

- Sci. U.S.A.* **82**, 168 (1986); R. H. Haynes and B. A. Kunz, in *The Molecular Biology of the Yeast Saccharomyces: Life Cycle and Inheritance*, J. N. Strathern, E. W. Jones, J. E. Broach, Eds. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1981), pp. 371-414; C. W. Lawrence, *Adv. Genet.* **21**, 173 (1982).
17. M. O. Dayhoff, in *Atlas of Protein Sequence and Structure*, M. O. Dayhoff, Ed. (National Biomedical Research Foundation, Washington, DC, 1978), vol. 5, suppl. 3, p. 345.
 18. S. Jentsch, J. P. McGrath, A. Varshavsky, *Nature* **329**, 131 (1987).
 19. The CDC34-containing Sca I-Fnu DII fragment (Fig. 3) was subcloned with Hind III linkers (New England Biolabs) into the Hind III site of the vector pKK223-3 [J. Brosius and J. R. Lupski, *Methods Enzymol.* **153**, 54 (1987)], downstream of the vector's *tac* promoter, in both the expressing and opposite (presumably nonexpressing) orientations to generate plasmids pK34-1 and pK34-2, respectively. To construct pK34-3, a single copy of a Bgl II linker, CAGATCTG (New England Biolabs), was inserted into plasmid pK34-1 at its unique Eco RV site (see Fig. 3), yielding the mutation *cdc34-3*, which encodes a truncated protein. Insertion of the linker generates a Gln-Ile insertion at amino acid residue 170 of the wild-type CDC34 protein (Fig. 3), followed (at the nucleotide sequence level) by a TGA stop codon.
 20. *Escherichia coli* cells harboring the pK34-1, pK34-2, or pK34-3 plasmid (19) were induced (19) and lysed by freeze-thawing in 5 volumes of 50 mM tris-HCl (pH 7.5), 1 mM EDTA, and 1 mM dithiothreitol (DTT). The assay for ubiquitin-histone conjugation was carried out in 50 μ l of 50 mM tris-HCl (pH 7.5), 2 mM ATP, 5 mM MgCl₂, and 0.2 mM DTT, in the presence of 50 pmol of ¹²⁵I-labeled ubiquitin (~1 \times 10⁵ cpm) (21, 22), 0.2 μ g of bovine histone H2B (Boehringer Mannheim), and 4 pmol of the purified yeast E1 enzyme (18). The assay for ubiquitin thioesters was carried out as above except that histone H2B was omitted, and the reaction products were incubated for 5 min at room temperature in an SDS-containing electrophoretic sample buffer lacking reducing agents before a polyacrylamide-SDS gel electrophoresis at 4°C (22). Reduction of thioesters was performed by heating the samples prior to electrophoresis at 100°C for 10 min in an SDS sample buffer containing 4M 2-mercaptoethanol (18, 21).
 21. D. Finley and A. Varshavsky, *Trends Biochem. Sci.* **10**, 343 (1985); A. Hershko and A. Ciechanover, *Prog. Nucleic Acid Res. Mol. Biol.* **33**, 19 (1986); M. Rechsteiner, *Annu. Rev. Cell Biol.* **3**, 1 (1987).
 22. D. Finley, A. Ciechanover, A. Varshavsky, *Cell* **37**, 43 (1984); A. Ciechanover, D. Finley, A. Varshavsky, *ibid.*, p. 57; A. Bachmair, D. Finley, A. Varshavsky, *Science* **234**, 179 (1986); D. Finley, E. Özkaynak, A. Varshavsky, *Cell* **48**, 1035 (1987).
 23. C. M. Pickart and I. A. Rose, *J. Biol. Chem.* **261**, 10210 (1985).
 24. S. Jentsch, J. P. McGrath, A. Varshavsky, unpublished data.
 25. E. Ball *et al.*, *Cell* **51**, 221 (1987); T. Yarden *et al.*, *Nature* **323**, 226 (1986); M. Siegelman *et al.*, *Science* **231**, 823 (1986).
 26. H. Busch, *Methods Enzymol.* **106**, 238 (1984); R. D. Mueller, H. Yasuda, C. L. Hatch, W. M. Bonner, E. M. Bradbury, *J. Biol. Chem.* **260**, 5147 (1985); L. Levinger and A. Varshavsky, *Cell* **28**, 375 (1982).
 27. The chromosome bearing the CDC34 gene was identified by probing a nitrocellulose filter containing chromosome-sized yeast genomic DNA separated by pulsed-field electrophoresis [G. Carle and M. V. Olson, *Nucleic Acids Res.* **12**, 5647 (1984)] with ³²P-labeled pCDC34-79 DNA (29). The results (8) tentatively localized CDC34 to chromosome IV. Diploid strains containing *cdc34-1* and various genetic markers present on the right arm of chromosome IV (32) were then constructed. The results of tetrad analyses with these strains are presented in Table 1.
 28. R. K. Mortimer and D. Schild, *Microbiol. Rev.* **49**, 181 (1985).
 29. P. W. J. Rigby, M. Dieckmann, C. Rhodes, P. Berg, *J. Mol. Biol.* **113**, 237 (1977). Hybridization conditions [G. M. Church and W. Gilbert, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 1991 (1984)]: overnight incubation in 7% SDS, 1 mM EDTA, and 0.5M sodium phosphate (pH 7.2) at 65°C, followed by a 5-min washing with 5% SDS and 40 mM sodium phosphate (pH 7.2) at 65°C, and further 5-min washings (five times) with 1% SDS and 40 mM sodium phosphate (pH 7.2) at room temperature.
 30. E. M. Southern, *J. Mol. Biol.* **98**, 503 (1975).
 31. T. Maniatis, E. F. Fritsch, J. Sambrook, *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982).
 32. F. Sherman, G. R. Fink, J. Hicks, *Methods in Yeast Genetics* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1981).
 33. A 3.2-kb Bgl II-Sph I fragment and a 3.3-kb Hind III-Sal I fragment from pCDC34-79 (Fig. 1) were subcloned into pGEM-1 (Promega Biotech, Madison, WI); the resulting plasmids were designated pGEM34 B/S and pGEM34 H/S, respectively.
 34. We thank L. Hartwell, W. Fangman, A. Hopper, J. Hurley, D. Finley, and B. Bartel for advice during the preparation of the manuscript. We thank C. Mann for the gift of plasmid CMp170 and R. McCarroll for the Southern blot of separated yeast chromosomes. Supported by grants from the National Institute of General Medical Sciences to B.B. (GM18541) and A.V. (GM31530).

9 March 1988; accepted 29 June 1988

Physical Analysis of Transcription Preinitiation Complex Assembly on a Class II Gene Promoter

MICHAEL W. VAN DYKE, ROBERT G. ROEDER, MICHÈLE SAWADOGO

Transcription of protein-encoding genes by human RNA polymerase II requires multiple ancillary proteins (transcription factors). Interactions between these proteins and the promoter DNA of a viral class II gene (the major late transcription unit of adenovirus) were investigated by enzymatic and chemical footprinting. The experiments indicated that the assembly of functionally active RNA polymerase II-containing transcription preinitiation complexes requires a complete set of transcription factors, and that both specific protein-DNA and protein-protein interactions are involved. This allows individual steps along the transcription reaction pathway to be tested directly, thus providing a basis for understanding basic transcription initiation mechanisms as well as the regulatory processes that act on them.

GENE ACTIVITY IN EUKARYOTES IS frequently regulated at the level of DNA transcription. Understanding the molecular mechanism behind this transcriptional control of gene expression has become one of the major goals of modern biology. Mutational analysis has revealed that the activity of protein-encoding genes is governed by several kinds of promoter elements. A basal level of transcription is usually observed when all but a small region of the promoter DNA has been deleted. This minimum or core promoter is often centered on a TATA box sequence (1). Two classes of cis-acting DNA elements are involved in modulating the TATA box-driven transcription. Upstream promoter elements act at short distance, while enhancer elements can be located up to several thousand base pairs from the transcription initiation site (2). The various components of the human class II transcription machinery are being elucidated by the biochemical dissection of crude cell-free systems (3, 4). This approach has revealed that specific transcription initiation by RNA polymerase II requires the coordinated action of multiple protein factors. These can be classified into two categories: general transcription factors and upstream element-binding proteins (2). Through both kinetic analysis and the use of

various inhibitors, several steps have been defined within the transcription mechanism (5-7). These include: (i) a commitment of the template after binding of a subset of the transcription factors, (ii) formation of an activated state through the action of other transcription factors and RNA polymerase II, (iii) fulfillment of an energy requirement (hydrolysis of the β - γ phosphate bond of either ATP or dATP), (iv) initiation of transcription (formation of the first phosphodiester bond), and (v) transcription elongation, with some transcription components remaining committed to the template. Within these various steps, the exact roles of the different protein factors, their assembly into active transcription complexes, and the various transitions that lead to the initiation event, remain for the most part unknown.

The major late (ML) promoter of adenovirus provides a useful model system for in vitro analysis of specific transcription by human RNA polymerase II. This promoter is quite simple structurally, with only two essential DNA elements: a TATA box at position -28 and a single upstream element at position -58 (8-11). At least three transcription factors (designated TFIIB, TFIID,

Laboratory of Biochemistry and Molecular Biology, The Rockefeller University, New York, NY 10021.

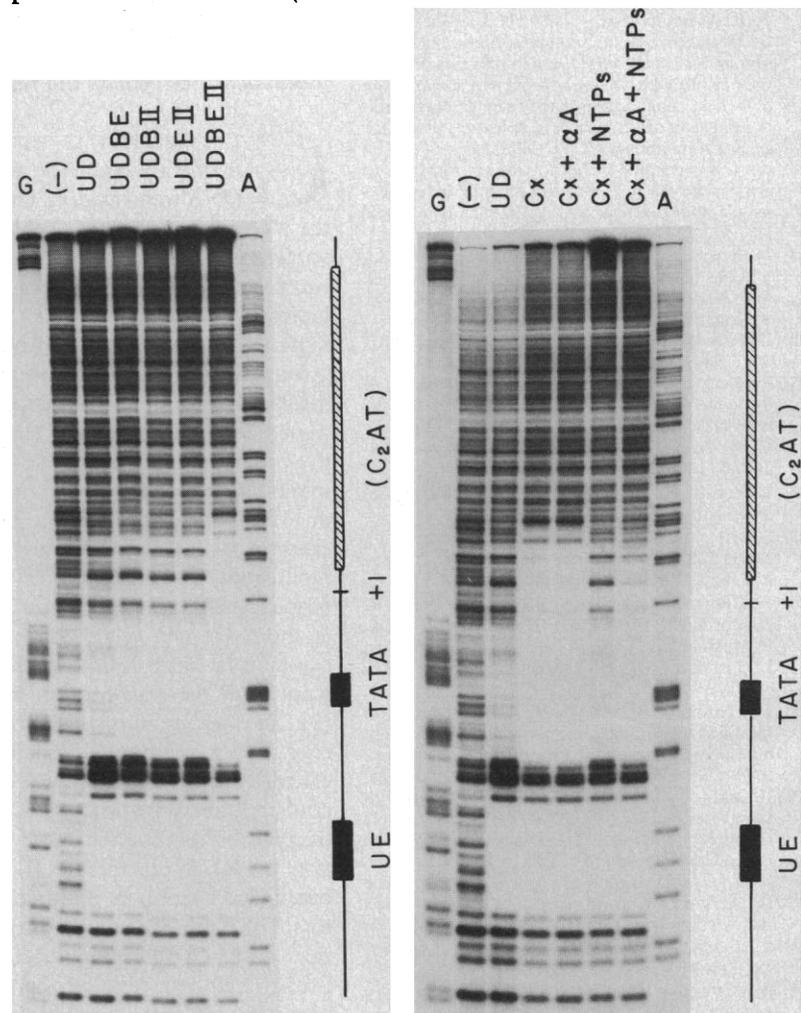
and TFIIE) are absolutely necessary, in addition to the RNA polymerase II itself, for a basal level of expression driven by the ML TATA box (4). Maximal activity of the promoter requires in addition a gene-specific transcription factor designated upstream stimulatory factor (USF) (10) [or MLTF, UEF (11)]. DNA cleavage protection methods (footprinting) have been used successfully to analyze the primary protein-DNA interactions necessary for ML transcription, namely the binding of TFIID to the TATA element (10, 12) and of USF to the ML upstream element (10). We report here that the same methods can be used to visualize directly subsequent steps along the reaction pathway.

Our *in vitro* reconstituted transcription system, composed of HeLa cell-derived transcription factors TFIIB, TFIID, TFIIE, USF, and RNA polymerase II (13), has been described (4, 10). For a footprinting analysis of transcription complexes, USF and TFIID were first incubated with a small amount of ³²P-labeled DNA containing the ML promoter under suitable conditions for complete binding (10). These preformed complexes were then subjected to a second incubation with the other protein fractions under transcription conditions (except for the absence of nucleotide triphosphates and the addition of a large concentration of non-specific carrier DNA). We first investigated the transcription factor requirement for the assembly of complete preinitiation complexes (Fig. 1). The deoxyribonuclease I (DNase I) footprinting pattern observed for the USF-TFIID-ML promoter complex (lane 3)

is characterized by complete cleavage protections centered over the upstream element and TATA boxes, with a series of periodic enhanced cleavages observed between positions -2 and +30 (10). This pattern of protection did not change when subsets of the remaining transcription factors were subsequently added (lanes 4 to 6). Only with the simultaneous addition of TFIIB, TFIIE, and RNA polymerase II was a novel DNase I footprint observed (lane 7). This footprint was characterized by a loss of the periodic pattern of DNase I-enhanced cleavages in the downstream portion of the TFIID footprint, and the appearance of a new hypersensitive site downstream at position +35. This observation concurred with the earlier studies indicating that at least one protein in each of these fractions was absolutely required for specific transcription initiation at the ML promoter (4). Our inability to detect interactions between the USF-TFIID-ML promoter complex and subsets of the remaining transcription factors could reflect greater instabilities and lower concentrations of these intermediates. By contrast, the whole preinitiation complex could be stabilized by multiple protein-protein and protein-DNA interactions (see below).

Both the involvement of DNA sequences deemed important by mutational studies and the requirement for all previously identified transcription factors supported the idea that the above-described novel footprint reflected the assembly of transcription preinitiation complexes onto the promoter of the ML gene. However, it was essential to determine whether a majority of the complexes formed under these conditions were fully functional, that is, responsive to known RNA polymerase II effectors (substrates, inhibitors, and cofactors). For example, transcription initiation should take place upon addition of ribonucleoside triphosphates. In the template construction used for these experiments, the ML promoter drives transcription of an artificial DNA fragment lacking G residues on the RNA-like strand (C₂AT cassette) (4). Thus transcription through this region only requires the three nucleotides ATP, CTP, and UTP. In the absence of the fourth nucleotide, progression of the RNA polymerase should cease at the end of the cassette. Finally, a low concentration of the specific inhibitor α -amanitin should prevent transcription initiation by RNA polymerase II.

Fig. 1 (left). Transcription factor requirement for preinitiation complex assembly. Transcription complexes were assembled on the ML promoter of adenovirus [Xba I to Nar I 690-bp restriction fragment from plasmid pML(C₂AT)19 Δ -127f (10) labeled upstream of the promoter] after a two-step incubation protocol. First, TFIID [D, 1.7 units, ω -aminooctyl agarose fraction (23)] and USF (U, 80 fmole; DE52 fraction) were incubated for 60 min at 30°C with 2 ng of the probe and 20 ng of carrier poly[d(GC)] in a 25- μ l reaction. Subsequent incubation with the other transcription factors TFIIB (B, 2.3 units, single-stranded DNA agarose fraction), TFIIE (E, 2.3 units, Bio-Gel A-1.5m fraction), and RNA polymerase II (II, 20 units, phosphocellulose fraction) was carried out at 30°C for 30 min in a 50- μ l final volume under transcription conditions (4) except for the absence of nucleoside triphosphates and the presence of 800 ng of carrier poly[d(GC)]. These complexes were then analyzed by DNase I footprinting as described (10, 16). The products of adenosine- (A) and guanosine-specific (G) chemical-sequencing reactions (24) were used as markers. A schematic representation of the DNA fragment is shown at right, indicating the location of the USF-binding element (UE), TATA box, initiation site (+1), and transcription cassette (C₂AT). **Fig. 2 (right).** Footprinting analysis of transcription complexes in the presence of α -amanitin and nucleotides. Transcription preinitiation complexes (Cx) were assembled as described in Fig. 1. Where indicated, α -amanitin (α A, 2 μ g/ml) or nucleotides (NTP's, 0.3 mM each of ATP, CTP, UTP; 0.2 mM 3'-O-methyl GTP) were present during the last 10 min of the second incubation. Control reactions contained either no transcription factors (-) or TFIID and USF only (UD).



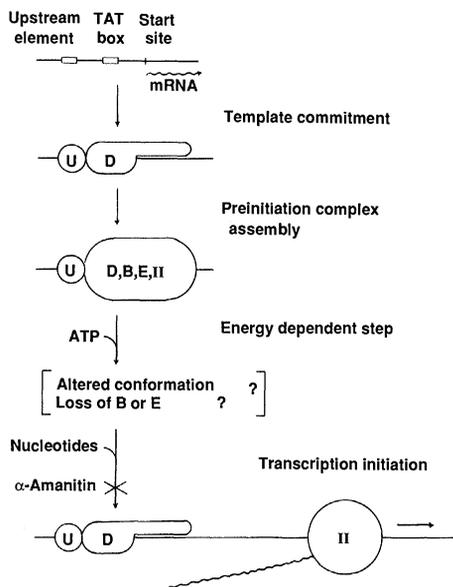


Fig. 5. Model for the mechanism of specific transcription initiation by RNA polymerase II.

cassette, protein-DNA contacts in this region are presumably DNA sequence-independent. In the presence of nucleotides and α -amanitin, the protein-DNA interactions within the transcription preinitiation complexes were significantly altered (Fig. 3, lanes 3; Fig. 4, row 3). This transition was detected by DNase I footprinting as a retreat at the downstream edge of the protected region, together with a loss of the distinctive enhanced cleavage at its 3' border. The concomitant change in the MPE cleavage protection was most noticeable, with a disappearance of the downstream footprint and a somewhat puzzling alteration of the cleavage pattern around the initiation site.

Our experiments demonstrate that physical analyses, such as DNA cleavage-protection methods, can be used to complement functional assays in dissecting the mechanism of specific transcription initiation by eukaryotic RNA polymerase II. In view of these latest results, the model illustrated schematically in Fig. 5 summarizes our current understanding for transcription of the ML gene by human RNA polymerase II. Sequence specificity of transcription initiation is governed by the initial binding of TFIID and USF to the TATA box and upstream element, respectively. This ternary complex is further stabilized by a direct interaction between the two specific DNA binding proteins (10). The initial (rate-limiting) step observed in the functional assay most likely corresponds to the assembly of this primary complex which results in commitment of the promoter to the transcription process (17). A subsequent step leads to the formation of the whole transcription

preinitiation complex and requires the other transcription factors (TFIIB and TFIIE) as well as the RNA polymerase itself. Formation of this complex results in the appearance of specific, though sequence-independent, protein-DNA contacts. Thus, these contacts must be controlled primarily by specific protein-protein interactions (for example, physical contacts between the RNA polymerase and the TFIID-USF complex). In this, the assembly of RNA polymerase II-containing preinitiation complexes presents similarities with the binding of *Acanthamoeba* RNA polymerase I and TIF factor to the cognate ribosomal gene promoter (18). Therefore, a general feature of eukaryotic RNA polymerases may be that promoter recognition is primarily governed by specific protein-protein interactions, rather than through sequence-specific DNA binding as is the case for prokaryotic RNA polymerases (19). Fulfillment of an energy-requiring step leads to a transition in the complex which could reflect the dissociation of a component or a global conformational change (20). This step would perhaps involve a melting of the DNA double helix, a prerequisite for transcription of double-stranded DNA templates. The energy-dependent step for eukaryotic RNA polymerase II would then be the equivalent of the closed-to-open transition for prokaryotic transcription (19). After transcription initiation and promoter clearance by the RNA polymerase, some remnant of the transcription complex remains associated with the ML promoter (minimally USF and TFIID) (21). This concurs with the earlier observation of post-transcriptional template commitment promoting reinitiation (7). However, multiple transcripts are seldom synthesized from the same promoter in our more purified reconstituted system. Thus, while the observed post-elongation complex may be necessary for template commitment, it is operationally insufficient for transcription reinitiation. Elaboration of our model requires additional studies concerning the molecular architecture of the various transcription complexes. These studies should extend our understanding of the general transcription mechanisms and provide a basis for investigating the action of regulatory factors (22).

REFERENCES AND NOTES

1. R. Breathnach and P. A. Chambon, *Annu. Rev. Biochem.* **50**, 349 (1981).
2. W. S. Dynan and R. Tjian, *Nature* **316**, 774 (1985); T. Maniatis, S. Goodbourn, J. A. Fischer, *Science* **236**, 1237 (1987).
3. T. Matsui, J. Segall, P. A. Weil, R. G. Roeder, *J. Biol. Chem.* **255**, 11992 (1980); M. Samuels, A. Fire, P. A. Sharp, *ibid.* **257**, 14419 (1982); J. D. Dignam, P. L. Martin, B. S. Shastry, R. G. Roeder,

Methods Enzymol. **101**, 582 (1983); V. Moncollin, N. G. Miyamoto, X.-M. Zheng, J.-M. Egly, *EMBO J.* **5**, 2577 (1986).

4. M. Sawadogo and R. G. Roeder, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 4394 (1985).
5. B. L. Davison, J.-M. Egly, E. R. Mulvihill, P. Chambon, *Nature* **301**, 680 (1983); A. Fire, M. Samuels, P. A. Sharp, *J. Biol. Chem.* **259**, 2509 (1984); D. K. Hawley and R. G. Roeder, *ibid.* **260**, 8163 (1985); D. Reinberg, M. Horikoshi, R. G. Roeder, *ibid.* **262**, 3322 (1987); H. Cai and D. S. Luse, *ibid.* **262**, 298 (1987).
6. D. Bunick, R. Zandomeni, S. Ackerman, R. Weinmann, *Cell* **29**, 877 (1982); M. Sawadogo and R. G. Roeder, *J. Biol. Chem.* **259**, 5321 (1984).
7. D. K. Hawley and R. G. Roeder, *J. Biol. Chem.* **262**, 3452 (1987).
8. J. Corden *et al.*, *Science* **209**, 1406 (1980); S. L. Hu and J. L. Manley, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 820 (1981); M. Concino, R. A. Goldman, M. H. Caruthers, R. Weinmann, *J. Biol. Chem.* **258**, 8493 (1983).
9. R. Hen, P. Sassone-Corsi, J. Corden, M. P. Gaub, P. Chambon, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 7132 (1982); Y. T. Yu and J. L. Manley, *Nucleic Acids Res.* **12**, 9309 (1984).
10. M. Sawadogo and R. G. Roeder, *Cell* **43**, 165 (1985).
11. R. W. Carthew, L. A. Chodosh, P. A. Sharp, *ibid.*, p. 439; N. G. Miyamoto, V. Moncollin, J.-M. Egly, P. Chambon, *EMBO J.* **4**, 3563 (1985).
12. N. Nakajima, M. Horikoshi, R. G. Roeder, *Mol. Cell. Biol.*, in press.
13. For convenience, we refer to our partially purified chromatographic fractions as transcription factors. Although further purification of these fractions has so far failed to reveal additional components, this possibility cannot yet be totally excluded. There is also a variable requirement for a factor designated TFIIA which is discussed in (4).
14. M. W. Van Dyke and M. Sawadogo, unpublished observations.
15. R. P. Hertzberg and P. B. Dervan, *J. Am. Chem. Soc.* **104**, 313 (1982).
16. M. W. Van Dyke and R. G. Roeder, *Mol. Cell. Biol.* **7**, 1021 (1987).
17. M. W. Van Dyke, M. Sawadogo, R. G. Roeder, in preparation.
18. E. Bateman, C. T. Ieda, P. Kownin, M. R. Paule, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 8004 (1985); E. Bateman and M. R. Paule, *Cell* **47**, 445 (1986).
19. P. H. von Hippel, D. G. Bear, W. D. Morgan, J. A. McSwiggen, *Annu. Rev. Biochem.* **53**, 389 (1984).
20. An apparent retreat of RNA polymerase II preinitiation complexes in the presence of ATP had been observed in a more indirect way by exhaustive DNase I digestion of DNA templates in crude HeLa nuclear extracts [H. Cai and D. S. Luse, *Mol. Cell. Biol.* **7**, 3371 (1987)].
21. After transcription initiation, cleavage protections indistinguishable from those of USF and TFIID were observed by both DNase I and MPE footprinting.
22. Data from our laboratory has demonstrated that the upstream stimulatory factor ATF acts on transcription of the adenovirus E4 gene by altering the TFIID-promoter interaction, thereby facilitating preinitiation complex assembly (M. Horikoshi, T.-Y. Hai, Y.-S. Lin, M. Green, R. G. Roeder, *Cell* **54**, 1033 (1988)).
23. Transcription factors were prepared as previously described: TFIID (12), USF (10), and TFIIB and TFIIE (4). Human RNA polymerase II was purified from HeLa cells by a protocol essentially identical to that of G. A. Bitter, *Anal. Biochem.* **128**, 294 (1983). See (4) for the definition of units.
24. B. L. Iverson and P. B. Dervan, *Nucleic Acids Res.* **15**, 7823 (1987); A. M. Maxam and W. Gilbert, *Methods Enzymol.* **65**, 499 (1980).
25. We thank M.-C. Weichs and der Glon for technical assistance and P. B. Dervan for providing MPE. Supported by NIH grants GM 38212 (M.S.) and CA 42567 (R.G.R.), by general support from the PEW Trust to the Rockefeller University, and by a Leukemia Society of America Special Fellowship (M.S.).

13 May 1988; accepted 9 August 1988