

*This contribution is part of the special series of Inaugural Articles by members of the National Academy of Sciences elected on April 30, 1996.*

## YY1 transcriptional initiator: Protein interactions and association with a DNA site containing unpaired strands

(transcription initiation/transcription factor YY1/transcription factor IIB)

ANNY USHEVA AND THOMAS SHENK\*

Howard Hughes Medical Institute, Department of Molecular Biology, Princeton University, Princeton, NJ 08544-1014

Contributed by Thomas Shenk, September 12, 1996

**ABSTRACT** The Ying–Yang 1 protein (YY1) DNA-binding site functions as an initiator element at which YY1, transcription factor IIB (TFIIB), and RNA polymerase II sponsor basal transcription from a supercoiled DNA template. We show that TFIIB binds to YY1, stabilizing its interaction with DNA, and YY1 contacts the large subunit of polymerase II, directing it to the initiation site. YY1 directs initiation from linear DNA containing mismatched sequences within its binding site, leading us to infer that supercoiling facilitates the separation of DNA strands and to suggest that YY1 likely remains bound to the start site as DNA strands separate during initiation. These results provide a mechanistic basis for transcriptional initiation directed by YY1 in the absence of the TATA box-binding protein.

The sequence surrounding the initiation site in many promoters can direct RNA polymerase II (pol II) to the start site in the absence of binding sites for other factors (reviewed in refs. 1–3). Sequences with this property are termed initiator elements (4). Initiator elements display considerable sequence variability (1, 2, 5), suggesting that a variety of factors might function at these sites. The auxiliary transcription factor TFIID has been shown to interact with initiator elements (6–9), and TFIIB binds to the initiator element of the adenovirus major late promoter (10, 11).

Binding sites for the Ying–Yang 1 protein (YY1) transcription factor (12–15) in the adeno-associated virus (AAV) P5 promoter (16) and the murine cytochrome *c* oxidase subunit Vb promoter (17) display the properties of initiator elements, and YY1 is essential for transcription from these sites (17, 18). In contrast to many initiator-containing promoters where TFIID is required for accurate initiation (6, 19–21), TFIID does not stimulate transcription from the AAV P5 YY1 initiator element *in vitro*. YY1, TFIIB, and RNA pol II are sufficient to mediate accurate basal transcription from the YY1 initiator when it is present on a supercoiled DNA (18). Under special conditions, subsets of the known auxiliary factors can mediate basal transcription on a variety of promoters (22–27), but each of these reactions requires the TBP subunit of TFIID, which plays a central role in transcription by all three eukaryotic RNA polymerases (2, 28). The YY1 initiator provides the only example of transcription initiation in the absence of a TATA box-binding protein (TBP).

Here we demonstrate that YY1 functions like TBP, binding to the DNA template and recruiting polymerase into an initiation complex. TFIIB and RNA pol II bind directly to YY1, as is known to be the case for TBP (29–31). TFIIB stabilizes the binding of YY1 to DNA, and YY1 contacts the carboxyl-terminal repeated domain of the large subunit of pol II, directing it to the initiation site. We also show that YY1 directs transcription from linear DNA templates containing

mismatches that form a 5 base “bubble” within the YY1 binding site. This observation indicates that YY1 potentially remains associated with the DNA as strands are separated during initiation, and it leads us to infer that supercoiling facilitates DNA strand separation at the site of YY1-mediated initiation.

### MATERIALS AND METHODS

**Assays for the Interaction of YY1 with TFIIB.** TFIIB, YY1, GST-YY1, and glutathione *S*-transferase (GST) protein were produced in *Escherichia coli* and purified as described (18). To test for protein–protein interactions in solution, TFIIB (10  $\mu$ g) and YY1 (30  $\mu$ g) were incubated alone or in combination for 30 min at 22°C in buffer A (20 mM Hepes, pH 7.9/0.15 M NaCl/5 mM dithiothreitol/5  $\mu$ M leupeptin/20% glycerol/200  $\mu$ g/ml bovine serum albumin) and subjected to gel filtration on Superose 6. YY1 and TFIIB were localized in the elution profile by protein blot using polyclonal antibodies (18). To assay protein interactions using GST fusion proteins (32), TFIIB (30  $\mu$ g) was incubated with either GST protein (30  $\mu$ g) or GST–YY1 fusion protein (10  $\mu$ g) immobilized on glutathione Sepharose for 30 min in buffer B (20 mM Tris·HCl, pH 8.0/20 mM NaCl/2 mM EDTA/2 mM dithiothreitol/5% glycerol/1 mg/ml BSA/1 mg/ml GST protein). After washing, TFIIB was eluted in buffer B containing the indicated additives. TFIIB was monitored by protein blot assay. To test for interactions in cell extracts, a HeLa nuclear extract was fractionated on phosphocellulose P11, and the 0.5 M KCl eluate containing TFIIB and YY1 was dialyzed in buffer A, applied to a single-stranded DNA cellulose matrix, and eluted in buffer A with a linear gradient of 0.15–0.7 M NaCl. The fractions containing both TFIIB and YY1 were incubated with purified YY1-specific polyclonal IgG bound to Affi-gel 102 ( $\approx$ 1.1 mg IgG per packed ml of beads in buffer A with 0.6 M NaCl and 1 mg/ml BSA) for 1 hr at 30°C; the beads were pelleted and washed in buffer A, and then the proteins were released by boiling in buffer with SDS and detected by protein blot assay.

The interaction between TFIIB and YY1 on DNA containing the YY1 binding site that serves as an initiator element in the AAV P5 promoter was also monitored. Band-shift assays received oligonucleotide (1.2 nM), YY1 (10  $\mu$ M), TBP (10  $\mu$ M), and TFIIB (50  $\mu$ M). DNA–protein complexes were analyzed by electrophoresis in either standard native poly-

Abbreviations: pol II, polymerase II; TBP, TATA box-binding protein; GST, glutathione *S*-transferase; CTD, carboxyl-terminal domain; AAV, adeno-associated virus.

\*To whom reprint requests should be addressed. e-mail: tshenk@molbiol.princeton.edu.

acrylamide gels or in native agarose gels. Agarose gel electrophoresis was performed at 4°C in a 2% gel with buffer containing 10 mM Tris·HCl (pH 8.5) and 90 mM glycine. DNA-containing bands were visualized by staining with ethidium bromide after electrophoresis.

Band-shift assays were used to determine under nonequilibrium conditions the observed binding constant ( $K_{obs}$ ) of YY1 and the TFIIB–YY1 complex for the YY1 binding site at the initiator element of the AAV P5 promoter. A constant amount of labeled 22-bp DNA oligonucleotide (2.5 pmol/15  $\mu$ l reaction) and poly [d(I-C)] (75  $\mu$ g/ml) was incubated with YY1 alone or a TFIIB–YY1 mixture. The YY1 concentration was varied, and when TFIIB was included it was added at four times the molar concentration of YY1. With DNA in excess,  $\approx$ 95% of the recombinant YY1 could bind to a YY1 recognition motif, while with DNA and YY1 in excess  $\approx$ 70% of the recombinant TFIIB could enter into a TFIIB–YY1–DNA complex (data not shown). DNA–protein mixtures were incubated for 15 min at 30°C to allow binding, and the resulting complexes were resolved by electrophoresis in a native polyacrylamide gel. The amount of free DNA was determined by scanning the gels on a PhosphorImager (Molecular Dynamics), and  $K_{obs}$  was the concentration of YY1 that reduced by one-half the amount of unbound DNA, taking the proportions of active reactants into account (33).

**Assays for the Interaction of YY1 with RNA pol II.** GST or GST–YY1 proteins were incubated with either purified pol II A, pol IIA/0, or pol IIB (1  $\mu$ g) or a carboxyl-terminal domain (CTD) polypeptide (100  $\mu$ g/ml) composed of amino acids encoded by base pairs 5040–5510 (34) of the large subunit of pol II. To prepare TBP–CTD complexes immobilized on glutathione-Sepharose beads, 600 ng TBP was mixed with 200 ng GST–CTD (this fusion protein contained the same CTD sequences as described above for the CTD polypeptide). To test for interactions in cell extracts, RNA pol II was partially purified from a HeLa cell nuclear extract by sequential chromatography on phosphocellulose P11, DEAE–Sepharose, and phenyl-Sepharose. The partially purified material was subjected to immunoaffinity purification by incubating with YY1-specific IgG bound to beads. The large subunit of the polymerase was monitored by protein blot assay using an antibody specific for the CTD or the amino acid sequence encoded by exon 5 of the mRNA encoding the large subunit (gift of R. Weinmann, Bristol-Myers Squibb Pharmaceutical Research Institute).

**In Vitro Transcription.** *In vitro* transcription reactions were performed as described (18). In some reactions, the template was a supercoiled DNA containing either the AAV P5 promoter or adenovirus ML promoter (100 fmol per 30  $\mu$ l reaction), and in other reactions the template was an 86-bp oligonucleotide (100 fmol per 30  $\mu$ l reaction) designed so that the transcription initiation sites were positioned 25–26 bp from the nearer end. RNA pol II (1 ng per 30  $\mu$ l reaction) was purified by the procedure of Lu *et al.* (35) with modifications (18), and TFIIB (10 ng/30  $\mu$ l reaction) and YY1 (4 ng/30  $\mu$ l reaction) were purified recombinant proteins (18). Transcription was assayed by primer extension using avian myeloblastosis virus reverse transcriptase after digestion of the DNA template with RNase-free DNase I.

## RESULTS

**TFIIB Binds to YY1.** The ability of TFIIB to bind to YY1 was tested by using purified proteins produced in *E. coli*. The formation of a TFIIB–YY1 complex in solution was assayed by gel filtration chromatography (Fig. 1A). Before mixing, TFIIB and YY1 were eluted from the column as expected for their molecular weights. When the two proteins were mixed before chromatography, a substantial portion of TFIIB and YY1 coeluted as a larger complex. The formation of a TFIIB–YY1

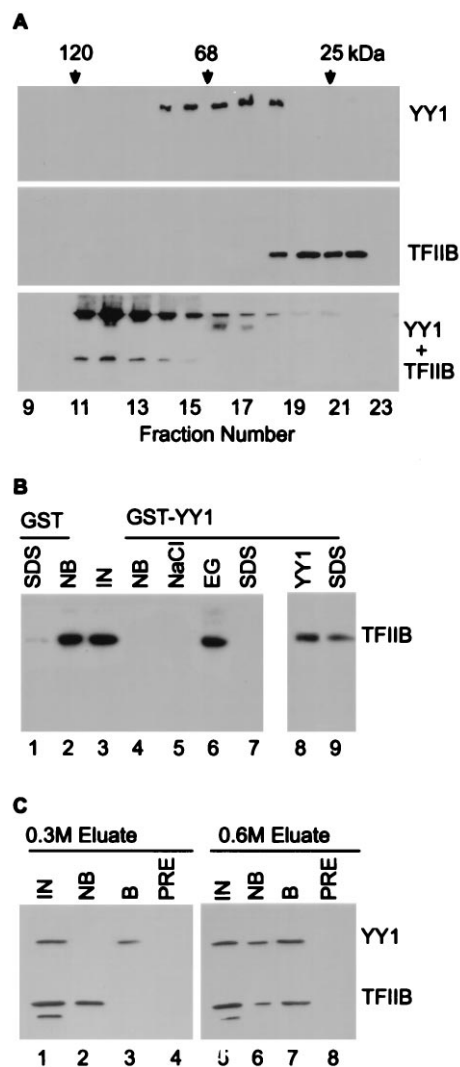


FIG. 1. TFIIB binds to YY1. (A) TFIIB binds to YY1 in solution. Purified TFIIB and YY1 were subjected to gel filtration chromatography alone or after mixing, and localized in the elution profile by protein blot assay using polyclonal antibody (18) to YY1 (Top), TFIIB (Middle), or a mixture of the antibodies (Bottom). Chymotrypsinogen A (25 kDa), BSA (68 kDa), and  $\beta$ -galactosidase (120 kDa) were markers. (B) TFIIB binds to GST–YY1. TFIIB was monitored by protein blot assay. Lanes: IN, TFIIB input to each reaction; NB, TFIIB not bound; NaCl, TFIIB eluted in buffer containing 1.0 M NaCl; EG, TFIIB eluted in buffer containing 50% ethylene glycol plus 100 mM NaCl; SDS, TFIIB eluted by boiling in buffer with detergent, either immediately after application to the affinity matrix (lanes 1 and 9) or following elution with ethylene glycol (lane 7); YY1, TFIIB eluted in buffer containing 60  $\mu$ g/ml YY1. (C) Copurification of TFIIB and YY1 from a HeLa cell nuclear extract. Two fractions from a single-stranded DNA cellulose column (0.3 M and 0.6 M NaCl) containing both YY1 and TFIIB were assayed on a YY1-specific IgG matrix. TFIIB and YY1 were detected by protein blot assay. Lanes: IN, TFIIB and YY1 input to each reaction; NB, input which is not bound; B, input which is bound; PRE, input bound by preimmune IgG. Figures were produced using PHOTOSHOP and FREEHAND software.

complex was also evaluated using a GST–YY1 fusion protein bound to glutathione beads; TFIIB, after capture or elution, was assayed by Western blot (Fig. 1B). When TFIIB was applied to beads containing GST protein without a fusion partner,  $>$ 97% of the factor remained in the unbound fraction (Fig. 1B, lanes 1–3). In contrast, GST–YY1 protein captured  $>$ 99% of the input TFIIB; the captured protein was not released by washing with buffer containing 1 M NaCl, but it was quantitatively released by washing with buffer containing

0.1 M NaCl and 50% ethylene glycol, leaving no protein for release upon boiling in buffer containing SDS (Fig. 1B, lanes 3–7). Ethylene glycol weakens hydrophobic interactions (36), so it is likely that hydrophobic contacts occur at the interface of the two proteins. Soluble YY1 could partially compete for interaction with TFIIB and release a portion of the factor from the GST–YY1 matrix (Fig. 1B, lanes 8 and 9), arguing that the TFIIB–YY1 interaction is specific and dynamic.

To further investigate the TFIIB–YY1 interaction, we asked if the two proteins could be copurified from a HeLa nuclear extract. Sequential fractionation on phosphocellulose and single-stranded DNA cellulose produced two fractions containing both TFIIB and YY1. YY1 was captured from each of these two fractions using a YY1-specific antibody, and the immunoaffinity purified material was assayed for the presence of TFIIB (Fig. 1C). TFIIB in the 0.3 M NaCl fraction did not copurify with YY1, while TFIIB in the 0.6 M fraction was captured by the YY1-specific antibody together with YY1. It is not clear why YY1 and TFIIB are present in a complex in one fraction and not in the other. Perhaps additional proteins in the 0.3 M NaCl fraction bind to YY1 or TFIIB and preclude their interaction. Nevertheless, we can conclude that a portion of the TFIIB and YY1 present in a HeLa cell can be copurified as a complex from a nuclear extract, consistent with their ability to interact *in vitro*.

The functional consequence of the YY1–TFIIB interaction was explored by performing band-shift assays with purified proteins (Fig. 2). TFIIB substantially increased the efficiency with which YY1 bound to its DNA recognition site (Fig. 2A Top, lanes 1 and 2). However, the electrophoretic migration of the shifted complex did not change in the presence of TFIIB. TFIIB also increased the efficiency with which TBP bound to a TATA motif (Fig. 2A, lanes 7 and 9), but in this case the addition of TFIIB retarded the migration of the protein–DNA complex. To ascertain whether TFIIB and YY1 were both present in a complex with DNA, the complexes were transferred from the band-shift gel to a membrane and the proteins were assayed by Western blot. When TFIIB was included, it was present in the shifted complex (Fig. 2A Bottom, lane 1); and inclusion of TFIIB increased the amount of YY1 present in the shifted complex (Fig. 2A Middle, lanes 1 and 2). TFIIB alone did not interact with the oligonucleotide containing a YY1-binding site (Fig. 2A Middle, lane 6), and no YY1 or TFIIB was detected in the blots if DNA containing a wild-type YY1-binding site was not included (Fig. 2A Middle and Bottom, lanes 3–5). This is expected since both proteins are basic and migrate toward the cathode in the assay.

TFIIB altered the migration of a YY1–DNA complex when electrophoresis was performed in an agarose gel (Fig. 2B, lanes 2 and 3) rather than in the conventional polyacrylamide gel. This experiment also indicated that the inclusion of TFIIB led to an increase in the amount of DNA complexed with protein since, in contrast to the reaction receiving YY1 alone where only a portion of the DNA was bound, very little of the probe migrated at the position corresponding to free oligonucleotide in the presence of both TFIIB and YY1. However, the TFIIB–YY1–DNA complex was apparently unstable during electrophoresis in the agarose gel because the amount of complex formed in the presence of TFIIB was not substantially greater than the amount of complex produced in the presence of YY1 alone. Control experiments showed that TFIIB alone did not induce a shifted band and that TFIIB plus YY1 did not interact with an oligonucleotide containing a mutated YY1 binding site (Fig. 2B, lanes 4 and 5). To confirm that the DNA–protein complex formed in the presence of TFIIB and YY1 contained both factors, shifted bands were excised from an agarose gel and their protein constituents were assayed by electrophoresis in an SDS-containing polyacrylamide gel (Fig. 2C). The complex formed in the presence of YY1 and DNA contained YY1; and the complex formed in the presence of

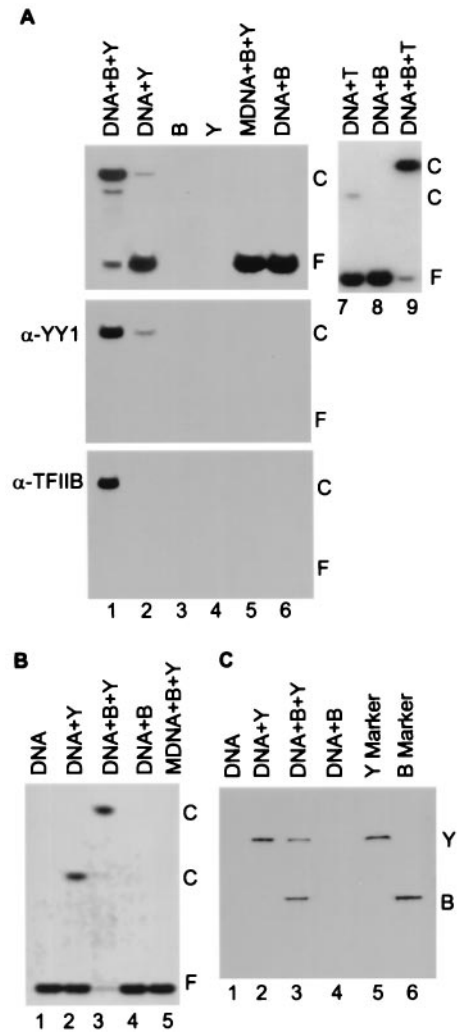


FIG. 2. TFIIB forms a complex on DNA with YY1. (A) Band-shift assay of YY1–DNA interactions in a polyacrylamide gel. The components added to band-shift assays are indicated at the top of the autoradiograms. Lanes: B, TFIIB; Y, YY1; T, TBP; DNA, a 22-bp oligonucleotide with the YY1 binding site present at +1 in the AAV P5 promoter or a 20-bp oligonucleotide with the TATA motif present in the adenovirus major late promoter; and MDNA, a 22-bp derivative of the YY1-binding site to which YY1 cannot bind (14). (Top) Band-shift assay displays an autoradiogram where  $^{32}$ P-labeled probe DNA is detected. (Middle and Bottom) Protein blots using antibody to YY1 ( $\alpha$ -YY1) or TFIIB ( $\alpha$ -TFIIB). Reactions for band-shift assays in which proteins were transferred to membranes received 25-fold more of each protein than was used in standard band-shift assays. The positions of free DNA, F, and protein–DNA complexes, C, are indicated. (B) Band-shift assay in an agarose gel showing that TFIIB, YY1 and DNA form a complex. Components added to assays are as in A. The YY1–TFIIB–DNA complex was unstable during analysis in agarose, as evidenced by the substantial reduction in the amount of free DNA without a concomitant increase in the band corresponding to the complex (lane 3). (C) Analysis of proteins present in shifted complexes. Complexes present in the agarose gel displayed in B were excised, and their protein constituents were analyzed by electrophoresis in an SDS-containing polyacrylamide gel. Purified YY1 (Y) and TFIIB (B) were included as markers.

TFIIB, YY1, and DNA contained both proteins in approximately equimolar amounts (Fig. 2C, lanes 2 and 3). A control experiment demonstrated that TFIIB in the absence of YY1 did not migrate in the agarose gel at the position of the TFIIB–YY1–DNA complex (Fig. 2C, lane 4).

The substantial increase in the amount of DNA–YY1 complex formed in the presence of TFIIB (Fig. 2A) suggested that TFIIB might increase the affinity of YY1 for DNA. Accord-

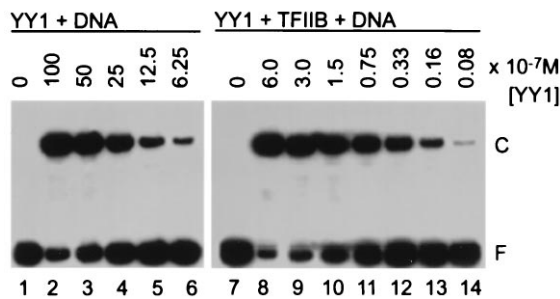


FIG. 3. TFIIB increases the apparent affinity of YY1 for its DNA-binding site. The amount of YY1 added to each band-shift reaction is recorded above the lanes, and the positions of free DNA (F) and protein-DNA complexes (C) are indicated.

ingly, we measured the dissociation constant for the YY1-DNA interaction in the presence and absence of TFIIB (Fig. 3). Complexes and free DNA were quantified by band-shift assay as a function of YY1 concentration (33), and the observed dissociation constant,  $K_{obs}$ , was  $4 \times 10^{-6}$  M for the YY1-DNA complex and  $1 \times 10^{-7}$  M for the TFIIB-YY1-DNA complex. The absolute  $K_{obs}$  values are not reliable since the experiment was performed under nonequilibrium conditions. However, the relative difference in  $K_{obs}$  seen for the interactions is instructive. TFIIB enhanced the apparent affinity of YY1 for its recognition site by a factor of 40.

To further evaluate the TFIIB-YY1 interaction with DNA, we examined a deleted derivative containing amino acids 113-316 of TFIIB. This truncated TFIIB interacts with the TBP-DNA complex and the structure of this complex has been determined (37). Band-shift analysis revealed that the TFIIB variant also interacted with the YY1-DNA complex, stabilizing its interaction with DNA (data not shown).

**RNA pol II Binds to YY1.** The repeated CTD of the large subunit of RNA pol II is known to bind to TBP (29, 30), and this interaction is thought to help direct the polymerase to the promoter. Two forms of the polymerase can be isolated that differ by the extent to which the CTD is phosphorylated, and TBP binds preferentially to the hypophosphorylated form of the polymerase, pol IIA (30).

We asked whether pol IIA might be directed to the promoter through an interaction with YY1 in the absence of TBP. The assay utilized GST-YY1 fusion protein bound to glutathione beads, and pol IIA was monitored by protein blot using an antibody specific for its large subunit. Pol IIA was quantitatively bound to the GST-YY1 matrix; it was not detectably released by washing in buffer containing 1.0 M NaCl, but it was released into buffer containing 0.1 M NaCl and 50% ethylene glycol, leaving no detectable protein to be eluted by boiling in buffer with SDS (Fig. 4A, lanes 1-5). Soluble CTD polypeptide could partially compete for interaction with GST-YY1, releasing pol IIA from the GST-YY1 matrix, and the bound CTD was subsequently eluted with ethylene glycol (Fig. 4A, lanes 6-8). The CTD competition experiment indicated that the interaction of YY1 with pol IIA likely occurred through the CTD of its large subunit. Capture experiments were also performed with a mixture of pol IIA and pol IIO, the hyperphosphorylated form of the polymerase (Fig. 4A, lanes 9-11); and with pol IIB, a form of polymerase lacking the CTD (Fig. 4A, lanes 12-14). The polymerase was captured by GST-YY1 whether or not the CTD of the large subunit was hyperphosphorylated (pol IIO was bound), but the interaction did not occur in the absence of a CTD repeat (pol IIB did not bind). The GST-YY1 protein also captured a CTD polypeptide (Fig. 4B). The bound protein was not released by 1.0 M NaCl, it was quantitatively released by buffer containing 0.1 M NaCl and 50% ethylene glycol, and it was partially released by washing in buffer containing soluble YY1, leaving a portion on

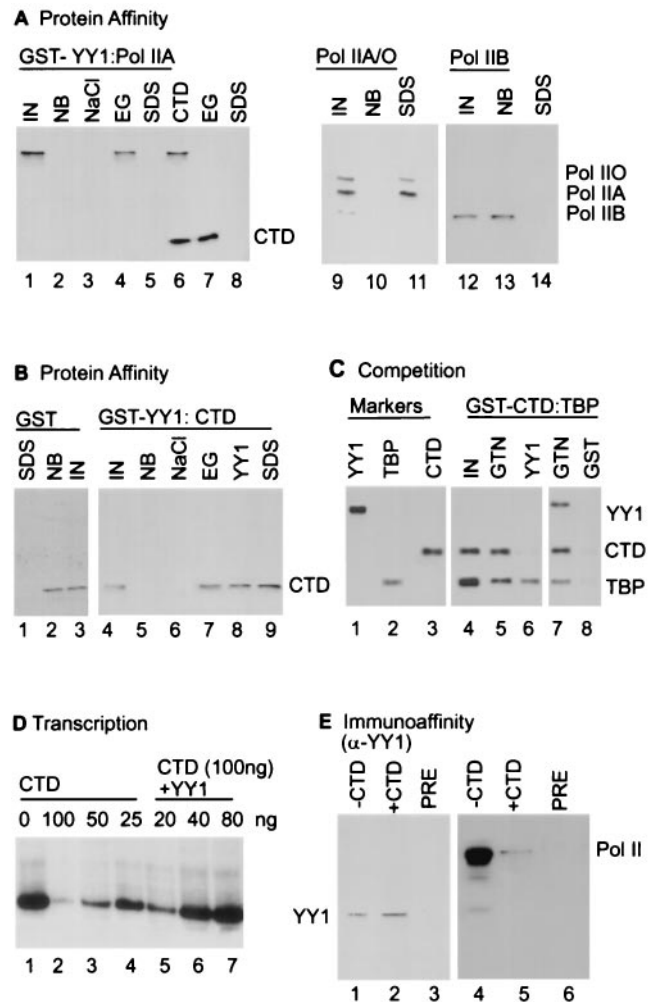


FIG. 4. RNA pol II interacts with YY1. (A) The large subunit of RNA pol II binds to GST-YY1. RNA pol IIA (pol IIA) mixed RNA pol IIA and pol IIO (pol IIA/O) and RNA pol IIB (pol IIB) were tested for interaction with GST-YY1. Polymerase was monitored by protein blot assay using antibodies specific for the CTD (lanes 1-8) or the amino acid sequence encoded by exon 5 of the mRNA encoding the large subunit (lanes 9-14). Lanes: IN, polymerase input; NB, polymerase not bound; NaCl, polymerase eluted in buffer containing 1.0 M NaCl; EG, polymerase eluted in buffer containing 50% ethylene glycol plus 100 mM NaCl; CTD, polymerase eluted by inclusion of a CTD polypeptide (100  $\mu$ g/ml) in the washing buffer; SDS, polymerase eluted by boiling in buffer with detergent, after elution with ethylene glycol (lanes 5 and 8) or immediately after loading onto GST-YY1 matrix (lanes 11 and 14). (B) The CTD binds to GST-YY1. The CTD polypeptide was detected by protein blot, and the designations for lanes are as for A. (C) YY1 competes with TBP for interaction with the CTD. TBP was bound to GST-CTD fusion protein on beads, complexes on beads were washed with binding buffer, TBP was displaced from the GST-CTD-TBP complex with binding buffer containing soluble YY1 (100  $\mu$ l, 4 mg/ml), and reactants were assayed by protein blot using a mixture of antibodies to YY1, TBP, and CTD (lanes 1-3, 7, and 8) or TPB and CTD (lanes 4-6). YY1-specific antibody was excluded from the membrane including lane 6 because it contained a huge quantity of YY1. Lanes: IN, CTD-TBP complex input; GTN, elution with glutathionine; YY1, elution with YY1; GST, elution with GST protein without a fusion partner. (D) CTD inhibits transcription directed by the AAV P5 promoter in the presence of YY1, TFIIB, and pol IIA. Reactions received the indicated amounts of CTD polypeptide and GST-YY1 fusion protein. The 79-nt primer extension products are labeled. (E) Copurification of pol II and YY1 from a HeLa nuclear extract. Partially purified polymerase was incubated with YY1-specific antibody on beads in the presence or absence of excess CTD polypeptide,  $\pm$ CTD. Material bound to YY1-specific antibody,  $\alpha$ -YY1, was eluted and assayed by protein blot using either an antibody to YY1 (lanes 1 and 2) or an antibody to the polymerase (lanes 3 and 4).

the matrix to be eluted by boiling in buffer containing SDS (Fig. 4B, lanes 4–9). GST protein without a fusion partner did not bind to the CTD (Fig. 4B, lanes 1–3). This experiment demonstrated that YY1 interacts directly with the CTD, since no other polymerase components were present. TBP bound to the GST–CTD fusion protein was partially displaced by soluble YY1 (Fig. 4C), reinforcing the conclusion that both TBP and YY1 interact within the CTD. The CTD polypeptide inhibited a reconstituted YY1-dependent transcription reaction and excess YY1 relieved the inhibition (Fig. 4D), consistent with a role for a YY1–CTD interaction in initiation.

A portion of YY1 and pol II exist as a complex in HeLa cell extracts. Partially purified polymerase was subjected to immunoaffinity chromatography using an antibody to YY1, and the proteins that bound to the matrix were assayed by protein blot. YY1 and the large subunit of the polymerase were both captured by the YY1-specific antibody (Fig. 4E). To confirm that the polymerase was captured indirectly due to its interaction with YY1, CTD polypeptide was added to the polymerase preparation before it was applied to the antibody matrix. The CTD competed with the polymerase for YY1 binding, and much less polymerase was captured by the YY1-specific antibody in its presence (Fig. 4E, compare lanes 4 and 5).

**Supercoiling Mediates DNA Strand Separation.** Houbaviy *et al.* (38) have determined the three dimensional structure of the YY1 zinc-finger domain bound to the initiator DNA sequence. While there are multiple contacts by YY1 on both DNA strands upstream of the start sites and on the template strand downstream of the two adjacent start sites, there is only one contact between the protein and DNA downstream of the start sites on the coding strand of DNA (Fig. 5A, oligonucleotide a, YY1 contacts are represented by dots and arrows). The dearth of YY1 contacts on the coding strand downstream of the start sites suggested that YY1 might bind to an artificial DNA template in which nucleotides in this region are substituted. We synthesized a “bubble” template (Fig. 5A, oligonucleotide b) with 5 nt substituted on the coding strand, and band-shift analysis indicated that YY1 was able to bind to oligonucleotide b, although with reduced efficiency as compared with the wild-type site on oligonucleotide a (Fig. 5B, compare lanes 3 and 4). In contrast, an oligonucleotide with a 5-nt substitution at the corresponding position on the template strand (Fig. 5A, oligonucleotide c) was not detectably bound by YY1 (Fig. 5B, lane 2), consistent with the multiple YY1 contacts on this strand.

The ability of YY1 to bind to oligonucleotide b provided the opportunity to ask if the bubble resulting from the mismatched sequence in the initiator sequence would substitute for supercoiling and permit YY1-dependent transcription from a linear template. When transcription was assayed by primer extension at nucleotide sequence resolution, oligonucleotide b generated transcripts with 5' ends identical to the ends generated on a supercoiled template (Fig. 5C, compare lanes 2 and 5). As expected, oligonucleotides a and c were not transcribed, and oligonucleotide b did not yield specific transcripts in a reaction lacking YY1 (Fig. 5C, lanes 4, 6, and 7). Oligonucleotides b and c both generated shorter products whose 5' ends correspond to an end of the bubble. These products are not dependent on YY1 (Fig. 5C, lane 7), and have been observed when bubble-containing templates are transcribed by *E. coli* RNA polymerase (39, 40). YY1-dependent transcription from oligonucleotide b, but not a or c, was confirmed by a run-off assay (Fig. 5D). Again, two shorter products are generated in reactions programmed with oligonucleotides b and c, corresponding in length to transcripts extending from the edges of the bubble to either end of the template. Taken together, the data show that a bubble activates YY1-dependent transcription of a linear DNA. This implies that supercoiling facilitates strand separation at the site of initiation.

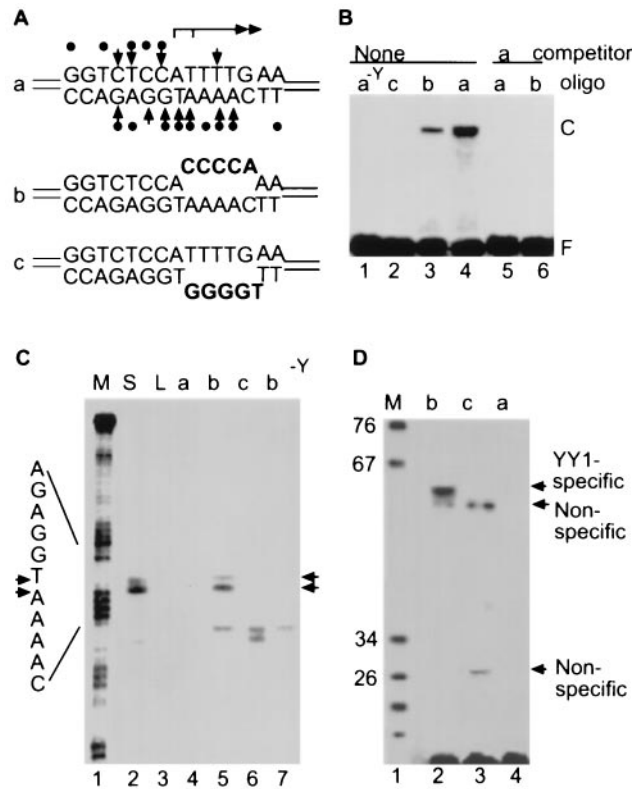


FIG. 5. YY1-directed transcription from linear DNA containing a mismatched sequence at the initiation site. (A) Sequences of oligonucleotides with normal and altered YY1 binding sites. Interactions between YY1 and its recognition site (37) are indicated for oligonucleotide a. (B) Band-shift assay comparing the ability of YY1 to interact with its normal recognition site and modified sites containing mismatched sequences. <sup>32</sup>P-labeled and unlabeled competitor oligonucleotides are identified at the top of the lanes. Lanes: a<sup>-Y</sup>, received oligonucleotide but no YY1; F, free DNA; and C, YY1–DNA complexes. (C) *In vitro* transcription assayed by primer extension. Transcription reactions received as template supercoiled (S) or linear (L) plasmid DNA with the AAV P5 promoter, the indicated oligonucleotide, or oligonucleotide b in a reaction lacking YY1 (b<sup>-Y</sup>). Markers (M) were generated by chemical sequencing using the A > C reaction of oligonucleotide a; initiation sites and bands corresponding to YY1-dependent initiation events are designated with arrows. (D) *In vitro* transcription monitored by run off assay. Reactions received oligonucleotide a, b, or c as template. The sizes of marker DNAs are indicated in bases; specific transcripts originating from the YY1-directed initiation sites and nonspecific transcripts from the ends of the bubbles are labeled.

DISCUSSION

Our results, considered in light of the three-dimensional structure (38) of the YY1 zinc-finger domain bound to the initiator sequence, suggest a model for YY1-dependent initiation on supercoiled DNA. YY1 binds to the initiator, and the YY1–DNA interaction is stabilized by TFIIB (Figs. 1–3). This sets the start site for transcription. The asymmetric arrangement of YY1–DNA contacts (38) very likely mandates unidirectional transcription since the YY1–DNA structure predicts that DNA strands can separate on only one side of a bound YY1 molecule. RNA pol II joins the complex by interacting with YY1 (Fig. 4) and TFIIB (31). Since a bubble activates YY1-dependent transcription of a linear DNA (Fig. 5), we can infer that supercoiling likely provides the energy to separate DNA strands at the site of initiation. Supercoiling also might facilitate extension of the open complex during promoter clearance by the polymerase, as is the case for other promoters in the absence of TFIIE and TFIIH (24), but our assay for YY1-dependent transcription directed by a bubble-containing

template does not discriminate the initial strand separation event from promoter clearance. Finally, the ability of YY1 to bind and function on a DNA containing a 5-nt mismatch at the start site (Fig. 5) suggests that it could potentially remain associated with the template throughout the initiation process, indeed through multiple rounds of initiation, tolerating the partial separation of DNA strands within its binding site.

In this model, YY1 displays striking functional similarities to TBP. Like TBP, YY1 provides a site-specific DNA recognition function, selecting the transcription initiation site (18); and YY1 has the potential to remain bound at the promoter through multiple rounds of initiation, as is known to be the case for TBP (41). YY1 binds to TFIIB, as does TBP (31). TFIIB increases the affinity of YY1 for its binding site, and the same is true for TBP (42). In the case of TBP, the increased affinity probably results from the interaction of TFIIB with both TBP and DNA (37, 43). Additional studies are needed to determine whether a similar mechanism leads to the stabilization of the YY1-DNA interaction. Finally, YY1 binds to the CTD of the large subunit of pol II, as does TBP (29, 30). The details of the interaction of TBP and YY1 with the CTD differ, however, in that the TBP-CTD interaction is sensitive to the ionic strength of the buffer while the YY1-CTD interaction is hydrophobic in nature (Fig. 4 A and B). Also, whereas TBP exhibits a preference for binding to the hypophosphorylated form of the polymerase (30), our assays have not yet revealed an effect of CTD phosphorylation on YY1 binding (Fig. 4A). Phosphorylation of the CTD might modulate the efficiency with which the polymerase enters a preinitiation complex containing TBP (44), but conversion of pol IIA to the highly phosphorylated pol IIO form is not essential for initiation complex assembly or elongation in reconstituted, TBP-dependent basal transcription reactions (45). Further, ATP hydrolysis is not required for initiation complex formation or the production of a short transcript from the adenovirus major late promoter on a supercoiled template DNA (24, 25). Thus, it is not yet clear whether the difference in binding preference exhibited by TBP and YY1 for the hyperphosphorylated form of the CTD is functionally significant.

Although we can now put forward a reasonable model to explain the ability of YY1 to sponsor transcription in cooperation with TFIIB and RNA pol II *in vitro*, we have not yet shown that the TFIID-independent reaction occurs within the cell. TFIID might be required for the activity of the YY1 initiator element or stimulate its activity *in vivo*. It is not possible to resolve the question of a role for TFIID by mutating the template since TFIID can function at TATA-less promoters, presumably through a TAF-mediated interaction at the initiator sequence (7-9) or by entering the initiation complex through protein-protein interactions and interacting through TBP with non-TATA DNA sequences (6, 19, 21, 46). Indeed, we have found that YY1 can bind to TBP *in vitro* (18), suggesting that TFIID might reside at a YY1 initiator in the absence of an upstream TATA motif. Experiments are in progress to determine whether TFIID plays a role in transcriptional initiation at the AAV P5 promoter within cells.

We thank S. Burley, M. Dobbstein, A. Hirsch, N. Horikoshi, H. Houbaviy, and R. Perry for discussions and critical reading of this manuscript. This work was supported by a grant from the National Cancer Institute. T.S. is an American Cancer Society Professor and an Investigator of the Howard Hughes Medical Institute.

1. Weis, L. & Reinberg, D. (1992) *FASEB J.* **6**, 3300-3309.
2. Hernandez, N. (1993) *Genes Dev.* **7**, 1291-1308.
3. Gill, G. (1994) *Curr. Biol.* **4**, 374-376.
4. Smale, S. T. & Baltimore, D. (1989) *Cell* **57**, 103-113.
5. Javahery, R., Khachi, A., Lo, K., Zenzie-Gregory, B. & Smale, S. T. (1994) *Mol. Cell. Biol.* **14**, 116-127.
6. Zhou, Q., Lieberman, P. M., Boyer, T. G. & Berk, A. J. (1992) *Genes Dev.* **6**, 1964-1974.
7. Purnell, B. A. & Gilmour, D. S. (1993) *Mol. Cell. Biol.* **13**, 2593-2603.
8. Kaufmann, J. & Smale, S. T. (1994) *Genes Dev.* **8**, 821-829.
9. Verrijzer, C. P., Yokomori, K., Chen, J.-L. & Tjian, R. (1994) *Science* **264**, 933-941.
10. Roy, A. L., Meisterernst, M., Pognonec, P. & Roeder, R. G. (1991) *Nature (London)* **354**, 245-248.
11. Roy, A. L., Malik, S., Meisterernst, M. & Roeder, R. G. (1993) *Nature (London)* **365**, 355-359.
12. Hariharan, N., Kelley, D. E. & Perry, R. P. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 9799-9803.
13. Park, K. & Atchison, M. L. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 9804-9808.
14. Shi, Y., Seto, E., Chang, L.-S. & Shenk, T. (1991) *Cell* **67**, 377-388.
15. Flanagan, J. B., Becker, K. G., Ennist, D. L., Gleason, S. L., Driggers, P. H., Levi, B.-Z., Appella, E. & Ozato, K. (1992) *Mol. Cell. Biol.* **12**, 38-44.
16. Seto, E., Shi, Y. & Shenk, T. (1991) *Nature (London)* **354**, 241-245.
17. Basu, A., Park, K., Atchison, M. L., Carter, R. S. & Avadhani, N. G. (1993) *J. Biol. Chem.* **268**, 4188-4196.
18. Usheva, A. & Shenk, T. (1994) *Cell* **76**, 1115-1121.
19. Smale, S. T., Schmidt, M. C., Berk, A. J. & Baltimore, D. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 4509-4513.
20. Carcamo, J., Buckbinder, L. & Reinberg, D. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 8052-8056.
21. Pugh, B. F. & Tjian, R. (1991) *Genes Dev.* **5**, 1935-1945.
22. Parvin, J. D., Timmers, H. T. M. & Sharp, P. A. (1992) *Cell* **68**, 1135-1144.
23. Parvin, J. D. & Sharp, P. A. (1993) *Cell* **73**, 533-540.
24. Goodrich, J. A. & Tjian, R. (1994) *Cell* **77**, 145-156.
25. Parvin, J. D., Shikind, B. M., Meyers, R. E., Kim, J. & Sharp, P. A. (1994) *J. Biol. Chem.* **269**, 18414-18421.
26. Timmers, H. T. (1994) *EMBO J.* **13**, 391-399.
27. Tyree, C. M., George, C. P., Lira-DeVito, L. M., Wampler, S. L., Dahmus, M. E., Zawel, L. & Kadonaga, J. T. (1994) *Genes Dev.* **7**, 1254-1265.
28. Rigby, P. W. J. (1993) *Cell* **72**, 7-10.
29. Koleske, J. A., Buratowski, S., Nonet, M. & Young, A. R. (1992) *Cell* **69**, 883-894.
30. Usheva, A., Maldonado, E., Goldring, A., Lu, H., Houbaviy, C., Reinberg, D. & Aloni, Y. (1992) *Cell* **69**, 871-881.
31. Ha, I., Roberts, S., Maldonado, E., Sun, X., Green, M. & Reinberg, D. (1993) *Genes Dev.* **7**, 1021-1032.
32. Kaelin, W. G., Pallas, D. C., DeCaprio, J. A., Kaye, F. J. & Livingston, D. M. (1991) *Cell* **64**, 521-532.
33. Carey, J. (1991) *Methods Enzymol.* **208**, 103-117.
34. Allison, L. A., Moyle, M., Shales, M. & Ingles, C. J. (1985) *Cell* **42**, 599-610.
35. Lu, H., Flores, O. & Reinberg, D. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 10004-10008.
36. Thompson, N. E., Aronson, D. B. & Burgess, R. R. (1990) *J. Biol. Chem.* **265**, 7069-7077.
37. Nikolov, D. B., Chen, H., Halay, E. D., Usheva, A. A., Hisatake, K., Lee, D. K., Roeder, R. G. & Burley, S. K. (1995) *Nature (London)* **377**, 119-128.
38. Houbaviy, H., Usheva, A., Shenk, T. & Burley, S. K. *Proc. Natl. Acad. Sci. USA* **93**, 13577-13582.
39. Tripatara, A. & deHaseth, P. L. (1993) *J. Mol. Biol.* **233**, 349-358.
40. Aiyar, S. E., Helmann, J. D. & deHaseth, P. L. (1994) *J. Biol. Chem.* **269**, 13179-13184.
41. Zawel, L., Kumar, K. P. & Reinberg, D. (1995) *Genes Dev.* **9**, 1479-1490.
42. Imbalzano, A. N., Zaret, K. S. & Kingston, R. E. (1994) *J. Biol. Chem.* **269**, 8280-8286.
43. Malik, S., Lee, D. & Roeder, R. G. (1993) *Mol. Cell. Biol.* **13**, 6253-6259.
44. Payne, J. M., Laybourn, P. J. & Dahmus, M. E. (1989) *J. Biol. Chem.* **264**, 19621-19629.
45. Serizawa, H., Conaway, J. W. & Conaway, R. C. (1993) *Nature (London)* **363**, 371-374.
46. Zenzie-Gregory, B., Khachi, A., Garraway, I. P. & Smale, S. T. (1993) *Mol. Cell. Biol.* **13**, 3841-3849.