Photocross-linking of the RNA Polymerase I Preinitiation and Immediate Postinitiation Complexes

IMPLICATIONS FOR PROMOTER RECRUITMENT

Received for publication, October 28, 2003, and in revised form, April 23, 2004
Published, JBC Papers in Press, May 25, 2004, DOI 10.1074/jbc.M311828200

Anka Bric, Catherine A. Radebaugh, and Marvin R. Paule‡
From the Department of Biochemistry and Molecular Biology, Colorado State University, Fort Collins, Colorado 80523-1870

The architecture of eukaryotic rRNA transcription complexes was analyzed, revealing facts significant to the RNA polymerase (pol) I initiation process. Functional initiation and elongation complexes were mapped by site-specific photocross-linking to template DNA. Polymerase I is recruited to the promoter via protein-protein interactions with DNA-bound transcription initiation factor-IB. The latter’s TATA-binding protein (TBP) and TAFs photocross-link to the promoter from −78 to +10 relative to the tis (+1). Although TBP does not bind DNA using its TATA-binding saddle, it does photocross-link to a 22-bp sequence that does not resemble a TATA box. Only TAF96 (the mammalian TAF68, yeast Rrn7p homolog) overlaps significantly with the DNA interaction cleft of pol I based on modeling to the pol II crystal structure. None of the pol I-specific subunits that are localized on the lips of the cleft (A49 and A34.5) or the pol I-specific stalk (A43 and A14) cross-link to DNA. Pol I does not extend significantly upstream of the promoter-proximal border of the factor complex (∼11 to −14), and similarly in the promoter proximal elongation complex, the enzyme does not contact DNA upstream of its normal exit from the cleft.

RNA polymerase (pol) I transcribes exclusively the precursor that is processed into the 28, 18, and 5.8 S RNAs. These RNAs, along with ribosomal proteins and 5 S RNA, an RNA polymerase III product, make up the ribosome. Since ribosomal RNA comprises about 75% of the total RNA being transcribed in the growing cell, it is energetically critical to tightly regulate its transcription. More importantly, there has accumulated recently compelling evidence that excessive ribosome elaboration, including transcription by both pol I and pol III, is a key element of carcinogenesis. Pivotal tumor suppressors, Rb and p53, and the oncogene c-myc have been shown to be important regulators of rRNA, 5 S RNA, and tRNA transcription, regulators that are compromised in a majority of cancers. It is, indeed, an open question whether this misregulation is a consequence or a more direct cause of neoplasia.

In Acanthamoeba castellanii, transcription is regulated at the level of recruitment of pol I to the promoter (1), whereas in mammals transcription is either modulated by promoter clearance (2) or by pol I recruitment (3–6). Promoter recruitment involves interaction of pol I and/or of a factor tightly associated with polymerase with previously bound transcription factor(s) (3, 4, 7). However, the protein-protein interactions that mediate this step are unclear. In the yeast, Saccharomyces cerevisiae, the transcription factor Rrn3p interacts with both the pol I-specific A43 subunit and one of the TAFs, Rrn6p, that bind the promoter DNA as part of core factor, prompting Peyroche et al. (8) to suggest Rrn3p bridges between pol I and core factor to bring about recruitment. However, Aprikian et al. (9) showed by direct biochemical analysis that yeast pol I in the absence of Rrn3p can be bound to the promoter, but the resulting complex is inactive in subsequent steps leading to promoter clearance. Similarly, mouse pol I is recruited to its promoter in the absence of its Rrn3p ortholog, TIF-IA, which must be added before subsequent steps in initiation can occur (2, 6, 10). In contrast, human pol I is only recruited to the promoter in the presence of Rrn3/TIF-IA (3, 4).

TIF-IA association with pol I in mouse, human, yeast, and probably all other organisms is the target of growth regulation (4–6, 10, 12, 13). Whereas in the mammals, it appears that it is the TIF-IA itself that is regulated, in yeast, it may be the association of TIF-IA with pol I that is modulated, in this case by the phosphorylation state of the polymerase (12). In A. castellanii, pol I is not recruited to the promoter when isolated from transcriptionally inactive cells (1), suggesting that in this organism TIF-IA plays a critical role in recruitment (9). Thus, in order to fully understand regulation, the protein-protein interaction(s) that mediate pol I recruitment and promoter clearance need further elucidation.

A. castellanii, a small free-living amoeba, provides the best model for studying pol I transcription biochemically. The A. castellanii in vitro transcription system is unusually robust, allowing a meaningful structural analysis of active complexes using a variety of structural techniques (14–18); over 90% of assembled complexes are transcriptionally active. The components of the system are well characterized (19–24). The preinitiation complex (PIC) includes only three or four proteins that can be assembled stepwise in solution using highly purified components. Transcription initiation factor-IB (TIF-IB) is the factor that recognizes the promoter (7, 15) and is the ortholog of yeast core factor, mouse TIF-IB, and human SL1. TIF-IB is composed of TATA-binding protein (TBP) and four TBP-associated factors (TAFs); named for their sizes of 145, 99, 96, and 91 kDa (24). Unlike in the pol II and III functional homologs of this factor, TBP does not bind DNA using its normal TATA
interaction surface in the pol I factor (19). Transcription initiation factor IE (TIF-IE) is an accessory factor necessary for stable binding of TIF-IB to the core promoter (24), but it does not bind DNA (25). Pol I makes no sequence-specific contacts with the DNA but is recruited to the promoter by TIF-IB through protein-protein interactions (7). This step apparently requires a homolog of mouse TIF-IA that is associated with a small fraction of the pol I.

Footprinting and minor groove-binding drug inhibition studies (14) indicated that TIF-IB binds primarily in the minor groove. Thus, it seemed logical to test photocross-linking reagents that either targeted the minor groove or, better, were unbiased as to groove. In addition, our earlier cross-linking study (17) did not include pol I, and a similar photocross-linking study of RNA pol II showed it to cross-link far upstream of the transcription initiation site (tis) on the DNA (26). We were curious whether pol I also extended upstream and what implications that would have in regard to its contact with TIF-IB subunits. In order to further define TIF-IB binding and to determine where pol I contacts the DNA without regard to which groove is contacted, we have performed experiments using a derivative that is linked to the DNA sugar-phosphate backbone, thus extending into both the major and minor grooves.

We have found that TBP in TIF-IB cross-links over a 22-base pair stretch of DNA, about twice that of free yeast TBP, and excludes much of the TAFI binding. The unusually large TBP footprint is discussed in terms of the differential roles of TBP in rRNA and pol II transcription. TIF-IB TAFI8 cross-link both upstream and downstream of TBP, from −78 to +10 relative to the tis. Pol I was found to cross-link downstream to +16, in agreement with where it had been previously footprinted (15), and upstream no farther than −13. This pattern localizes the interaction between TAFI8 and pol I in the DNA-binding cleft of the polymerase. We also report that a partial clone of TAFI8 shows that it is the homolog of mammalian TAFI68 and yeast Rrn7p and that this suggests a mechanism for pol I recruitment involving multiple interactions between polymerase per se, in addition to that mediated by TIF-IA/Rrn3p.

**MATERIALS AND METHODS**

**Derivatized Promoter DNA Fragments—**Oligodeoxynucleotides were derivatized by the method of Bartholomew. All reactions before UV irradiation were performed under reduced lighting. Oligodeoxynucleotides 15–17 nucleotides in length with a phosphorothioate backbone, thus extending into both the major and minor grooves.

Derivatization was performed with water-saturated ethyl ether. Oligodeoxyribonucleotides extracted three times with water-saturated isobutyl alcohol and four times with water-saturated ethyl acetate. Oligodeoxyribonucleotides were phosphorylated with (γ32P)-ATP and annealed to pBSx(+) single-stranded plasmid (pBS+) 18 containing the RNA polymerase I promoter from −120 to +80 (17), followed by primer extension and ligation (27). Dithiothreitol was omitted from all reactions, since it reduces phenyl azides to photoinert phenyl amines (28). The plasmid was digested with BamHI and HindIII to produce a 221-bp fragment, which was purified on a 5% native polyacrylamide gel and passively eluted.

5′–3′ photocross-linking reactions contained 30 fmol of derivatized DNA fragment, 0 or 40 milliunits of A. castellanii RNA polymerase pol I (22), and no TIF-IB/TIF-IE or the amount needed to shift all template DNA in an electroforetic mobility shift assay, 20 μg HEPS, pH 7.9, 10 mM MgCl2, 1% Nonidet P-40 (U.S. Biochemical Corp.), 2.5 mg/ml bovine serum albumin, 10 mM 2-mercaptoethanol, 0.3 mg/ml pBR322, 0.1 M NaEDTA, and 10% glycerol. Reactions were incubated 20 min at 25 °C and then irradiated for 4 min at 365 nm (11 mJ/mm2) with UV light (Spectrolite model SB-100p). Complexes were digested as described (17) and electrophoresed on 7.5% and 12% SDS polyacrylamide gels, which were dried, exposed to phosphor storage screens (Eastman Kodak Co.) for 2–4 days, and visualized on a Storm PhosphorImager (Molecular Dynamics).

**Purification of Proteins—**RNA pol I was purified as described (29) through the heparin-Sepharose stage. TIF-IB/TIF-IE was purified through one round of the DNA promoter affinity column (24). This TIF-IB contains sufficient TIF-IE for the formation of stable complexes with the promoter DNA. In one test experiment, additional TIF-IE was added, but it was not found to cross-link to the DNA (data not shown).

**TIF-IB Peptide Sequencing—**TIF-IB was purified from 5 kg of A. castellanii cells through the promoter-DNA affinity column. Peptide sequencing of the purified subunits (TAFI8) was completed by the Beckman Research Institute of the City of Hope (Duarte, CA). TAFI96 peptide sequences T60 (GGQYYPYF), T113 (T12YATYTTYC), T65R, ACBGGTNGAGARATGGG) and antisense (T90, CGGRAARGGGGGTA), T65, CCCATYTCRTCNA, RCCAVGD) degenerate primer sequences.

**TAFI96 Genomic DNA Cloning and Sequencing—**Genomic DNA was isolated from 1 × 107 logarithmically growing A. castellanii cells using a GenomicPrep Cells and Tissue DNA isolation kit from Amersham Biosciences (available on the World Wide Web at www.ch.embnet.org). Multiple sequence database searches for orthologs to TAFI96 were completed using ClustalW and displayed using Boxshade at the Swiss Institute of Bioinformatics.

**Gene-specific RT-PCR—**Total cellular RNA was isolated from 2 × 107 logarithmically growing A. castellanii cells using Trizol reagent from Invitrogen. First strand cDNA was synthesized from 5 μg of total RNA and 10 pmol of either the T60 or 3′-1A (GGGAGAGGAGGCGTGCTGAGG) primers using Superscript II RNase H-free reverse transcriptase (Invitrogen) according to the manufacturer’s protocol. The second cDNA strand was synthesized using the following segments: three cycles of 94 °C for 1 min, 55 °C for 1 min, and 72 °C for 1 min, 3 min at 72 °C; three cycles of 94 °C for 1 min, 55 °C for 1 min, and 72 °C; three cycles of 94 °C for 1 min, 46 °C for 1 min, and 72 °C; 20 cycles of 94 °C for 1 min, 43 °C for 1 min, and 72 °C; and one cycle of 5 min at 72 °C. The −1.6-kbp PCR product was cloned and partially sequenced using standard methods.

**DeepView**—Modeling of the polymerase was performed using DeepView (version 3.7) available on the World Wide Web at www.ch.embnet.org. Starting from the 3D coordinates of human pol II (PDB ID: 1I50), a DNA Protein Data Bank file was created using DeepView (version 3.7) and electrophoresed on 7.5 and 12% SDS polyacrylamide gels, which were dried, exposed to phosphor storage screens (Eastman Kodak Co.) for 2–4 days, and visualized on a Storm PhosphorImager (Molecular Dynamics).

**Modeling—**Modeling (version 3.7) was used to create a single Protein Data Bank file from the 3D coordinates of human pol II (PDB ID: 1I50), plus a DNA Protein Data Bank file, and then Rasmol (version 2.6) was used to create an α-carbon backbone representation of the polymerase.

2 J. Gogain, unpublished results.

3 B. Bartholomew, personal communication.
RESULTS

Positions of DNA Derivatization and Factor Binding—Modified DNA templates were made by incorporating single derivatization sites located on primers into the A. castellanii rRNA promoter from −78 to +16 relative to tis. This part of the promoter was chosen, because it encompasses TIF-IB and pol I binding sites as shown by DNase I and methidiumpropyl-EDTA-Fe(II) (MPE) footprinting assays (7, 15). The modified derivatization sites were 2–6 base pairs apart, resulting in 27 DNA fragments each derivatized at a single, distinct site. Each primer used for primer synthesis was tested for derivatization by comparing its slower mobility with nonderivatized primers on a denaturing polyacrylamide gel (data not shown). Derivatization efficiency approached 100%. Templates were tested for TIF-IB binding using electrophoretic mobility shift assays (EMSAs). Sample assays are shown for two derivatized DNAs (Fig. 1). TIF-IB was titrated until almost 100% of the template forms complexed with slower mobility (lanes 2 and 4) than the free DNA template (lanes 1 and 3). Pol I binding was not tested by EMSAs, since the DNA/TIF-IB-pol I complex is so large that it stays in the well in this assay. We know, however, that pol I is active and that the full initiation complex is formed because we can induce transcription elongation by adding NTPs to the reaction mixture (e.g. see Fig. 5), leading to loss of photocross-linking starting first at the upstream sites, but complete loss is noted when the polymerase has cleared the promoter (Fig. 5). We also can footprint the bound pol I both before elongation and at several stalled sites downstream of the tis. In these assays, template usage is greater than 90%, and nearly all PICs formed are transcriptionally active (data not shown) (18, 31).

Specificity of Photocross-linking—In order to verify that TIF-IB and pol I were cross-linking specifically, we tested cross-linking under a variety of conditions. One example of these data is shown in Fig. 2. No consistent radiolabeled products are found in the absence of TIF-IB and pol I (lanes 1 and 9). However, we do observe a group of bands migrating between the 60- and 80-kDa markers that are irreproducibly present even when there are no added proteins other than the large excess of bovine serum albumin. These are marked with an asterisk in this and all subsequent figures. When these bands appear, there frequently are also bands between the 20- and 30-kDa markers in a 12% gel (e.g. Fig. 3B, lanes 1, 2, and 16–18). In contrast, when TIF-IB/TIF-IE is bound to the DNA derivatized at −1 (Fig. 2, lanes 1–8) or −24 (lanes 9–12), strong and reproducible cross-linking to the 96-kDa subunit occurs (lanes 2, 10, and 11). After the addition of pol I to the committed complex, a new strongly cross-linked band corresponding to the A133 subunit (the β subunit homolog (21)) is obtained (lane 3). Pol I does not cross-link in the absence of TIF-IB under the conditions used (lanes 4 and 12), in agreement with its inability to recognize promoters in the absence of TIF-IB. However, it is critical to note that if too much polymerase is added or not enough nonspecific competitor is used, pol I cross-links over the entire template, demonstrating its strong nonspecific affinity for DNA (data not shown). Therefore, in these experiments, the controlled titration of pol I and competitor DNA was carried out to eliminate this spurious background. TIF-IB cross-linking is abolished by the addition of specific competitor (lanes 5 and 6) but remains strong even in the presence of a large excess of nonspecific competitor (lane 7), demonstrating the sequence-specific interaction of TIF-IB/TIF-IE with the promoter DNA.

Identification of the polypeptides that are cross-linked is accomplished by comparison of the migration of the labeled bands on polyacrylamide gels with the electrophoretic mobility of the subunits from the purified protein, as described previously (17, 28). See Ref. 17 for a discussion of the identification of TIF-IB subunits.

A. castellanii rRNA Transcription Complexes—Photocross-linking was carried out on three different complexes. TIF-IB/TIF-IE was incubated with derivatized promoter DNA fragments to map TBP and the individual TAFs. In separate reactions, Pol I was also added to form the PIC. Pol I and TIF-IB were cross-linked at numerous sites along the promoter region to see how far each protein extends, both upstream and downstream, and to see how much overlap there is between TIF-IB and pol I. Cross-linking was also performed with the same set of proteins in the elongation complex, after pol I was moved to +7 by the addition of two NTPs, ATP and GTP. Our ability to move the polymerase in this way demonstrates that we have a complete, active A. castellanii rRNA transcription complex.

TAF6, Cross-linking—The TIF-IB TAFs cross-link from −78 to +10 on the rRNA promoter (Figs. 3–5). TIF-IE does not cross-link to the DNA, in accord with its not binding DNA (32). TAF photocross-linking was identical whether or not RNA polymerase I was present, so only the data for the quaternary
species in having only three TAF Is instead of four. Mammalian TAFI68 complex are shown. The largest TIF-IB subunit, TAFI145, is found from −78 to −17 (Fig. 3A), cross-linking mainly on one side of the helical axis. It does not cross-link to sites more proximal to the tis than −17 (Figs. 4 and 5).

TAFI96, the homolog of yeast Rrn7p and mammalian TAFI68 (see below), cross-links similarly to TAFI145 in the far upstream regions at −78, −73, −68, and −62 (Fig. 3A, lanes 1–5). It does not cross-link in most of the region where TBP is in proximity to the DNA (Fig. 3, A and B, lanes 6–9), with the exception of −35 (lane 13). Farther downstream, in contrast to TAFI145, there is pronounced cross-linking of TAFI96. The cross-linking is almost continuous from −40 all of the way to the start site at −1 (Fig. 3A, lanes 11–18, and Fig. 5A). TAFI96 also shows up one helical turn downstream of the start site, at +10 (Fig. 5B, lanes 9–12). TAFI91 cross-links only far upstream, at −78, −73, −68, and −62, and then once downstream at −17 (Fig. 3A), perhaps suggesting that it is located on the outside of TIF-IB relative to the DNA. Curiously, in the present study TAFI99 was not found to cross-link at any site, although in an earlier study (17) using 5-[(p-azidobenzoyl)-3-aminoallyl]-dUMP-derivatized template, this subunit was localized between −38 and −66 (see “Discussion”). Otherwise, the present study is in substantial agreement with the earlier one. It is interesting that in the earlier study, TAFI99 only cross-linked at positions where TAFI145 also was found, opening the possibility that TAFI99 is a degradation product of TAFI145. If so, this would bring A. castellanii TIF-IB into line with other species in having only three TAFIs instead of four.

A. castellanii TAFI96 Is the Homolog of Yeast Rrn7p and Mammalian TAFI68—A partial cDNA for TAFI96, missing the 5’-end of the gene, was cloned using PCR primers derived from peptide sequences of the isolated protein. The derived amino acid sequence of the clone was used as a query in a BLAST search of the nonredundant protein data base at the Swiss Institute of Bioinformatics. This search identified two regions of sequence similarity between TAFI96 and SpRrn7h from S. pombe. One of these regions has been shown by Boukhgalter et al. (33) to be conserved in all putative TAFI68/Rrn7p homologs.
The alignment of this region from TAF,96 with S. pombe spRrn7h, S. cerevisiae Rrn7p, murine TAF,68, and human TAF,63 is shown in Fig. 6A. As can be seen from the alignment, A. castellanii TAF,96 is very similar to all four of the other proteins in this region and is clearly an ortholog of these subunits in other eukaryotes.

The second region of similarity identified between TAF,96 and spRrn7h is also found in S. cerevisiae Rrn7p. In this region, TAF,96 is 24% identical and 38% similar over 274 amino acids to S. pombe spRrn7h and is 28% identical and 53% similar over 56 amino acids to S. cerevisiae Rrn7p. The alignment of the three proteins in this region is shown in Fig. 6B.

In analyzing the two regions of similarity between TAF,96 and the fungal proteins, it became apparent that there has been an insertion in TAF,96 and spRrn7h. The two conserved sequence blocks are contiguous in S. cerevisiae Rrn7p (see Fig. 6). However, in TAF,96 and S. pombe spRrn7h, these two regions are separated by 173 and 48 amino acids, respectively. The additional amino acids in the A. castellanii protein are not unexpected, given its larger size and the size variation seen in other homologs (i.e. 60.3 kDa for Rrn7p from S. cerevisiae to 101.8 kDa for TAF,68 from Drosophila melanogaster).

TBP—TBP cross-links in one area of the promoter, from 57 to 35 (Fig. 3B). Especially in the upstream region, it appears to exclude the cross-linking of the TAF,96, suggesting that it is in close proximity to the DNA. This region does not include a TATA element, and TBP is not believed to use its TATA-binding surface to interact with the DNA in the Pol I factor (19). Interestingly, it cross-links over a much larger region, 22 bp, than yeast TBP in pol II initiation complexes, 10 bp (see “Discussion”).

RNA Polymerase I Cross-linking Overlaps TAF,96/Rrn7p/ TAF,68—Pol I was cross-linked in both the initiation and in the elongation complexes. In the initiation complex, the two largest subunits, A185 and A133, were found to cross-link at various sites from 11 to 16 (Fig. 4, lanes 4, 6, and 8, and Fig. 5, A and B, lanes 2, 6, 10, and 14). A133 cross-links at every site tested in that region with the exception of +6. A185 cross-links at +4, +6, +10, and +16, every site tested downstream of +4. No pol I subunits were cross-linked at any of the sites from 14 to 78 (Figs. 3A and 4), so that the farthest upstream site at which we saw polymerase in the initiation complex was +11 (Fig. 4).

Pol I can be initiated and stalled at specific sites along the template by supplying a limited number of the four NTPs (18). For example, the addition of ATP and GTP allows the formation of a heptamer RNA, but not an octamer RNA (5′-AAAGG-GAC). In the elongation complex, with the polymerase stalled at +7, the farthest upstream we see A133 cross-linking is −4 (Fig. 5A, lane 11); cross-linking at −8 and −11 is seen before initiation is lost (Fig. 5A, lanes 3 and 7). This is between 4 and 7 nucleotides downstream from the upstream limit where A133 was cross-linked in the PIC (Fig. 4, lane 4), approximately the same distance the polymerase was moved. A185 cross-linking, whose most upstream site before initiation was +4 (Fig. 5B, lane 2), disappears altogether (Fig. 5B, lanes 3, 7, 11, and 15). These data suggest that either this region of polymerase has moved even farther downstream than +16 or, more likely, the DNA is situated such that the reactive derivative is not juxtaposed with a reactive portion of the polymerase (see “Discussion”).

It is particularly noteworthy that none of the pol I-specific subunits, homologs of yeast A49, A43, A34.5, A14, or A12.2, cross-link. Several of these subunits have recently been mapped by cryoelectron microscopy to the lips of the DNA binding cleft or to a stalk unique to pol I that is positioned similarly to the pol II CTD (34) (see “Discussion”).

**DISCUSSION**

We have used site-specific photo-cross-linking to map the subunits of TIF-IB/TIF-IE and RNA polymerase I on the A. castellanii rRNA promoter preinitiation complex, pol I in the elongation complex, and TIF-IB-TIF-IE after promoter clearance. The results have given us significant insights into the mechanism of pol I recruitment to the rRNA promoter.

**TIF-IB**—The different subunits of TIF-IB cross-link continuously along the promoter region from −78 to the start site; only two tested sites were without any cross-linking (derivatives at −37 and −14; cross-linking is summarized in Fig. 7). This agrees with uranyl nitrate, hydroxyl radical, and dimethyl sulfate footprinting experiments, which have shown that TIF-IB wraps around and protects DNA along the entire region from −67 to −17 (14). TAF,96 also cross-links at a site 10 bp downstream of the start site. However, this interaction does not result in strong footprinting with either DNase I, uranyl ion, methidiumpropyl-EDTA-Fe(II), or EDTA-Fe(II) in solution (14, 15). Photo-cross-linking is, therefore, more sensitive than footprinting at both ends of the interaction region in detecting protein in proximity to the DNA.

Our cross-linking indicates that TAF,145, TAF,91, and TBP interact with the region from −78 down to −17 but not farther downstream. Only TAF,96, which we also show here is the A. castellanii ortholog of yeast Rrn7p and mammalian TAF,68, extends farther downstream. It makes sense that the DNA region near the start site should be in less contact with the factor in order for pol I to readily make contact with the DNA in its DNA-binding groove formed by the two largest subunits. Indeed, this idea agrees with the cross-linking results for pol I (see below). Experiments with deletion mutations of the DNA...
ments, in accordance with the above hypothesis. In yeast, \textit{S. cerevisiae} TIF-IB/core factor subunit Rrn6p (8), we have found that the bridging interaction of Rrn3p with pol I subunit A43 and in two-hybrid and glutathione S-transferase pull-down experiments, in accordance with the above hypothesis. In yeast, pol I can be recruited to the promoter in complex with the yeast homolog of TIF-IB by promoter-bound UAF, in the absence of any other known factor (9). This is in agreement with the notion that interaction between Rrn7p and pol I A190 serves a significant role in polymerase recruitment in this organism in addition to the bridging interactions of Rrn6p and Rrn3p.

**RNA Polymerase I**—We have determined that the polymerase cross-links to about 27 bp of DNA, from −11 to +16 relative to \textit{tis}. This is in close agreement with methidiumpropyl-EDTA-Fe(II) footprinting, which estimated 26 bp protected at a number of stalled sites close to the \textit{tis} (18). The footprint extends 13–15 bp upstream and 15–17 bp downstream of +1. K\textsubscript{2}MnO\textsubscript{4} footprinting shows a melted transcription bubble from −7 to +2 in the preinitiation complex, widening to −15 to +4 in the stalled complex used in the present study (18).

The two largest subunits of polymerase, which have been shown in the pol II crystal structure to form the “jaws” of the polymerase (30, 39), are the only subunits to make substantial contact with the DNA. We do not observe cross-linking to the bacterial \textalpha subunit ortholog (AC39 in \textit{A. castellanii}), although such cross-linking is observed in the yeast pol II elongation complex at the upstream edge of its footprinted region, −28/−28 (40). This upstream edge is at −11 in \textit{A. castellanii} (18), and we see no small polymerase subunits photocross-linking in this or more upstream portions of the template (data not shown). In agreement with a study of elongating pol II (40), we find that the upstream RNA-like strand of the DNA primarily contacts A133, the homolog of the bacterial \textalpha subunit and of the yeast pol II second largest (RPB2) subunit, whereas the downstream DNA makes contact with both A185 (RPB1) and A133 (RPB2), with a bias toward RPB1. This suggests the DNA takes a path similar to that in pol II. In contrast, we only occasionally observed photocross-linking at +6 of ABC22.5, the \textit{A. castellanii} homolog of yeast ABC27 (RPB5), (data not shown). This subunit lies at the “tips” of the jaw near the entrance of the downstream DNA into the cleft in the x-ray structure of pol II (30, 39, 41). It was reported to cross-link to the DNA in pol II between +5 and +15 (26) and between +11 and +16 in pol III (42). In pol I, however, the intermittent nature of the cross-linking suggests the DNA might take a slightly different path past this region of the enzyme, or, more likely, RNA polymerase I itself has a slightly different conformation so that this subunit is not as accessible to the DNA as in pol II or pol III.

Several yeast polymerase I subunits have recently been positioned by cryoelectron microscopy adjacent to the DNA cleft’s lips (34, 43). Based on this analysis, \textit{A. castellanii} A35 (yeast A34.5) is found near one end of the cleft, adjacent to the entry...

---

* C. A. Radebaugh, B. V. McConnell, and M. R. Paule, manuscript in preparation.
point for downstream DNA, and at a point predicted to contact the DNA; yeast A49 is found on the opposite side of the cleft about halfway between the entry and exit points of the DNA but outside the cleft proper. In the present study, neither of these subunits was observed to cross-link to the DNA. Yeast A14, A43, and ABC23 form a prominent stalk outside of the cleft. Rrn5p, which interacts with A43 (8), could be positioned to interact with Rrn6p and Rrn7p. A. castellanii TAF6 (25) bound to the upstream portion of the initiation complex. In the present study, the position of the remaining TAF subunits is entirely consistent with such an interaction. Furthermore, it has been argued that the A43-A14 pair is functionally homologous to the Rpb4-Rpb7 pair in pol II (43) that is not stoichiometrically associated with pol II. Although Rpb4 cross-linked to DNA that was not in transcription complexes in the Wooddell and Burgess study, it did not cross-link in authentic elongation complexes (40). We also did not detect such cross-linking. In the yeast pol III system, the two largest subunits also cross-link along the entire stretch of the polymerase footprint, as is the case with A. castellanii pol I. However, unlike in both the pol I and pol II systems, as many as seven other subunits of pol III cross-link to the DNA, implying that there is some protein proximity to the DNA that is specific to the different polymerases (42).

In distinct contrast to cross-linking of the RNA polymerase II preinitiation complex, which showed cross-linking to upstream DNA all of the way to –53 (26), we do not find any pol I subunits cross-linking to the promoter upstream of –11, the upstream limit of its methidiumpropyl-EDTA(Fe(II)) footprint. There are several reasons why we believe that our results are an accurate reflection of the in vivo situation. From a technical viewpoint, we have carefully titrated competitor DNA in order to eliminate the nonspecific binding and cross-linking to DNA of polymerase alone (Fig. 2). Second, where cross-linking occurs, reproducible and clearly visible bands appear in the gels (Figs. 3–5), and even after extreme computer enhancement of the digitized data, we do not see any cross-linking at the upstream sites (data not shown). In the upstream region, other subunits of TIF-IB clearly cross-link at nearly every site, so there is no systematic problem with application of the method. More critically, however, our reconstituted system closely mimics the active in vivo complex; a completely homologous A. castellanii RNA transcription system, with all components from a single species, was used. We can elongate the transcript to the end of the DNA template, or we can stall the polymerase by depletion of two NTPs (Fig. 5, A and B). This reconstituted system demonstrates extremely efficient template usage; nearly 100% of the templates form active preinitiation complexes capable of promoter clearance (18), in contrast to the low efficiency by pol II-reconstituted systems. When this is done, cross-linking to pol I is strong before translocation and is lost completely or nearly completely following promoter clearance or shows the expected movement downstream in the stalled complex (Fig. 5). The data of Fig. 5 demonstrates the activity of the system used. We have also performed cross-linking experiments with yeast pol II transcription factors TBP, TFIIB, and TFIIF on the yeast HIS3 promoter. Our results are in substantial agreement with those of Kim et al. (26), who cross-linked human TFIIB and TFIIF on a viral promoter, demonstrating the reproducibility of the method in our hands. Also, since our downstream cross-linking of pol I is similar to that of pol II, there is no reason we should not detect upstream cross-linking were it actually present. Finally, a subsequent study of pol II elongation complexes, albeit different from the earlier preinitiation complex study, showed cross-linking only up to the upstream border of the footprint, the same as we see in pol I (40).

There are several reasons why, in our system, polymerase might not be expected to cross-link upstream. The transcription factor TIF-IB is so large that it may preclude any other protein binding to the promoter. We have found that in footprinting assays, neither DNase nor MPE-Fe(II) have access to the region of DNA bound by TIF-IB (7), and it is therefore unlikely that a protein as large as polymerase would have that access. The cross-linking done with pol II included transcription factors TBP, TFIIB, and TFIIF, which together are only about half the size of TIF-IB and cross-linked a span of only 33 bp (26), as opposed to 88 bp for TIF-IB. Perhaps significantly, this is not the minimal functional pol II transcription initiation complex, which would include TFIIB, TFIIF, and, on many promoters, the TAFs as well. The effect these might have on cross-linking of the upstream DNA was not evaluated (26). In addition, different TBP interactions in pol I and pol II PICs may have a significant effect on the trajectory of the upstream DNA as it exits the enzyme (see the model below). Finally, when we move pol I downstream by seven nucleotides, the upstream border of cross-linking to polymerase also moves downstream by the same distance. This is not surprising, since previous methidiumpropyl-EDTA-Fe(II) footprinting assays of TIF-IB and pol I have shown that there is very little overlap of the two factors on the promoter in the vicinity of the DNA (18, 35). In contrast to pol II, our results are in substantial agreement with the similarly functional RNA polymerase III complex. The upstream border of pol III cross-linking is at –21 in the PIC and moves the predicted distance upon stalling the polymerase at +17 (42, 44). Thus, our results with the pol I complex (Fig. 5) mirror closely the pol III data, lending universality to our results.

Model of the Path of DNA through RNA Polymerase I—Fig. 8 shows a model for the path of DNA through pol I derived from the x-ray crystallographic structure of yeast RNA polymerase II (30, 39) and the data described above. The three eukaryotic RNA polymerases share five subunits in common, and in addition, the two largest subunits are largely conserved between enzymes and across species. Thus, although there are bound to be subtle differences between the enzymes (34), because there

---

\[5\] C. Radebaugh, manuscript in preparation.

\[6\] M. M. Robinson, A. Bric, M. R. Paule, and L. A. Stargell, in preparation.
is no crystal structure for polymerase I, as a first approximation we used the crystal structure for yeast pol II on which to base a low resolution mating of the DNA and the enzyme. In the model, the DNA is seen to lie in the open groove formed by the two largest subunits of the enzyme, with the downstream DNA entering from the left and proceeding until it encounters the “wall” at the active site, whereupon it makes a nearly right angle bend and then exits the polymerase. This general scheme is based upon the models for the Taq polymerase (45) and the co-crystal structure of elongating yeast RNA polymerase II (39). This model is entirely consistent with the photocross-linking data from this study. However, because the DNA largely contacts only the two largest subunits, there is little detail revealed about the exact path of the DNA. A steady state melted bubble of 19 bp is present in the pol I elongation complex, with 9 bp of DNA-RNA hybrid. The size of this bubble and hybrid was revealed by reaction with KmmO₄ and diethylpyrocarbonate (18). In the model, these structures are not modeled, because not enough is known about their locations or paths through the polymerase.

The biggest difference between this model and that determined earlier for pol II (26) is the extent of contact between the upstream DNA, out to 78, precluding interaction with polymerase I (see above). Second, for the DNA to bend around the polymerase I factor, TBP does not similarly bend the DNA. This general scheme is based upon the models for the TATA box complexes (19). This bending may be critical to the interaction of upstream DNA with the TATA box, although they do so for pol II and pol III in parallel experiments. Finally, the nonconserved N-terminal domain might be in close enough proximity to the DNA in TIF-IB to cross-link, resulting in the extended reactivity. This domain is rather small in A. castellanii, however, making this explanation less likely (11).

Acknowledgments—We thank Richard Ebright and Blaine Bartholomew for advice on the use of the photocross-linking method and Anna Maria Al-Khouri for help with protein purifications and helpful discussions.

REFERENCES

1. Bateman, E., and Paule, M. R. (1986) Cell 47, 445–450.
2. Schnapp, A., and Grummt, I. (1991) J. Biol. Chem. 266, 24588–24595.
3. Miller, G., Panos, K. I., Friedrich, J. J., Trinkle-Mulhak, L., Lamon, A. I., and Kemederji, J. C. (2001) EMBO J. 20, 1373–1382.
4. Cavanaugh, A. H., Hirschler-Laszkiewicz, I., Hu, Q., Dunda, M., Smink, T., Mistich, T., and Rothblum, I. L. (2002) J. Biol. Chem. 277, 27423–27432.
5. Eistert, D., Pfugfelder, G., and Grummt, I. (1986) Nucleic Acids Res. 14, 8165–8180.
6. Schnapp, A., Pfeiferer, C., Rosenhauer, H., and Grummt, I. (1990) EMBO J. 9, 2857–2863.
7. Kowmin, P., Bateman, E., and Paule, M. R. (1987) Cell 50, 693–699.
8. Peyroche, G., Milkeriet, P., Bischler, N., Tschochner, H., Schultz, P., Sentenac, A., Carles, C., and Riva, M. (2000) EMBO J. 19, 5475–5482.
9. Apricag, P., Moorefeld, B., and Reeder, R. H. (2001) Mol. Cell. Biol. 21, 4847–4855.
10. Bodem, J., Doehrova, G., Hoffmann-Roehr, U., Ben, I., Sentzgraf, H., Delius, H., Vinrong, M., and Gennert, I. (2000) EMBO J. 19, 171–175.
11. Wong, J.-M., Liu, F., and Bateman, E. (1992) Gene 110, 23–32.
12. Spindler, S. R., Duester, G., Alessio, J. M., and Paule, M. R. (1979) Mol. Cell. Biol. 15, 9862–9870.
13. Milkereit, P., Schultz, P., and Tschochner, H. (1997) Biol. Chem. 478, 1433–1443.
14. Geiss, G. K., Radebaugh, C. A., and Paule, M. R. (1997) J. Biol. Chem. 272, 28243–28254.
15. Bateman, E., Iida, C. T., Kowmin, P., and Paule, M. R. (1985) Proc. Natl. Acad. Sci. U. S. A. 82, 8004–8008.
16. Marsily, M., Radebaugh, C., Geiss, G. K., Laybourn, P. J., and Paule, M. R. (1995) Mol. Genet. Genomics 267, 781–791.
17. Gong, X., Radebaugh, C. A., Geiss, G. K., Simon, M. N., and Paule, M. R. (1995) Mol. Cell. Biol. 15, 4896–4906.
18. Kahl, B. L., and Paule, M. R. (2000) J. Mol. Biol. 299, 75–89.
19. Radebaugh, C. A., Matthews, J. L., Geiss, G. K., Liu, F., Wong, J.-M., Bateman, E., Camier, S., Sentenac, A., and Paule, M. R. (1994) Mol. Cell. Biol. 14, 785–795.
20. D’Alessio, J. M., Spindler, S. R., and Paule, M. R. (1979) J. Biol. Chem. 254, 4905–4909.
21. D’Alessio, J. M., Perna, P. J., and Paule, M. R. (1979) J. Biol. Chem. 254, 11282–11287.
22. Spindler, S. R., Duester, G. L., D’Alessio, J. M., and Paule, M. R. (1978) J. Biol. Chem. 253, 4699–4675.
23. Spindler, S. R., Duester, G. L., D’Alessio, J. M., and Paule, M. R. (1978) J. Biol. Chem. 253, 3642–3648.
24. Radebaugh, C. A., Kubasova, W. M., Hoffman, L. H., Stiffer, K., and Paule, M. R. (1998) J. Biol. Chem. 273, 27708–27715.
25. Al-Khouri, A. M., and Paule, M. R. (2002) Mol. Cell. Biol. 22, in press.
26. Kim, T. K., Largrange, T., Wang, Y. H., Griffith, J. D., Reineberg, D., and El Haghtig, R. (1997) Proc. Natl. Acad. Sci. U. S. A. 94, 12268–12273.
27. Travers, A. A., and Buckle, M. D. (2000) DNA, Protein Interactions: A Practical Approach, Oxford University Press, Oxford, UK.
28. Bartholomew, B., Kassavetis, G. A., Braun, B. R., and Geiduschek, E. P. (1990) EMBO J. 9, 2197–2205.
29. Iida, C. T., and Paule, M. R. (1992) Nucleic Acids Res. 20, 3211–3221.
30. Cramer, P., Bushnell, D. A., and Kornberg, R. D. (2001) Science 292, 1863–1876.
31. Kahl, B. L., and Paule, M. R. (2001) in DNA-Protein Interactions: Principles and Protocols (Moss, T., ed) pp. 63–75, Humana Press, Totowa, NJ.
32. Al-Khouri, A. M., and Paule, M. R. (2002) Mol. Cell. Biol. 22, 750–761.
33. Boukhgaltier, B., Liu, M., Guo, A., Tripp, M., Tran, K., Huyhn, C., and Pape, L. D. (2002) Gene (Amst.) 291, 187–201.
34. Bischler, N., Bruno, L., Carles, C., Riva, M., Tschochner, H., Mailouh, V., and Schultz, P. (2002) EMBO J. 21, 4136–4144.
35. Paule, M. R., Bateman, E., Hoffman, L. Ida, C., Imboden, M., Kubasak, W., Kowmin, P. L., I, Lequst, A., Rasi, P., Yang, Q., and Zwick, M. (1991) Mol.
Cross-linking RNA Polymerase I Transcription Complexes

31267

36. Radebaugh, C. A., Gong, X., Bartholomew, B., and Paule, M. R. (1997) J. Biol. Chem. 272, 3141–3144
37. Perna, P. J., Harris, G. H., Iida, C. T., Kownin, P., Bugren, S., and Paule, M. R. (1992) Gene Expr. 2, 71–78
38. Keys, D. A., Yu, L., Steffan, J. S., Dodd, J. A., Yamamoto, R. T., Nogi, Y., and Nomura, M. (1994) Genes Dev. 8, 2349–2362
39. Gnatt, A. L., Cramer, P., Fu, J. H., Bushnell, D. A., and Kornberg, R. D. (2001) Science 292, 1876–1882
40. Wooddell, C. I., and Burgess, R. R. (2000) Biochemistry 39, 13405–13421
41. Cramer, P., Bushnell, D. A., Fu, J. H., 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, 640–649
42. Bartholomew, B., Durkovich, D., Kassavetis, G. A., and Geiduschek, E. P. (1993) Mol. Cell. Biol. 13, 942–952
43. Peyroche, G., Levillain, E., Siaut, M., Callebaut, I., Schultz, P., Sentenac, A., Riva, M., and Carles, C. (2002) Proc. Natl. Acad. Sci. U. S. A. 99, 14670–14675
44. Bartholomew, B., Braun, B. R., Kassavetis, G. A., and Geiduschek, E. P. (1994) J. Biol. Chem. 269, 18090–18095
45. Zhang, G. Y., Campbell, E. A., Minakhin, L., Richter, C., Severinov, K., and Darst, S. A. (1999) Cell 98, 811–824
46. Kownin, P., Bateman, E., and Paule, M. R. (1988) Mol. Cell. Biol. 8, 747–753
47. Campbell, K. M., Ranallo, R. T., Stargell, L. A., and Lumb, K. J. (2000) Biochemistry 39, 2633–2638
48. Bateman, E., and Paule, M. R. (1988) Mol. Cell. Biol. 8, 1940–1946
