

Delineation of two functional regions of transcription factor TFIIB

(protease-resistant domain/protein-protein interactions/transcription complex/RNA polymerase recruitment/GAL4-VP16 chimeric activator)

ALCIDE BARBERIS*, CHRISTOPH W. MÜLLER*, STEPHEN C. HARRISON*†, AND MARK PTASHNE*

*Department of Biochemistry and Molecular Biology, and †Howard Hughes Medical Institute, Harvard University, Cambridge, MA 02138

Contributed by Mark Ptashne, March 1, 1993

ABSTRACT Human transcription factor TFIIB, a protein of 316 amino acids, was subjected to limited proteolysis in order to define stable structural domains. We find that the C-terminal region of TFIIB, residues 106–316, is relatively stable, while the N-terminal region is very sensitive to proteases. Like full-length TFIIB, the stable domain, which we refer to as TFIIBc, interacts with the TATA-binding protein (TBP) on DNA. However, TFIIBc is unable to substitute for TFIIB in an *in vitro* transcription assay. We show by gel mobility-shift experiments that TFIIBc arrests formation of the transcription complex after binding to TBP, and we conclude that the N-terminal region of TFIIB, which is missing from TFIIBc, is responsible for the recruitment of RNA polymerase II to the promoter. We also show that TFIIBc inhibits transcription by competing with full-length TFIIB for the interaction with TBP, either in the presence or in the absence of the TBP-associated factors. The acidic transcriptional activator GAL4-VP16 does not favor the assembly of the functional transcription complex over the nonfunctional complex containing TFIIBc. Thus, if the function of GAL4-VP16 is enhancement of the interaction between TFIIB and the TFIID-DNA complex, then this function can also be exerted on the protease-resistant domain TFIIBc.

Several proteins, called general transcription factors, are required in addition to RNA polymerase II for accurate transcription initiation. They include the TATA-binding factor TFIID, as well as TFIIA, TFIIB, TFIIE, and TFIIF (1–3). The general transcription factors are believed to assemble in an ordered fashion on a promoter along with RNA polymerase II to form a preinitiation complex. The first step in the assembly of this complex is the binding of TFIID to the TATA box, a process facilitated by TFIIA. Subsequent binding of TFIIB to this DNA-protein complex is necessary for the recruitment of RNA polymerase II to the promoter. This recruitment requires in turn the presence of RAP30, one of the two subunits of TFIIF which bind to RNA polymerase II. The further association with TFIIE and other general factors results in a complex capable of accurately initiating RNA synthesis (3–5). The extent of transcriptional activity elicited by these general factors along with RNA polymerase II in a reconstituted *in vitro* system is defined as basal-level transcription. Transcriptional activity is stimulated above the basal level when a sequence-specific transcriptional activator is added to the reconstituted *in vitro* system (2, 6). The mechanism of action of transcriptional activators is not understood. However, certain experiments suggest that TFIIB, a single polypeptide of 316 amino acids, and TFIID, which is composed of the TATA-binding protein (TBP) and the associated factors (TAFs), are targets for sequence-specific activators (7–10). These interactions may either facilitate one or more steps of the assembly of the transcrip-

tional machinery, or cause a conformational change which would trigger initiation by RNA polymerase II, or both.

In this study we use limited proteolysis to define structural domains of TFIIB. We find that this protein is composed of a protease-resistant domain, spanning from residue 106 to the C terminus, and a protease-sensitive N-terminal region. The stable domain, which we refer to as TFIIBc, binds to the TBP-DNA complex like full-length TFIIB, but it is unable to recruit RNA polymerase II. We therefore suggest that this recruitment is accomplished by the N-terminal region of TFIIB, which is missing from TFIIBc. We also discuss results indicating that the proposed action of an acidic activator on the interaction between TFIIB and the TFIID-DNA complex (7) may not require the N-terminal region of TFIIB.

MATERIALS AND METHODS

Proteolytic Digestion of TFIIB. TFIIB (1 mg/ml) was incubated at 37°C with trypsin (Worthington; 5 µg/ml). After 0, 1, 3, 10, 20, 30, 60, and 90 min, aliquots of the reaction mixture were taken and the reaction was stopped by adding phenylmethylsulfonyl fluoride to a final concentration of 10 mM. Subsequently, the samples were mixed with a 2× SDS buffer, heated at 98°C for 4 min, and loaded on an SDS/15% polyacrylamide gel. After electrophoresis, the gel was stained with Coomassie blue.

Transcription Factors. The pET11a expression vector (Novagen) carrying the human TFIIB gene under the control of the T7 promoter was kindly provided by D. Reinberg (11). Recombinant TFIIB was purified from *Escherichia coli* as follows. Cells from a 6-liter culture were harvested and suspended in 300 ml of buffer X (10 mM Tris, pH 8/25 mM EDTA/0.1 M KCl/10 mM 2-mercaptoethanol/1 mM phenylmethylsulfonyl fluoride/20 µM benzamidine hydrochloride). Cell disruption was performed by sonication and the lysate was cleared of debris by centrifugation at 10,000 × g for 20 min. Polyethyleneimine (pH 7) was slowly added to the cleared lysate to a final concentration of 0.1%, and the precipitated nucleic acids were removed by centrifugation at 10,000 × g for 20 min. Recombinant TFIIB was precipitated from the supernatant by ammonium sulfate at 0.25 g/ml. After centrifugation at 10,000 × g for 20 min, the protein pellet was suspended in 30 ml of buffer A (20 mM Hepes, pH 7.5/10 mM 2-mercaptoethanol/1 mM phenylmethylsulfonyl fluoride/20 µM benzamidine hydrochloride), and NaCl was added until this buffer reached the conductivity of a 0.1 M NaCl solution. Soluble proteins were loaded on a phosphocellulose column (5 mg per ml of resin) preequilibrated with 0.1 M NaCl/buffer A, and eluted with a gradient of NaCl (0.1–0.8 M) in buffer A. The TFIIB-containing fractions (0.4–0.5 M NaCl) were dialyzed against buffer D [20 mM Hepes, pH 7.5/0.1 M KCl/0.1 mM EDTA/20% (vol/vol)

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: TBP, TATA-binding protein; TAF, TBP-associated factor.

glycerol/1 mM dithiothreitol/0.1 mM phenylmethylsulfonyl fluoride] and frozen in liquid nitrogen.

The TFIIB proteolytic fragment TFIIBc was purified as follows. Trypsin digestion of TFIIB was stopped after 3 min by adding an excess of phenylmethylsulfonyl fluoride. This solution was adjusted to 0.1 M NaCl and loaded on a phosphocellulose column preequilibrated with 20 mM Hepes, pH 7.5/2 mM dithiothreitol/0.1 M NaCl. TFIIBc was eluted with a gradient of NaCl (0.1–1.0 M) in the same buffer. Purified TFIIBc was subjected to N-terminal sequence analysis and molecular weight determination by mass spectrometry at the Harvard Microchemistry Facility, Harvard University. A plasmid for the overexpression of the TFIIBc polypeptide in *E. coli* was constructed as follows. A fragment of the TFIIB gene spanning from the methionine codon 107 to the stop codon was generated by PCR with *Pfu* DNA polymerase (Stratagene) and 5' and 3' primers containing *Nhe* I and *Bam*HI sites, respectively (introducing the *Nhe* I site changed the serine codon 108 into an alanine codon). The PCR products were digested with *Nhe* I and *Bam*HI and cloned into the *Nhe* I/*Bam*HI sites present downstream of the T7 promoter on the plasmid pET11a. Recombinant TFIIBc was purified according to the protocol that was used for TFIIB, with the following modifications. Ammonium sulfate at 0.45 g/ml was used to precipitate TFIIBc from the cleared *E. coli* lysate. After elution from the phosphocellulose column, the TFIIBc-containing solution was adjusted to 0.5 M NaCl and loaded on a hydroxylapatite column preequilibrated with the same buffer. TFIIBc was eluted with a gradient of potassium phosphate (0.0–0.5 M) in buffer A and then was dialyzed against buffer D. The purity of TFIIBc after this step was >95%.

Recombinant human TBP was tagged with six histidines at the N terminus, expressed in *E. coli* from the T7 promoter of the pET11a vector, and purified over a nickel-agarose column (Qiagen) according to the manufacturer's protocol. TFIID, TFIIA, TFIIE/F (a chromatographic fraction containing both TFIIE and TFIIF activities), and RNA polymerase II were purified essentially as described (12, 13). The more recently identified transcription factors TFIIF and TFIIF are believed to be present in these chromatographic fractions (3). Purified GAL4-VP16, a chimeric protein that is an acidic

transcriptional activator, was kindly provided by M. Carey (14).

Gel Mobility-Shift Assays. The 20- μ l binding reaction mixtures typically contained 10 mM Hepes (pH 7.5), 60 mM KCl, 7.5 mM MgCl₂, 0.5 mM dithiothreitol, 10% glycerol, 1 μ g of bovine serum albumin, 1 μ g of poly(dG)·poly(dC), and 100 pM ³²P-labeled DNA fragment containing the sequence from the adenovirus major late promoter from –53 to +33. The protein components indicated in the figures were added to the reaction mixtures in the following amounts: 30 nM recombinant human TBP, 30 nM recombinant TFIIB, 30 nM recombinant TFIIBc, 1 μ l of TFIIA fraction, 1 μ l of TFIIE/F fraction, and 1 μ l of RNA polymerase II. After 30 min at 30°C, the mixture was loaded onto 4% polyacrylamide gels containing 25 mM Tris base, 25 mM boric acid, and 0.65 mM EDTA. Electrophoresis was carried out in gel buffer at 10 V/cm at room temperature.

In Vitro Transcription Reactions. The DNA template G5E4T, kindly provided by M. Carey (14), contains five tandem GAL4-binding sites upstream of the TATA box of the adenovirus E4 gene. The 40- μ l mixture typically contained 20 ng of supercoiled G5E4T template, 280 ng of pGEM-3 (Promega) carrier DNA, 10 mM Hepes (pH 7.5), 60 mM KCl, 7.5 mM MgCl₂, 0.5 mM dithiothreitol, 10% glycerol, and 2 μ g of bovine serum albumin. Transcription factors indicated in the figures were added in the following amounts: 15 nM GAL4-VP16, 30 nM recombinant TFIIB (unless otherwise indicated), 30 nM recombinant TFIIBc (unless otherwise indicated), 30 nM recombinant TBP (or 3 μ l of TFIID), 3 μ l of TFIIA, 3 μ l of TFIIE/F, and 2 μ l of RNA polymerase. This mixture was preincubated for 30 min at 30°C; nucleotides were then added to a final concentration of 0.5 mM and the transcription mixtures were incubated for 30 min at 30°C. E4 transcripts were analyzed by primer extension as described by Carey *et al.* (14).

RESULTS

Proteolytic Analysis of TFIIB Defines a Relatively Stable Domain. Recombinant human TFIIB, a 316-amino acid protein (11, 15), was expressed in *E. coli* and purified as described in *Materials and Methods*. The protein was subjected to limited proteolysis by trypsin. A 1-min incubation at

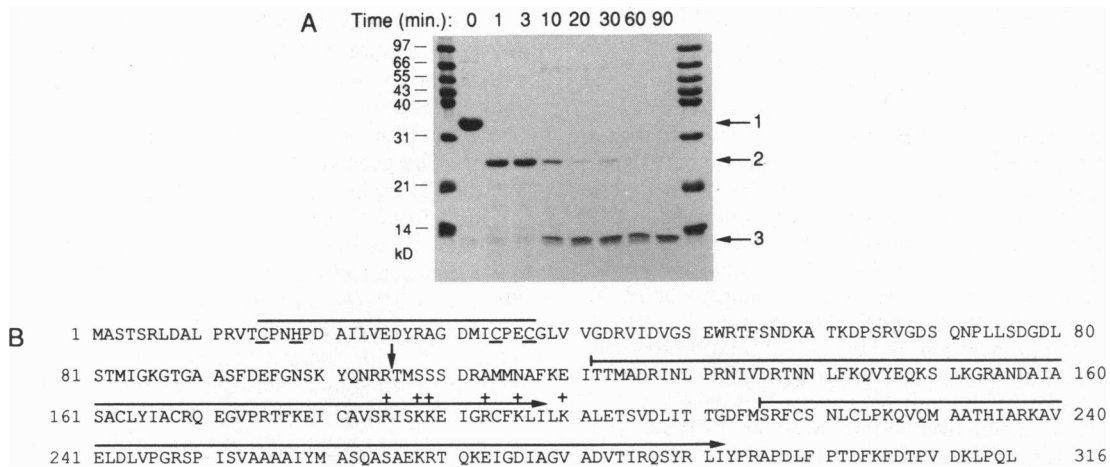


FIG. 1. Limited proteolysis of human TFIIB defines structural domains. (A) Analysis of TFIIB proteolytic fragments by SDS/polyacrylamide gel electrophoresis followed by Coomassie blue staining. Purified recombinant human TFIIB (1 mg/ml) was incubated with trypsin (5 μ g/ml) at 37°C for the times indicated above the panel. Lanes at the extreme left and right contain protein molecular weight markers ($M_r \times 10^{-3}$ indicated at left). The major band 1 corresponds to full-length TFIIB, whose calculated molecular weight is 34,800. Band 2 is the TFIIB proteolytic fragment that we have analyzed in this study, TFIIBc. The proteolytic fragment corresponding to band 3 has not been characterized by sequence analysis. (B) Amino acid sequence of human TFIIB in the single-letter code. The structural features predicted from the TFIIB sequence are denoted as follows: the putative zinc finger is indicated by an overline; three cysteines and one histidine that may bind one zinc ion are underlined. The direct repeat is indicated by unidirectional arrows. The basic amino acid residues of the amphipathic α -helix are denoted by plus signs. The vertical arrow indicates the trypsin cleavage site between amino acid residues 105 and 106, which forms the N terminus of TFIIBc.

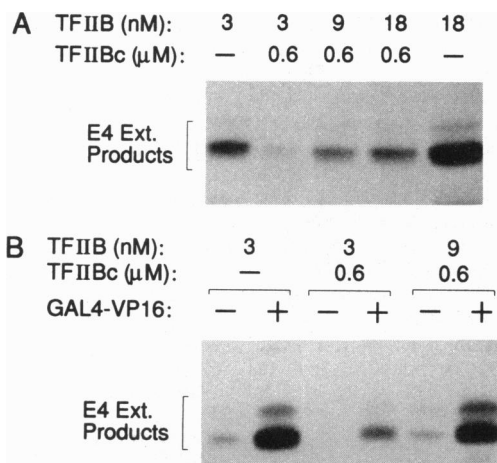


FIG. 4. TFIIBc inhibits basal and activated transcription by competing with full-length TFIIB. (A) Basal transcription with recombinant human TBP substituting for TFIID. *In vitro* transcription of the E4 gene was performed with recombinant general factors and RNA polymerase II. Transcripts were quantitated by primer extension. The concentrations of TFIIB and TFIIBc used in the various reactions are indicated at the top. (B) Basal and activated transcription with TFIID combined with the other general factors and RNA polymerase II in the absence (–) or presence (+) of the activator GAL4-VP16. Reaction conditions were as in A. The concentrations of TFIIB and TFIIBc are indicated.

The ability of TFIIBc to substitute for full-length TFIIB in a transcriptionally active complex was tested by *in vitro* transcription with purified transcription factors. While a reconstituted transcription complex with full-length TFIIB was able to support basal-level transcription as well as to respond to the activator GAL4-VP16, TFIIBc was unable to substitute for TFIIB, and no transcripts were detected either in the presence or in the absence of GAL4-VP16 (Fig. 2B). Transcription was absolutely dependent on the presence of TFIIB, and no difference in activity was observed between the purified proteolytic fragment and recombinant TFIIBc (data not shown). These results suggest that TFIIBc, although capable of interacting with TBP, is unable to participate in the assembly of a functional transcription complex.

The N-Terminal Region of TFIIB Is Necessary for the Recruitment of RNA Polymerase II to the Promoter. We analyzed the formation of the transcription complex by mobility-shift assay to determine which step was arrested in the presence of TFIIBc. While the TBP-TFIIA-TFIIB (TAB) and TBP-TFIIA-TFIIBc (TABc) complexes formed with equal efficiency on DNA, only TAB, but not TABc, was able to form a more slowly migrating complex with RNA polymerase II and the TFIIE/F fraction containing RAP30 (Fig. 3). These results suggest that the N-terminal region of TFIIB, which is missing from TFIIBc, is necessary for the recruitment of RNA polymerase II to the promoter. In the absence of TBP, none of the factors showed any DNA-binding activity (Fig. 3, lanes 6 and 7).

TFIIBc Can Block Basal and Activated Transcription by Competing with Full-Length TFIIB. An excess of TFIIBc was added to a reconstituted *in vitro* transcription system with TBP substituting for the TFIID fraction and TFIIB present at low concentrations. An excess of TFIIBc inhibited basal transcription (Fig. 4A). Similar high concentrations of full-length TFIIB did not repress transcription (data not shown). Such inhibition was counteracted by increasing the concentration of full-length TFIIB, suggesting that TFIIBc and TFIIB were competing for the interaction with TBP, whereas only TFIIB could participate in the formation of a functional transcription complex.

In a second set of experiments, we tested whether TFIIBc would have the same dominant negative effect on transcription with the TFIID fraction (TBP plus TAFs) replacing the recombinant TBP in a reconstituted transcription complex and whether a transcriptional activator might change the equilibrium of the TFIIB/TFIIBc competition. In the presence of the TFIID fraction, TFIIBc inhibited transcription by competing with intact TFIIB as it did when recombinant TBP was used (Fig. 4B, compare lanes 1, 3, and 5). Moreover, the presence of the acidic activator GAL4-VP16 did not affect the competition between the two molecules (compare lanes 2, 4, and 6).

DISCUSSION

We have used limited proteolysis to examine structural features of human TFIIB. We found that the C-terminal region of TFIIB, comprising residues 106–316, is resistant to protease digestion, suggesting that it forms a tightly folded domain. In contrast, the N-terminal region of TFIIB is very sensitive to protease digestion, and it is therefore likely to be in a more extended or flexible conformation. Of the structural features predicted from the TFIIB primary sequence (11, 15, 16), the direct repeat and the putative basic amphipathic α -helix are located within the stable C-terminal domain, which we refer to as TFIIBc, whereas the zinc-finger motif resides in the N-terminal region of TFIIB (Fig. 1B).

We have shown that, like full-length TFIIB, TFIIBc can interact with TBP on DNA, in either the absence or the presence of TFIIA (Figs. 2A and 3). It has been observed in DNase I footprinting assays that TFIIB contacts DNA downstream of the TATA box upon interaction with TBP (4, 17, 18). Our observation that both TFIIBc and full-length TFIIB stabilize the TBP–DNA interaction in a gel mobility-shift assay suggests a cooperative DNA binding of TBP with TFIIB or TFIIBc. This interpretation implies that TFIIBc contains a surface for interaction with DNA as well as one for binding to TBP. It is therefore unlikely that the zinc-finger motif located near the N terminus of TFIIB and missing from TFIIBc is involved in DNA binding (16).

TFIIBc is unable to substitute for full-length TFIIB in an *in vitro* transcription assay, despite its ability to form a complex with TBP on the promoter (Fig. 2B). Consistent with these results, we have found that TFIIBc arrests the assembly of the transcription complex after binding to TBP, because it is unable to recruit RNA polymerase II (Fig. 3). We therefore conclude that the N-terminal region of TFIIB, which is missing from TFIIBc, is required for the recruitment of RNA polymerase II to the promoter. In agreement with this conclusion, the paper by Buratowski and Zhou (19) in this issue presents results of site-directed mutagenesis experiments showing that the cysteines of the zinc finger motif near the N terminus of TFIIB are essential for the recruitment of RNA polymerase II. We do not know whether the TFIIB N-terminal region recruits RNA polymerase II by directly interacting with the enzyme or by binding to RAP30, one of the two TFIIF subunits that are required for promoter recognition by RNA polymerase II (20). If the latter case is correct, RAP30 would act as a bridging (docking) protein between RNA polymerase II and TFIIB.

One prediction, based on the functional properties of TFIIBc discussed above, is that TFIIBc should have a dominant negative effect in an *in vitro* transcription assay when present in excess over full-length TFIIB. Indeed, we have shown that TFIIBc can inhibit transcription by competing with full-length TFIIB for the interaction with TBP (Fig. 4A). We have also found that the TAFs, which are present in the TFIID fraction and are required for the response of a transcription complex to an activator (21–24), do not appear to change the accessibility of TBP for TFIIB

in a way that might favor the interaction of TBP with either the full-length protein or TFIIBc. Thus TFIIBc inhibited transcription by a preparation containing complete TFIID as well as it did when purified recombinant TBP was used. Furthermore, the presence of the acidic transcriptional activator GAL4-VP16 does not counteract inhibition by TFIIBc. This result indicates that the activator does not change the equilibrium of the TFIIB/TFIIBc competition; in other words, it does not facilitate the assembly of the functional transcription complex over the nonfunctional complex containing TFIIBc. One proposed function of an acidic activator is enhancement of the interaction between TFIIB and the TFIID-DNA complex (7). Our results suggest that, if this function of an activator is real, it can be exerted also on TFIIBc and, therefore, it does not require the protease-sensitive N-terminal region of TFIIB.

We thank William Lane for sequence analysis and molecular weight determination of TFIIBc. We thank Drs. S. Buratowski and D. Reinberg for communicating results prior to publication. We thank Jun Ma, Joe Pearlberg, Josh Brickman, and Richard Reece for comments on the manuscript and Luisella Barberis-Maino for superb technical assistance. A.B. was supported by fellowships from the European Molecular Biology Organization and the Swiss National Science Foundation. C.W.M. was supported by a fellowship from the Deutsche Forschungsgemeinschaft. This work was further supported by grants from the National Institutes of Health to M.P. and S.C.H.

1. Sawadogo, M. & Sentenac, A. (1990) *Annu. Rev. Biochem.* **59**, 711-754.
2. Roeder, R. G. (1991) *Trends Biochem. Sci.* **16**, 402-408.
3. Zawel, L. & Reinberg, D. (1992) *Curr. Opin. Cell. Biol.* **4**, 488-495.
4. Buratowski, S., Hahn, S., Guarente, L. & Sharp, P. A. (1989) *Cell* **56**, 549-561.
5. Flores, O., Lu, H., Killeen, M., Greenblatt, J., Burton, Z. F. & Reinberg, D. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 9999-10003.
6. Ptashne, M. & Gann, A. (1990) *Nature (London)* **346**, 329-331.
7. Lin, Y. S. & Green, M. R. (1991) *Cell* **64**, 971-981.
8. Sundseth, R. & Hansen, U. (1992) *J. Biol. Chem.* **267**, 7845-7855.
9. Horikoshi, M., Carey, M., Kakidani, H. & Roeder, R. G. (1988) *Cell* **54**, 665-669.
10. Stringer, K. F., Ingles, C. J. & Greenblatt, J. (1990) *Nature (London)* **345**, 783-786.
11. Ha, I., Lane, W. S. & Reinberg, D. (1991) *Nature (London)* **352**, 689-695.
12. Dignam, J. D., Martin, P. L., Shastry, B. S. & Roeder, R. G. (1983) *Methods Enzymol.* **101**, 582-598.
13. Reinberg, D. & Roeder, R. G. (1987) *J. Biol. Chem.* **262**, 3310-3321.
14. Carey, M., Leatherwood, J. & Ptashne, M. (1990) *Science* **247**, 710-712.
15. Malik, S., Hisatake, K., Sumimoto, H., Horikoshi, M. & Roeder, R. G. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 9553-9557.
16. Pinto, I., Ware, D. E. & Hampsey, M. (1992) *Cell* **68**, 977-988.
17. Maldonado, E., Ha, I., Cortes, P., Weis, L. & Reinberg, D. (1990) *Mol. Cell. Biol.* **10**, 6335-6347.
18. Moncollin, V., Fischer, L., Cavallini, B., Egly, J.-M. & Chambon, P. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 397-401.
19. Buratowski, S. & Zhou, H. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 5633-5637.
20. Killeen, M., Coulombe, B. & Greenblatt, J. (1992) *J. Biol. Chem.* **267**, 9463-9466.
21. Peterson, M. G., Tanese, N., Pugh, B. F. & Tjian, R. (1990) *Science* **248**, 1625-1630.
22. Dynlacht, B. D., Hoey, T. & Tjian, R. (1991) *Cell* **55**, 563-576.
23. Meisterernst, M. & Roeder, R. G. (1991) *Cell* **67**, 557-567.
24. Zhou, Q., Liebermann, P. M., Boyer, T. G. & Berk, A. J. (1992) *Genes Dev.* **6**, 1964-1974.