. W., Kubasek, W. L., Vorachek, W. R., and von Hippel, P. H. (1993) *J. Biol. Chem. 268*, 15712−15720.
19. Buckle, M., and Buc, H. (1994) in *Transcription: Mechanisms and Regulations* (Conaway, R. C., and Conaway, J. W., Eds.) pp 207−225, Raven Press Ltd., New York.
20. Eichenberger, P., Déthiollaz, S., Buc, H., and Geiselmann, J. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 9022−9027.
21. Sastry, S. S., Spielmann, H. P., Hoang, Q. S., Phillips, A. M., Sancar, A., and Hearst, J. E. (1993) *Biochemistry 32*, 5526−5538; (1994) *Biochemistry 33*, 1616 (Correction).
22. Sastry, S. S. (1996) *Biochemistry 35*, 13519−13530.
23. He, B., Rong, M., Lyakhov, D., Gartenstein, H., Diaz, G., Castagna, R., McAllister, W. T., and Durbin, R. K. (1997) *Protein Expression Purif. 9*, 142−151.
24. King, G. C., Martin, C. T., Pham, T. T., and Coleman, J. E. (1986) *Biochemistry 25*, 36−40.
25. Martin, C. T., and Coleman, J. E. (1987) *Biochemistry 26*, 2690−2696.
26. Buckle, M., Geiselmann, J., Kolb, A., and Buc, H. (1991) *Nucleic Acids Res. 19*, 833−840.
27. Adelman, K., Brody, E. N., and Buckle, M. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 15247−15252.
28. Panyutin, I. G., Kovalsky, O. I., and Budowsky, E. I. (1989) *FEBS Lett. 258*, 274−276.
29. Martin, T. C., Muller, D. K., and Coleman, J. E. (1988) *Biochemistry 27*, 3966−3974.
30. Bonner, G., Lafer, E. M., and Sousa, R. (1994) *J. Biol. Chem. 269*, 25120−25128.
31. Sasse-Dwight, S., and Gralla, J. D. (1989) *J. Biol. Chem. 264*, 8074−8081.
32. Siebenlist, U., Simpson, R. B., and Gilbert, W. (1980) *Cell 20*, 269−281.
33. Spassky, A., Kirkegaard, K., and Buc, H. (1985) *Biochemistry 24*, 2723−2731.
34. Buckle, M., and Buc, H. (1989) *Biochemistry 28*, 4388−4396.
35. Jorgensen, E. D., Durbin, R. K., Risman, S. S., and McAllister, W. T. (1991) *J. Biol. Chem. 266*, 645−651.
36. Buc, H. (1986) *Biochem. Soc. Trans. 14*, 196−199.
37. Kuhnke, G., Fritz, H. J., and Ehring, R. (1987) *EMBO J. 6*, 507.

38. Heumann, H., Richetti, M., and Werel, W. (1988) *EMBO J. 7*, 4379.
39. Zhang, L., and Gralla, J. D. (1989) *Gene Dev. 3*, 1814−1822.
40. Rees, W. A., Keller, R. W., Vesenka, J. P., Yang, G., and Bustamante, C. (1993) *Science 260*, 1646−1649.
41. Rivetti, C., Guthold, M., and Bustamante, C. (1998) in *Practical Approach Series* (Buckle, M., and Travers, A., Eds.) Oxford University Press, Cambridge (in press).
42. Kim, T. K., Lagrange, T., Wang, Y., Griffith, J. D., Reinberg, D., and Ebright R. H. (1997) *Science 94*, 12268−12273.
43. Suck, D., Lahm, A., and Oefner, C. (1988) *Nature 332*, 464−468.
44. Schultz, S. C., Shields, G. C., and Steitz, T. A. (1991) *Science 253*, 1001−1007.
45. Osterman, H. L., and Coleman, J. E. (1981) *Biochemistry 20*, 4884−4892.
46. Diaz, G. A., Raskin, C. A., and McAllister, W. T. (1993) *J. Mol. Biol. 229*, 805−811.
47. Rong, M., He, B., McAllister, W. T., and Durbin, R. K. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 515−519.
48. Schick, C., and Martin, C. T. (1995) *Biochemistry 34*, 666−672.
49. Li, T., Ho, H. H., Maslak, M., Schick, C., and Martin, C. T. (1996) *Biochemistry 35*, 3722−3727.
50. Hockensmith, J. W., Kubasek, W. L., Vorachek, W. R., and von Hippel, P. H. (1986) *J. Biol. Chem. 261*, 3512−3518.
51. Pemberton, I. K., Buckle, M., and Buc, H. (1996) *J. Biol. Chem. 271*, 1498−1506.
52. Spassky, A. (1986) *J. Mol. Biol. 186*, 99−103.

BI982689E