Transcription Factor B Contacts Promoter DNA Near the Transcription Start Site of the Archaeal Transcription Initiation Complex

Matthew B. Renfrow‡§, Nikolai Naryshkin¶, L. Michelle Lewis†***, Hung-Ta Chen‡‡, Richard H. Ebright¶, and Robert A. Scott§§

From the ‡Department of Biochemistry and Molecular Biology, and Center for Metalloenzyme Studies, University of Georgia, Athens, Georgia 30602-2256 and the ¶Howard Hughes Medical Institute, Waksman Institute, and Department of Chemistry, Rutgers University, Piscataway, New Jersey 08854

Received for publication, October 17, 2003
Published, JBC Papers in Press, November 3, 2003, DOI 10.1074/jbc.M311433200

Transcription initiation in all three domains of life requires the assembly of large multiprotein complexes at DNA promoters before RNA polymerase (RNAP)-catalyzed transcript synthesis. Core RNAP subunits show homology among the three domains of life, and recent structural information supports this homology. General transcription factors are required for productive transcription initiation complex formation. The archaeal general transcription factors TATA-element-binding protein (TBP), which mediates promoter recognition, and transcription factor B (TFB), which mediates recruitment of RNAP, show extensive homology to eukaryal TBP and TFIIB. Crystallographic information is becoming available for fragments of transcription initiation complexes (e.g. RNAP, TBP-TFB-DNA, TBP-TFIIB-DNA), but understanding the molecular topography of complete initiation complexes still requires biochemical and biophysical characterization of protein-protein and protein-DNA interactions. In published work, systematic site-specific protein-DNA photocrosslinking has been used to define positions of RNAP subunits and general transcription factors in bacterial and eukaryal initiation complexes. In this work, we have used systematic site-specific protein-DNA photocrosslinking to define positions of RNAP subunits and general transcription factors in an archaeal initiation complex. Employing a set of 41 derivatized DNA fragments, each having a phenyl azide photoactivatable crosslinking agent incorporated at a single, defined site within positions −40 to +1 of the gdh promoter of the hyperthermophilic marine archaea, Pyrococcus furiosus (Pf), we have determined the locations of P/RNAP subunits P/TBP and P/TFB relative to promoter DNA. The resulting topographical information supports the striking homology with the eukaryal initiation complex and permits one major new conclusion, which is that P/TFB interacts with promoter DNA not only in the TATA-element region but also in the transcription-bubble region, near the transcription start site. Comparison with crystallographic information implicates the P/TFB N-terminal domain in the interaction with the transcription-bubble region. The results are discussed in relation to the known effects of substitutions in the TFB and TFIIB N-terminal domains on transcription initiation and transcription start-site selection.

Extensive structural and biochemical characterization of transcription initiation complexes has revealed similarity of transcription systems across the three domains of life (1–4). Transcription initiation in archaea closely resembles eukaryal class II transcription initiation and is mediated by a single RNA polymerase (RNAP) and two general transcription factors, TATA-element binding protein (TBP) and transcription factor B (TFB) (5, 6). In vitro transcription initiation has been demonstrated with RNAP, TBP, and TFB in Pyrococcus (7–10), Methanococcus (11), Methanosarcina (12), and Sulfolobus (13) cell-free transcription systems.

Recent crystallographic structures of bacterial RNAP and eukaryal RNAP II reveal striking structural homology (3, 14–20), homology that apparently also extends to archaeal RNAP (3, 21–23). Archaeal and eukaryal general transcription factors also exhibit striking structural homology (24–29). Archaeal TFB belongs to the family of TATA-element-binding proteins that bind to promoter TATA elements, bend DNA, and nucleate the formation of initiation complexes (24, 25, 30). Archaeal TFB belongs to the TFIIB family, whose members bind promoter DNA, bind RNAP, and serve as bridges between the TBP-TATA-element complex and RNAP (26–29). Archaeal TFB has the characteristic domain organization of other TFIIB family members (31, 32), comprising a C-terminal domain (TFBc) that mediates interactions with the TBP-TATA-element complex (7, 27, 28) and an N-terminal domain (TFBn) that mediates interactions with RNAP (33, 34). The TFB N-terminal domain consists of a conserved metal binding region (“zinc ribbon”; Refs. 35 and 36) followed by a short, highly conserved region (“con-
served sequence block" (CSB); Ref. 33). Substitutions in the conserved sequence block in archaeal TFB indicate a role in transcription-initiation NTP concentration dependence (33), and substitutions in eukaryal TFIIB indicate roles in transcription-initiation efficiency and start-site selection (37–42). Although the roles of the TFB and TFIIB N-terminal domains in transcription initiation and start-site selection have been well characterized, the structural and mechanistic basis of these roles has remained undefined.

In published work, systematic site-specific protein-DNA photocrosslinking has been used to characterize the structural organization of eukaryal transcription initiation complexes (43–45), bacterial transcription initiation complexes (46, 47), and the part of an archaeal transcription initiation complex at and downstream of the transcription start site (positions −1 to +20 of the template strand (21)). Here, we report the use of systematic site-specific protein-DNA photocrosslinking to define the structural organization of the part of an archaeal transcription initiation complex upstream of the transcription start site (positions −40 to +1; nontemplate and template strands), the part that contains binding determinants for RNAP subunits and general transcription factors. Our results confirm the anticipated high degree of homology between the archaeal and eukaryal transcription systems and identify an unanticipated interaction between the TFB N-terminal domain and promoter DNA in the transcription-bubble region, upstream of and at the transcription start site. The results immediately suggest a mechanistic basis for the role of TFB and TFIIB N-terminal domains in transcription initiation and start-site selection.

EXPERIMENTAL PROCEDURES

P/RNAP—Under anaerobic conditions, frozen Pyrococcus furiosus cell paste (15 g, prepared as in Bryant and Adams (48)) was resuspended in 82 ml of ice-cold 50 mM Tris (pH 7.5), 22 mM NH₄Cl, 10 mM EDTA, and 10% glycerol, and cells were lysed in two passes through a French press operating at 1200 p.s.i. After the addition of Polymin P to 0.4% w/v and stirring for 30 min at room temperature, the sample was centrifuged at 100,000 × g for 4 h at 4 °C. The pellet was extracted 3 times with 100 ml of 50 mM Tris (pH 7.5), 1.2 mM NH₄Cl, 10 mM EDTA, and 10% glycerol, extracts were pooled, ammonium sulfate was added to 70% saturation, the sample was stirred for 1 h on ice, and the sample was again centrifuged at 100,000 × g for 4 h at 4 °C. The resulting pellet was aerobically resuspended in 100 ml of buffer A (50 mM Tris (pH 7.5), 50 mM KCl, 10 mM EDTA, and 20% glycerol) and dialyzed against three changes of 1 liter of the same buffer for 12 h at 4 °C (12–14-kDa molecular weight cutoff dialysis tubing (Spectrapor)). The sample was centrifuged at 200,000 × g for 1 h at 4 °C. The sample was loaded onto a 220-ml DEAE-Sepharose column (Amersham Biosciences) pre-equilibrated in buffer A and eluted with 2 liters of a linear gradient of 100% buffer A to 100% buffer B (50 mM Tris (pH 7.5), 1 mM KCl, 10 mM EDTA, and 20% glycerol). Fractions containing DNA-dependent RNA polymerase (RNAP) activity (49, 50) were pooled, concentrated, and loaded onto a 30-ml heparin-Sepharose column (Amersham Biosciences) equilibrated with 15% buffer B. The column was eluted with a 300-mM linear gradient of 15 to 100% buffer B. Pooled fractions containing RNAP activity were concentrated and dialyzed against buffer A for 12 h at 4 °C. The sample was centrifuged at 200,000 × g for 30 min at 4 °C and loaded onto a Mono Q 10/10 column (Amersham Biosciences). The column was eluted with a 100-mM linear gradient of 10–100% buffer B. Pooled fractions containing polymerase activity were concentrated and stored as aliquots at −80 °C. Yield, −47 µg.

P/TBP—Plasmid P/TBP/pT7–7, which encodes P/TBP under control of the bacteriophage T7 gene 10 promoter, was constructed by replacing the Ndel-BamHI segment of plasmid pT7–7 (51) with a Ndel-BamHI DNA fragment containing the P/TBP-coding sequence (prepared by ad hoc PCR using P. furiosus genomic DNA as template 2 (a gift of F. Jenney, University of Georgia, Athens, GA)). Transformants of Escherichia coli strain BL21 DE3 (Novagen) with plasmid P/TBP/pT7–7 were cultured at 37 °C in 4 liters of LB medium (Fisher) to an A₅₀₀ of 0.1, cultured an additional 4 h at 37 °C (A₅₀₀ ~ 2.0), and harvested by centrifugation at 3000 × g for 20 min at 4 °C. Cells pellets (10–12 g) were frozen and stored at −80 °C until used. Cells were thawed by suspension in 60 ml of lysate buffer (50 mM potassium phosphate (pH 7.8), 10 mM 2-mercaptoethanol, 1 mM phenylmethylsulfonyl fluoride, 1 µg/ml leupeptin, and 1 µg/ml DNase I) and were lysed by sonication; the lysate was centrifuged at 30,000 × g for 50 min at 4 °C. After incubation of the supernatant for 30 min at 70 °C, the sample was centrifuged at 30,000 × g for 50 min at 4 °C, the supernatant was loaded onto a 2.5 × 10-cm Q-Sepharose benchtop column (Amersham Biosciences), the column was washed with 500 ml of 50 mM potassium phosphate buffer (pH 7.8), and the column was eluted with a 150-mL linear gradient of 0–800 mM KC1 in the same buffer. Pooled fractions containing P/TBP (50 ml; 6 × 10⁵ KCl) were further pooled into 50 mM potassium phosphate (pH 7.8) and concentrated to 2 ml by ultrafiltration through a YM3 membrane (Millipore). The sample was loaded onto a 1 × 10-cm Mono Q column (Amersham Biosciences), and the column was eluted with a 55-mL linear gradient of 0–600 mM KC1 in 50 mM potassium phosphate (pH 7.0). Pooled fractions containing P/TBP (5 ml; 4 × 10⁴ KCl) were exchanged into 50 ml of 50 mM potassium phosphate (pH 7.8), 500 mM KC1, concentrated by ultrafiltration through a YM3 membrane (Millipore) at 4 °C, and stored in aliquots at −70 °C. SDS-PAGE indicated an ~21-kDa protein, corresponding to the expected molecular mass of P/TBP. Matrix-assisted laser desorption ionization time-of-flight mass spectrometry revealed a single molecular ion series of 21,285 atomic mass units (predicted, 21,311 atomic mass units). The yield was ~40 µg, and purity was ~95%.

TFPB—Plasmid pMLTfB (52), which encodes P/TBP under control of the bacteriophage T7 gene 10 promoter, was introduced by transformation into E. coli strain BL21 DE3 (Novagen). Growth, induction, harvesting, and clarification were performed as described above for P/TBP. The supernatant was loaded onto a 2.5 × 20-cm phosphocellulose column (Whatman), the column was washed extensively with 50 mM potassium phosphate buffer (pH 7.8), and the column was eluted with 150 ml of 50 mM potassium phosphate (pH 7.8) and 500 mM KC1. The sample was concentrated to ~2 ml by ultrafiltration through a YM3 membrane (Millipore, 7000 M.W. cutoff) and dialyzed fractions of the same buffer. Pooled fractions containing P/TBP (20 µl) were stored in aliquots of 500 µl at −70 °C. SDS-PAGE analysis indicated a ~34-kDa protein, corresponding to the expected molecular mass of P/TBP. The yield was ~25 µg, and the purity was ~95%.

Derived Promoter DNA Fragments—M13mp18-gdh and M13mp19-gdh were constructed by replacement of the EcoRI-Sphl segments of, respectively, M13mp18 and M13mp19 (New England Biolabs) by EcoRI-Sphl DNA fragments containing positions −7 to +55 of the P. furiosus gdh promoter (prepared by ad hoc PCR using as template plasmid pLUI479 (Ref. 8; a generous gift from M. Thomm, University of Regensburg, Regensburg, Germany). Promoter DNA fragments containing a phenyl azide photoactivatable cross-linking agent incorporated at a single, defined DNA phosphate and containing an adjacent 32P label were prepared using M13mp18-gdh (for analysis of the template strand) and M13mp19-gdh (for analysis of the nontemplate strand) as templates by the method of Naryshkin et al. (47). The samples were added to a primer extension and ligation products were digested with FcpI and Avall, 5′ overhanging ends were filled in using DNA polymerase Klone fragment (Promega) and dATP, dGTP, dCTP, and dTTP (Amersham Biosciences), and 5′ phosphates at ends and nicks were removed using calf intestinal alkaline phosphatase (Promega; 3 units for 20 min at 37 °C). Filling of 5′ overhanging ends and removal of 5′ phosphates from ends and nicks significantly reduced nonspecific cross-linking by end- and nick-bound P/RNAP.

Site-specific Protein-DNA Photocrosslinking—Reaction mixtures (43 µl) contained 0 or 2 nM P/RNAP, 1 nM P/TBP, 1 nM P/TBP, 1 nM derivatized promoter DNA fragment (200 Bq/µmol), and 40 µg/ml heparin (for transcription buffer) to generate HEPES buffer (pH 7.5), 250 mM KC1, 2.5 mM MgCl₂, 0.1 mM EDTA, 10 mM β-mercaptoethanol, 6 mg/ml chymostatin, 1 µg/ml aprotinin, 0.4 µg/ml pepstatin, 0.4 µg/ml leupeptin, 0.2 µg/ml 4-(2-aminoethyl)benzenesulfonyl fluoride, and 8% glycerol) at 70 °C. Reaction mixtures were prepared by the incubation of 2 nM P/TBP and 2 nM derivatized promoter DNA fragment (200 Bq/µmol) in 10 mM transcription buffer at 70 °C for 30 min. After incubation at 0 or 2 nM P/RNAP and 2 nM P/TBP in 20 µl of transcription buffer and incubation for 10 min at 70 °C. This was followed in experiments

a M. B. Renfrow, unpublished data.
performed in the presence of PRNAP by the addition of 3 μl of 583 μg/ml heparin and incubation for 2 min at 70 °C. Crosslinking, nuclease digestion, and product analysis were performed by a modification of the procedures of Naryshkin et al. (47). Reaction mixtures were UV-irradiated for 60 s (11 mJ/mm² at 365 nm) in a Spectrolinker XL-1000 UV crosslinker (Fisher). Reaction vessels were siliconized polypropylene microcentrifuge tubes contained within borosilicate glass culture tubes (16 / 100 mm) filled with water pre-equilibrated to 70 °C. Crosslinked polypeptides were identified by performing nuclease digestion (the addition of 2 μl of 100 mM CaCl₂ and 10 units/μl DNase I (Sigma) and incubation for 10 min at 70 °C followed by the addition of 1 μl of 30 units/μl S1 nuclease (Roche Applied Science) and incubation for 10 min at 37 °C), SDS-PAGE, and autoradiography.

RESULTS

Site-specific Protein-DNA Photocrosslinking—The procedure used consists of four steps (Fig. 1A; Refs. 43–47) as follows: (i) Chemical and enzymatic reactions were used to prepare a DNA fragment with a photoactivable crosslinking agent (R) and an adjacent radiolabel (*) incorporated at a single, defined site. Line 3, UV irradiation of the derivatized protein-DNA complex initiates crosslinking. Nuclease digestion eliminates uncrosslinked DNA and converts crosslinked DNA to a crosslinked, radiolabeled 3′–5′ nucleotide tag. (ii) The multiprotein-DNA complex of interest was formed using the site-specifically derivatized DNA fragment, and the multiprotein-DNA complex was UV-irradiated, initiating covalent crosslinking with polypeptides in direct physical proximity to the probe. (iii) Extensive nuclease digestion was performed, eliminating uncrosslinked DNA and converting crosslinked DNA to a crosslinked, radiolabeled 3′–5′ nucleotide “tag.” (iv) The tagged polypeptides were identified using denaturing polyacrylamide gel electrophoresis and autoradiography.

For analysis of the P. furiosus transcription initiation complex, we constructed 41 derivatized DNA fragments, each having a probe incorporated at a single, defined phosphate of the P. furiosus glutamate dehydrogenase promoter (gdh; each second phosphate on each strand from positions −40 to +1, Fig. 1B). For each DNA fragment, we analyzed crosslinking both in the P/TFB-P/TBP-promoter complex and in the P/PRNAP-P/TFB-P/TBP-promoter complex. For each DNA fragment, we
formed the complex, UV-irradiated the complex to initiate crosslinking, and identified crosslinked polypeptides.

Experiments were performed at 70 °C. In vitro transcription assays indicate that the PfRNAP-P/TFB-P/TBP-promoter complex exhibits optimal activity at 70 °C (with high activity at 60–90 °C (8–10)). Potassium permanganate footprinting experiments indicate that the PfRNAP-P/TFB-P/TBP-promoter complex is an open complex at 70 °C (with a single-stranded “transcription bubble” extending from at least position –7 to at least position +3 at 70 °C (53)). Heparin-challenge experiments further indicate that the PfRNAP-P/TFB-P/TBP-promoter complex is a promoter-specific and transcriptionally competent complex at 70 °C.3

Complications due to crosslinking within nonspecific, non-productive complexes were avoided by inclusion of heparin (which disrupts nonspecific and non-productive complexes on double-stranded DNA (54)) and by inclusion of filling-in and dephosphorylation steps in DNA-fragment preparation (which prevents formation of heparin-resistant nonspecific and non-productive complexes at DNA ends and DNA nicks). Identical results were obtained using an alternative procedure in which complications due to crosslinking within nonspecific and non-productive complexes were avoided by use of immobilized templates and stringent washing (see the procedures described in Kim et al. (45, 46)).3 As a direct control for specificity of crosslinking, three derivatized promoter DNA fragments with probes well outside the range of reported protein-DNA interactions in the eukaryal initiation complex (44, 45), i.e. at positions –70 and +45 on the nontemplate strand and –66 on the template strand, were constructed and analyzed. No crosslinking was detected using the three negative control derivatized promoter DNA fragments either in experiments with the PfTFB-P/TBP-promoter complex or in experiments with the PfRNAP-P/TFB-P/TBP-promoter complex (Fig. 2). Representative data for the PfTFB-P/TBP-promoter complex and the PfRNAP-P/TFB-P/TBP-promoter complex are shown in Fig. 2, and the results are summarized in Fig. 3.

PfRNAP Photocrosslinking—PfRNAP consists of 12 distinct polypeptide subunits (10, 55). Only two of these polypeptide subunits crosslinked to positions –40 to +1 of the PfRNAP-P/TFB-P/TBP-promoter complex: subunit A’ (homologous to the N-terminal portion of eukaryal RPB1) and subunit B (homologous to eukaryal RPB2) (Figs. 2B and 3B). PfRNAP subunit A’ crosslinked primarily to the transcription-bubble region on the template strand (positions –8 to –2). PfRNAP subunit B crosslinked over the entire region on both strands (positions –40 to +1).

PfTBP Photocrosslinking—PfTBP crosslinked exclusively within the TATA element (positions –28 to –22), both in the PfTFB-P/TBP-promoter complex and in the PfRNAP-P/TFB-P/TBP-promoter complex (Figs. 2 and 3). In the PfTBP-DNA complex PfTBP crosslinked with much lower efficiency,3 consistent with expectation based on the documented low stability

3 N. Naryshkin and M. B. Renfrow, unpublished data.
FIG. 3. Site-specific protein-DNA photocrosslinking of archaeal transcription complexes: summary of data. A, summary of data from experiments with the PfTFB-PfTBP-promoter complex. B, summary of data from experiments with the PfRNAP-PfTFB-PfTBP-promoter complex. In each panel, results for the nontemplate strand are shown above the sequence, and results for the template strand are shown beneath the sequence. Phosphates analyzed are indicated by arrows. Crosslinks are indicated by bars (strong, consistently reproducible crosslinks indicated by filled bars; weak crosslinks indicated by open bars). The transcription start site, the TATA element, and the TFB recognition element (BRE) are indicated by shading; the transcription-bubble region is indicated by a black rectangle; the melted transcription bubble in the PfRNAP-PfTFB-PfTBP-promoter complex is indicated in yellow; the positions that show PfRNAP-dependent enhancement of crosslinking efficiency are indicated by a yellow rectangle.

of PfTBP-DNA complexes and on the documented stabilization of PfTBP-DNA interactions by PfTFB (21, 56).

**PfTFB Photocrosslinking—**PfTFB crosslinked in the TATA-element region (positions -40 to -14; strong crosslinks at positions -40 to -30 and -22 to -14 of the nontemplate strand), both in the PfTFB-PfTBP-promoter complex and in the PfRNAP-PfTFB-PfTBP-promoter complex (Figs. 2 and 3). PfTFB also crosslinked in the transcription-bubble region (crosslinks between positions -10 and +1; strong crosslinks between positions -10 and -2), both in the PfTFB-PfTBP-promoter complex and in the PfRNAP-PfTFB-PfTBP-promoter complex (Figs. 2 and 3). Two crosslinks in the transcription-bubble region in immediate proximity to the transcription start site (crosslinks at positions -4 and -2 of the template strand) exhibited a marked dependence on PfRNAP presence, being weak or absent in the PfTFB-PfTBP-promoter complex but strong in the PfRNAP-PfTFB-PfTBP-promoter complex (Fig. 2, A and B, last two lanes below the promoter sequence; Fig. 3).

**DISCUSSION**

Our photocrosslinking results define the positions of PfRNAP subunits, PfTBP, and PfTFB relative to positions -40 to +1 of promoter DNA in the PfTFB-PfTBP-promoter and PfRNAP-PfTFB-PfTBP-promoter complexes in solution. In conjunction with the results of Bartlett et al. (21), which define positions of PfRNAP subunits relative to positions -1 to +20 of promoter DNA, our results provide an essentially complete description of protein-DNA interactions in the archaeal transcription initiation complex (Supplemental Fig. 1A).

**RNAP—**Only 2 of the 12 PfRNAP subunits were observed to crosslink to positions -40 to +1 of the PfRNAP-PfTFB-PfTBP-promoter complex: subunit A’ and subunit B (Figs. 2B and 3B). Subunit A’ crosslinked primarily in the transcription-bubble region close to the transcription start site. Subunit B crosslinked both in the transcription-bubble region and upstream of the transcription-bubble region. Our results for the archaeal initiation complex are similar to published results for the bacterial and eukaryal initiation complexes, for each of which only two RNAP subunits were observed to crosslink to positions -40 to +1: the largest subunit (the homolog of subunit A’) and the second largest subunit (the homolog of subunit B) (Supplemental Fig. 1; Refs. 3, 46, 57, 58). We infer that, in the archaeal initiation complex, RNAP interacts with promoter DNA in a manner equivalent to that in the bacterial and eukaryal initiation complexes. The RNAP active-center cleft, comprising determinants of subunit A’ and subunit B, interacts with the transcription-bubble region and transcription start site (Ref. 3, 46, 57, and 58), and the RNAP flap or wall, comprising determinants of subunit B, interacts with DNA upstream of the transcription-bubble region (Ref. 3, 46, 57, and 58).

**TBP—**PfTFB crosslinked exclusively within the TATA ele-
TFB—P/TFB crosslinked in the TATA-element region yielding strong crosslinks immediately upstream and downstream of the TATA element, both in the P/TFB-P/TBP-promoter complex and in the P/RNAP-P/TFB-P/TBP-promoter complex (Figs. 2 and 3). TFB-DNA crosslinking in the TATA-element region is as expected based on crystallographic structures of TBP-TFBc-DNA complexes (27, 28) and on footprinting experiments with TBP-TFBc-DNA complexes (7, 11, 13). TFB-DNA crosslinking immediately upstream of the TATA element is as expected for sequence-specific interaction between the BH4-BH5 helix-turn-helix motif of the TFB C-terminal domain and the TFB recognition element (Refs. 13, 28, 59, 61; see also Ref. 29), and TFB-DNA crosslinking immediately downstream of the TATA element is as expected for interaction between the BH2-BH3 loop of the TFB C-terminal domain and the DNA minor groove downstream of the TATA element (13; see also Ref. 29). The pattern of TFB-DNA crosslinking in the TATA-element region of the archaeal transcription initiation complex is similar to the pattern of TFIIIB-DNA crosslinking in the TATA-element region of the eukaryal RNAPII transcription initiation complex (Supplemental Fig. 1; Refs. 43–45 and 59), supporting the homology between the two transcription systems.

P/TFB also crosslinked in the transcription-bubble region, yielding crosslinks close to and at the transcription start site, both in the P/TFB-P/TBP-promoter complex and in the P/RNAP-P/TFB-P/TBP-promoter complex (Figs. 2 and 3). TFB-DNA crosslinking in the transcription-bubble region spanned positions −10 to +1 and involved both DNA strands. TFB-DNA crosslinking close to the transcription start site on the template strand exhibited a pronounced dependence on RNAP, being weak or absence in the P/TFB-P/TBP-promoter complex but strong in the P/RNAP-P/TFB-P/TBP-promoter complex (Fig. 2). We infer that residues of TFB contact or closely approach essentially the entire transcription-bubble region, including the transcription start site. Furthermore, we infer that residues of TFB contact or closely approach both strands of the melted transcription bubble within the P/RNAP-P/TFB-P/TBP-promoter complex and, thus, most likely are located between the two strands of the melted transcription bubble, within the RNAP active-center cleft in the P/RNAP-P/TFB-P/TBP-promoter complex. Finally, we infer that RNAP modulates proximity between TFB and DNA close to the transcription start site through interactions with TFB and/or through interactions with DNA (e.g. melting of the transcription bubble).

In analysis of eukaryal transcription initiation complexes, no crosslinks were observed between TFIIb and the transcription start site region (i.e. no TFIIb-DNA crosslinks were observed downstream of position −6 (44, 45)). However, we emphasize that this negative result permits no strong conclusions for two reasons: (i) the azidophenacyl crosslinking chemistry used is chemoselective (with an essentially absolute requirement that a probeproximal segment of a target polypeptide contains a lysine, histidine, or tyrosine residue for efficient reaction) and (ii) the eukaryal complexes analyzed were closed complexes (in the absence of ATP) or mixtures of closed complexes and open complexes (in the presence of ATP) (with a predominance of closed complexes). We favor the view that, in accord with the high sequence and structural homology between the eukaryal and archaeal initiation complexes, both archaeal TBF and eukaryal TFIIb closely approach the transcription start site region in their respective open initiation complexes but that for technical reasons the interaction is readily detectable for TFB but not readily detectable for TFIIb.

Structural Implications of P/TFB Downstream Crosslinking—Comparison of the observed TFB-DNA crosslinks (Figs. 2 and 3) with the crystallographic structures of complexes of TBP, TFB C-terminal domain, and DNA (27, 28) allows us to ascribe all TFB-DNA crosslinking in the TATA-element region (positions −40 to −14) to interactions between the TFB C-terminal domain and DNA (Supplemental Fig. 2).

However, we also observe TFB-DNA crosslinking in the transcription-bubble region, close to and at the transcription start site (positions −10 to +1; Figs. 2 and 3). Based on crystallographic structures of complexes of TBP, TFB C-terminal domain, and DNA (27, 28), this crosslinking cannot be ascribed to interactions between the TFB C-terminal domain and DNA. (Crystallographic structures of complexes of TBP, TFB C-terminal domain, and DNA show no interactions between TFB C-terminal domain and DNA downstream of position −16 (Supplemental Fig. 2).) There is only one structurally plausible explanation for the observed TFB-DNA crosslinking in the transcription-bubble region, near the transcription start site; i.e. a part of TFB other than the TFB C-terminal domain contacts or closely approaches the transcription-bubble region, near the transcription start site. We infer that residues of the TFB N-terminal domain and/or the TFB interdomain linker contact or closely approach the transcription-bubble region, near the transcription start site.

The proposed interaction of residues of the TFB N-terminal domain and/or interdomain linker with the transcription-bubble region, near the transcription start site, immediately suggests a structural and mechanistic explanation for roles of the TFB and TFIIb N-terminal domains in post-recruitment reactions involving the transcription start site: transcription-initiation NTP concentration dependence (33), transcription initiation efficiency (33, 38–40), and start-site selection (33, 37–42). We propose that residues of the TFB and TFIIb N-terminal domains contact or closely approach the transcription-bubble region near the transcription start site and mediate, through direct interactions near the transcription start site, roles in these post-recruitment reactions.

The TFB and TFIIb N-terminal domains each consists of a zinc ribbon (35, 36, 52) followed by a CSB. Available evidence suggests that the zinc ribbon is responsible for effects of TFB and TFIIb on recruitment of RNAP to initiation complexes (31, 34, 38, 39) and that the CSB is responsible for effects of TFB and TFIIb on transcription-initiation NTP concentration dependence, transcription-initiation efficiency, and start-site selection (33, 37–42). Thus, deletions or substitutions within the zinc ribbon prevent recruitment of RNAP (31, 34, 38, 39), whereas deletions or substitutions within the CSB do not affect recruitment of RNAP (33, 39, 40, 42) but do affect transcription-initiation NTP concentration dependence (33), transcription-initiation efficiency (33, 38–40), and start-site selection (37–42). We propose that residues of the TFB and TFIIb N-terminal-domain CSBs contact or closely approach the transcription-bubble region near the transcription start site and mediate, through direct interactions near the transcription start site, roles in transcription.

---

4 N. Naryshkin and R. H. Ebright, unpublished data.

5 T.-K. Kim, D. Reinberg, and R. H. Ebright, unpublished data.
tion-initiation NTP concentration dependence, transcription-initiation efficiency, and start-site selection.

Consistent with the above proposals, footprinting experiments indicate that TFB induces DNase I hypersensitivity within the transcription-bubble region (at positions ~6 and ~5 of the nontemplate strand) and that induction of DNase I hypersensitivity does not require the TFB zinc ribbon but does require the TFB CSB. In the bacterial transcription initiation complex, a region of the initiation factor σ contacts or closely approaches the transcription-bubble region near the transcription start site: σ region 3.2 (σR3.2, also known as the σR–σ linker (18, 19)). σR3.2 consists of two sub-regions: an extended segment that binds within the RNAP RNA exit channel (18, 19, 57) and a hairpin loop that binds within the RNAP active-center cleft between the two strands of the melted transcription bubble, near the transcription start site (18, 19). Deletion of the σR3.2 hairpin loop affects transcription-initiation NTP concentration dependence and transcription-initiation efficiency (18, 62). The σR3.2 hairpin loop can be crosslinked to the initiating NTP (63). The σR3.2 hairpin loop has been proposed to interact directly with the initiating NTP and thereby to facilitate de novo, unprimed initiation of RNA synthesis (18, 62). Thus, in transcription initiation, σ region 3.2 performs functions analogous to those of TFB and TFIIB N-terminal-domain CSBs.

Based on our crosslinking results (Figs. 2 and 3), and on documented roles of TFB and TFIIB N-terminal-domain CSBs in transcription-initiation NTP concentration dependence, transcription-initiation efficiency, and start-site selection (33, 37–42), we suggest that the TFB and TFIIB N-terminal domains in the archaeal and eukaryal transcription initiation complexes occupy positions analogous to those of σR3.2 in the bacterial transcription initiation complex. Specifically, we suggest that the TFB or TFIIB N-terminal-domain zinc ribbon binds in or near the RNAP RNA exit channel (as documented by protein-protein photocrosslinking and protein affinity cleaving (64)) and that the TFB and TFIIB N-terminal-domain CSBs bind in the RNAP active-center cleft, between the two strands of the melted transcription bubble, near the transcription start site. We further suggest that the TFB and TFIIB N-terminal-domain CSBs make direct interactions with the initiating NTP that facilitates de novo, unprimed initiation of RNA synthesis.

Consistent with the above proposals, in the recently determined crystallographic structure of a complex between yeast TFIIB and yeast RNAPII, the TFIIB N-terminal domain zinc ribbon is located in the RNAP RNA exit channel, and the TFIIB N-terminal domain CSB is located in the RNAP active-center cleft. 7

Acknowledgments—We thank Drs. S. Bell, D. Bushnell, and R. Kornberg for helpful discussion.

REFERENCES

1. Hampsey, M. (1998) Microbiol. Mol. Biol. Rev. 62, 465–503
2. Lee, T. I., and Young, R. A. (2000) Annu. Rev. Genet. 34, 77–137
3. Ebert, R. H. (2000) J. Mol. Biol. 304, 687–698
4. Woychik, N. A., and Hampsey, M. (2002) Cell 109, 453–463
5. Soppa, J. (1999) Mol. Microbiol. 35, 1295–1305
6. Bell, S. D., Magil, C. P., and Jackson, S. P. (2001) Biochem. Soc. Trans. 29, 392–395
7. Hauser, W., Wettach, J., Hethke, C., and Thomm, M. (1996) J. Biol. Chem. 271, 30144–30148
8. Hethke, C., Geering, A. C., Hauser, W., de Vos, W. M., and Thomm, M. (1996) Nucleic Acids Res. 24, 2369–2376
9. Hethke, C., Bergerat, A., Hauser, W., Fortherre, P., and Thomm, M. (1999) Genetics 152, 1325–1333
10. Lewis, L. M. (2000) Structural and Functional Characterization of the Transcription Preinitiation Complex from Pyrococcus furiosus. Ph.D. disserta-

6 S. Bell, personal communication.
7 D. Bushnell, K. Westover, R. Davis, and R. Kornberg, personal communication.