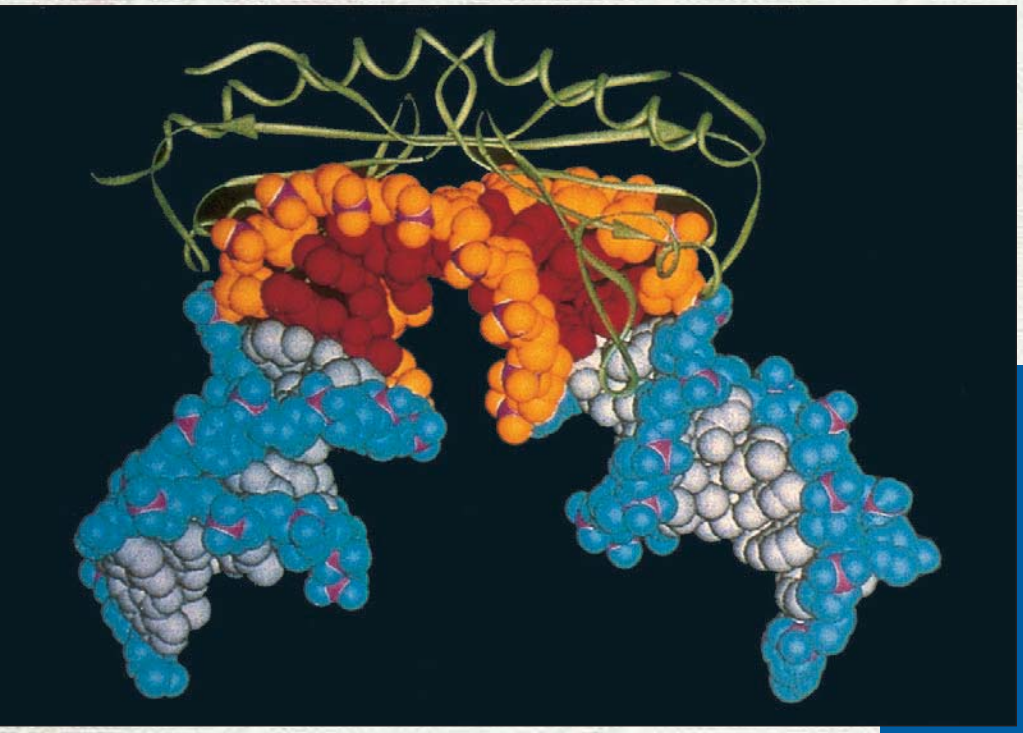
CHAPTER 11

General Transcription Factors in Eukaryotes



X-ray crystal structure of the TBP-TATA box complex. © Klug, A. Opening the gateway. *Nature* 365 (7 Oct 1993) p. 487, f. 2. © Macmillan Magazines Ltd.

Eukaryotic RNA polymerases, unlike their bacterial counterparts, are incapable of binding by themselves to their respective promoters. Instead, they rely on proteins called transcription factors to show them the way. Such factors are grouped into two classes: **general transcription factors** and gene-specific transcription factors (**activators**). Without activators, the general transcription factors can attract the RNA polymerases to their respective promoters, but only to a weak extent. Therefore, these factors can support only a basal level of transcription. Furthermore, general transcription factors and the three polymerases alone allow for only minimal transcription control, whereas activators help cells exert exquisitely fine control over transcription. Nevertheless, the task performed by the general transcription factors—getting the RNA polymerases together with their promoters—is not only vital, but also very complex because many

polypeptides are required to do the job. In this chapter we will survey the general transcription factors that interact with all three RNA polymerases and their promoters.

11.1 Class II Factors

The general transcription factors combine with RNA polymerase to form a **preinitiation complex** that is competent to initiate transcription as soon as nucleotides are available. This tight binding involves formation of an open promoter complex in which the DNA at the transcription start site has melted to allow the polymerase to read it. We will begin with the assembly of preinitiation complexes involving polymerase II. Even though these are by far the most complex, they are also the best studied. Once we see how the class II general transcription factors work, the class I and III mechanisms will be relatively easy to understand.

The Class II Preinitiation Complex

The class II preinitiation complex contains polymerase II and six general transcription factors named **TFIIA, TFIIB, TFIID, TFIIE, TFIIF,** and **TFIIH.** Many studies have shown that the class II general transcription factors and RNA polymerase II bind in a specific order to the growing preinitiation complex, at least in vitro. In particular, Danny Reinberg, as well as Phillip Sharp and their colleagues, performed DNA gel mobility shift and DNase and hydroxyl radical footprinting experiments (Chapter 5) that defined most of the order of factor binding in building the class II preinitiation complex.

Figure 11.1a presents the results of a gel mobility shift assay performed by Danny Reinberg and Jack Greenblatt

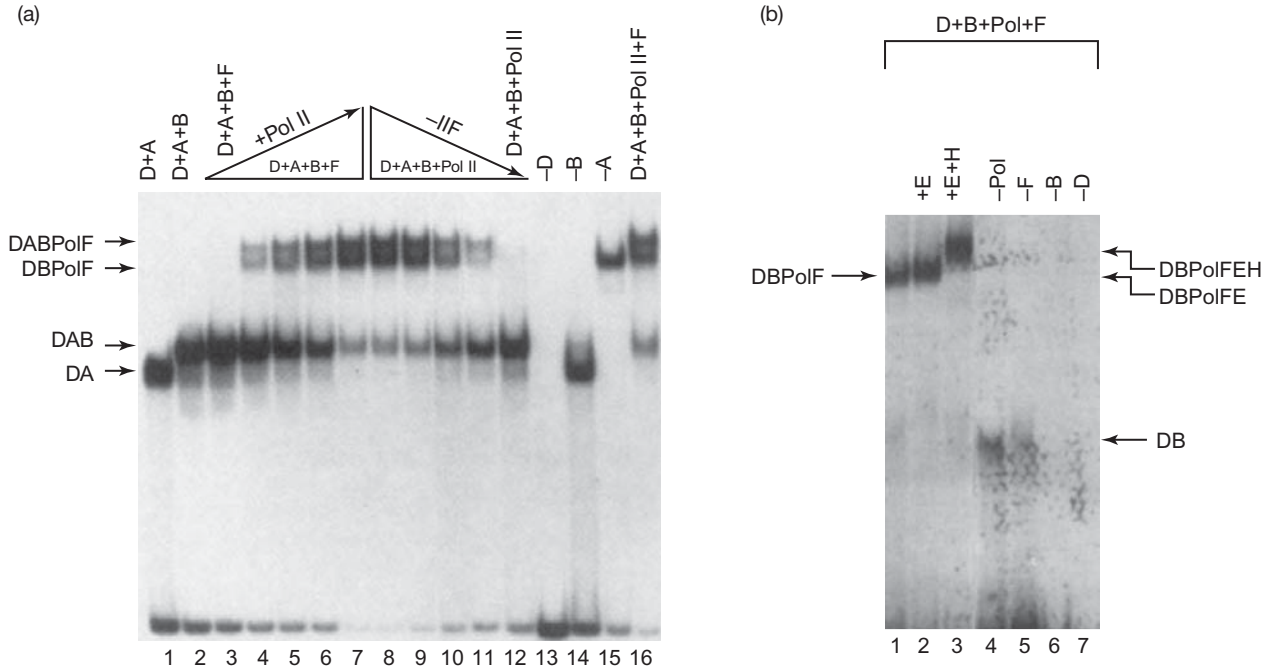


Figure 11.1 Building the preinitiation complex. (a) The DABPolF complex. Reinberg and colleagues performed gel mobility shift assays with TFIID, A, B, and F, and RNA polymerase II, along with labeled DNA containing the adenovirus major late promoter. Lane 1 shows the DA complex, formed with TFIID and A. Lane 2 demonstrates that adding TFIIB caused a new complex, DAB, to form. Lane 3 contained TFIID, A, B, and F, but it looks identical to lane 2. Thus, TFIIF did not seem to bind in the absence of polymerase II. Lanes 4–7 show what happened when the investigators added more and more polymerase II in addition to the four transcription factors: More and more of the large complexes, DABPolF and DBPolF, appeared. Lanes 8–11 contained less and less TFIIF, and we see less and less of the large complexes. Finally, lane 12 shows that essentially no DABPolF or DBPolF complexes formed when TFIIF was absent. Thus, TFIIF appears to bring polymerase II to the complex. The lanes on the right show what happened when Reinberg and colleagues left out one factor at a time. In lane 13, without TFIID, no complexes formed at all. Lane 14 shows that the DA complex, but no others, formed in the absence of TFIIB. Lane 15 demonstrates that DBPolF could still develop without TFIIA. Finally, all the large complexes appeared in the presence of all the factors (lane 16). **(b)** The DBPolFEH complex. Reinberg and colleagues started with the DBPolF complex (lacking TFIIA, lane 1) assembled on a labeled DNA containing the adenovirus major late promoter. Next, they added TFIIE, then TFIIH, in turn, and performed gel mobility shift assays. With each new transcription factor, the complex grew larger and its mobility decreased further. The mobilities of both complexes are indicated at right. Lanes 4–7 show again the result of leaving out various factors, denoted at the top of each lane. At best, only the DB complex forms. At worst, in the absence of TFIID, no complex at all forms. (*Sources:* (*a*) Flores, O., H. Lu, M. Killeen, J. Greenblatt, Z.F. Burton, and D. Reinberg, The small subunit of transcription factor IIF recruits RNA polymerase II into the preinitiation complex. *Proceedings of the National Academy of Sciences USA,* 88 (Nov 1991) p. 10001, f. 2a. (*b*) Cortes, P., O. Flores, and D. Reinberg. 1992. Factors involved in specific transcription by mammalian RNA polymerase II: Purification and analysis of transcription factor IIA and identification of transcription factor IIJ. *Molecular and Cellular Biology* 12: 413–21. American Society for Microbiology.)

and their colleagues using TFIIA, TFIID, TFIIB, and TFIIF, as well as RNA polymerase II. This experiment reveals the existence of four distinct complexes, which are labeled at the left of the figure. When the investigators added TFIID and A alone to DNA containing the adenovirus major late promoter, a *DA complex* formed (lane 1). When they added TFIIB in addition to D and A, a new, *DAB complex* formed (lane 2). The central part of the figure shows what happened when they added various concentrations of RNA polymerase II and TFIIF to the DAB complex. In lane 3, labeled D+A+B+F, all four of those factors were present, but RNA polymerase was missing. No difference was detectable between the complex formed with these four factors and the DAB complex. Thus, TFIIF does not seem to bind independently to DAB. But when the investigators added increasing amounts of polymerase (lanes 4–7), two new complexes appeared. These seem to include both polymerase and TFIIF, so the top complex is called the *DAB-PolF complex*. The other new complex *(DBPolF)* migrates somewhat faster because it is missing TFIIA, as we will see. After they had added enough polymerase to give a maximum amount of DABPolF, the investigators started decreasing the quantity of TFIIF (lanes 8–11). This reduction in TFIIF concentration decreased the yield of DABPolF, until, with no TFIIF but plenty of polymerase (lane 12), essentially no DABPolF (or DABPol) complexes formed. These data indicated that RNA polymerase and TFIIF are needed *together* to join the growing preinitiation complex.

Reinberg, Greenblatt, and colleagues assessed the order of addition of proteins by performing the same kind of mobility shift assays, but leaving out one or more factors at a time. In the most extreme example, lane 13, labeled −D, shows what happened when the investigators left out TFIID. No complexes formed, even with all the other factors present. This dependence on TFIID reinforced the hypothesis that TFIID is the first factor to bind; the binding of all the other factors depends on the presence of TFIID at the TATA box. Lane 14, marked −B, shows that TFIIB was needed to add polymerase and TFIIF. In the absence of TFIIB, only the DA complex could form. Lane 15, labeled −A, demonstrates that leaving out TFIIA made little difference. Thus, at least in vitro, TFIIA did not seem to be critical. Also, the fact that the band in this lane comigrated with the smaller of the two big complexes suggests that this smaller complex is DBPolF. Finally, the last lane contained all the proteins and displayed the large complexes as well as some residual DAB complex.

Reinberg and his coworkers extended this study in 1992 with TFIIE and H. Figure 11.1b demonstrates that they could start with the DBPolF complex and then add TFIIE and TFIIH in turn, producing a larger complex, with reduced mobility, with each added factor. The final preinitiation complex formed in this experiment was DBPolFEH. The last four lanes in this experiment show again that leaving out any of the early factors (polymerase II, TFIIF, TFIIB, or TFIID) prevents formation of the full preinitiation complex.

Thus, the order of addition of the general transcription factors (and RNA polymerase) to the preinitiation complex in vitro is as follows: TFIID (or TFIIA + TFIID), TFIIB, TFIIF + polymerase II, TFIIE, TFIIH. Now let us consider the question of where on the DNA each factor binds. Several groups, beginning with Sharp's, approached this question using footprinting. Figure 11.2 shows the results of a footprinting study on the DA and DAB complexes. Reinberg and colleagues used two different reagents to cut the protein–DNA complexes: 1,10-phenanthroline (OP)-copper ion complex, which creates hydroxyl radicals (lanes 1–4 in both panels), and DNase I (lanes 5–8 in both panels). Panel (a) depicts the data on the template strand, and panel (b) presents the results for the nontemplate strand. Panel (a), lanes 3 and 7 show that TFIID and A protect the TATA box. Lanes 3 and 7 in panel (b) show that the DA complex also protects the TATA box region on the nontemplate strand. Lanes 4 and 8 in panel (a) show no change in the template strand footprint after adding TFIIB to form the DAB complex. Essentially the same results were obtained with the nontemplate strand, but one subtle difference is apparent. As lane 8 shows, addition of TFIIB makes the DNA at position +10 even more sensitive to DNase. Thus, TFIIB does not seem to cover a significant expanse of DNA, but it does perturb the DNA structure enough to alter its susceptibility to DNase attack.

RNA polymerase II is a very big protein, so we would expect it to cover a large stretch of DNA and leave a big footprint. Figure 11.3 bears out this prediction. Whereas TFIID, A, and B protected the TATA box region (between positions −17 and −42) in the DAB complex, RNA polymerase II and TFIIF extended this protected region another 34 bases on the nontemplate strand, from position −17 to about position +17. Figure 11.4 summarizes what we have learned about the role of TFIIF in building the DABPolF complex. Polymerase II (red) and TFIIF (green) bind cooperatively, perhaps by forming a binary complex that joins the preformed DAB complex.

SUMMARY Transcription factors bind to class II promoters, including the adenovirus major late promoter, in the following order in vitro: (1) TFIID, apparently with help from TFIIA, binds to the TATA box, forming the DA complex. (2) TFIIB binds next. (3) TFIIF helps RNA polymerase bind to a region extending from at least position −34 to position +17. The remaining factors bind in this order: TFIIE and TFIIH, forming the DABPolFEH preinitiation complex. The participation of TFIIA seems to be optional in vitro.

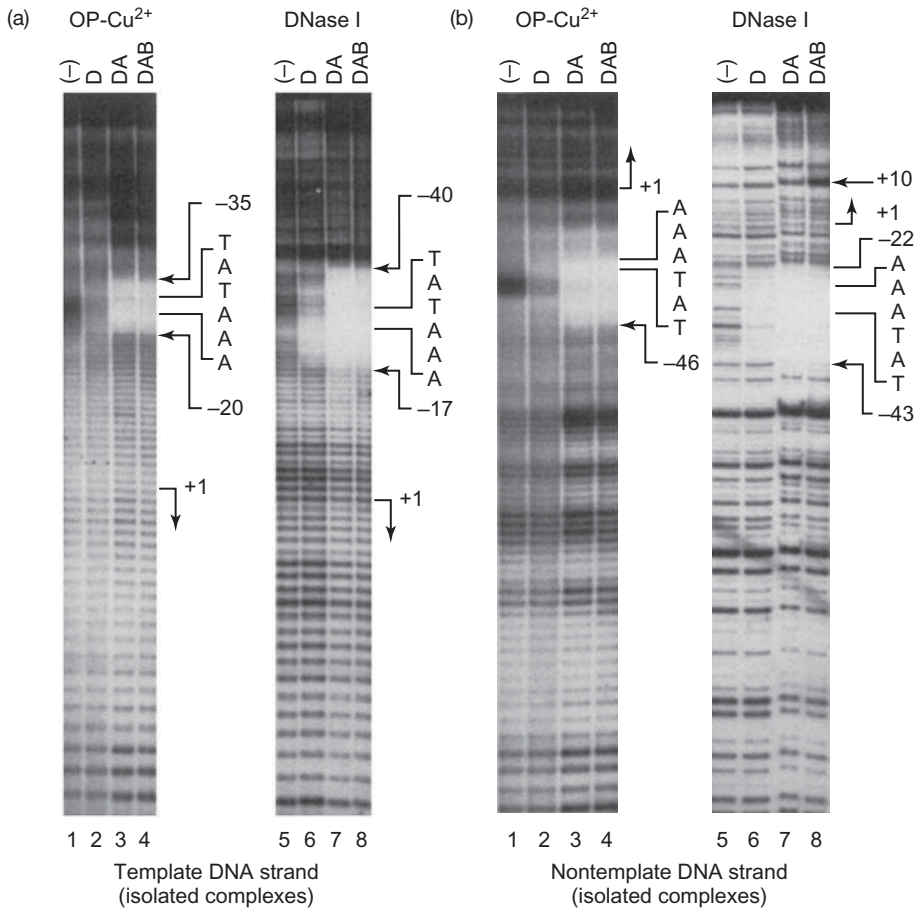


Figure 11.2 Footprinting the DA and DAB complexes. Reinberg and colleagues performed footprinting on the DA and DAB complexes with both DNase I (lanes 1–4) and another DNA strand breaker: a 1,10-phenanthroline-copper ion complex (OP-Cu^{2+}, lanes 5–8). **(a)** Footprinting on the template strand. The DA and DAB complexes formed on the TATA box (TATAAA, indicated at right, top to bottom). **(b)** Footprinting on the nontemplate strand. Again, the protected region in both the DA and DAB complexes was centered on the TATA box (TATAAA, indicated at right, bottom to top). The arrow near the top at right denotes a site of enhanced DNA cleavage at position +10. (*Source:* Adapted from Maldonado E., I. Ha, P. Cortes, L. Weiss, and D. Reinberg, Factors involved in specific transcription by mammalian RNA polymerase II: Role of transcription Factors IIA, IID, and IIB during formation of a transcription-competent complex. *Molecular and Cellular Biology* 10 (Dec 1990) p. 6344, f. 9. American Society for Microbiology.)

Structure and Function of TFIID

TFIID is a complex protein containing a **TATA-box-binding protein** (**TBP**) and 13 core **TBP-associated factors** (**TAFs,** or more specifically, **TAF_{II}s**). The subscript "II" was traditionally used when the context was unclear, because TBP also participates in transcription of class I and III genes and is associated with different TAFs (TAF_Is and TAF_{III}s) in class I and III preinitiation complexes, respectively. We will discuss the role of TBP and its TAFs in transcription from class I and III promoters later in this chapter. Let us first discuss the components of TFIID and their activities, beginning with TBP and concluding with the TAFs.

The TATA-Box-Binding Protein TBP, the first polypeptide in the TFIID complex to be characterized, is highly evolutionarily conserved: Organisms as disparate as yeast, fruit flies, plants, and humans have TATA-box-binding domains that are more than 80% identical in amino acid sequence. These domains encompass the carboxyl-terminal 180 amino acids of each protein and are very rich in basic amino acids. Another indication of evolutionary conservation is the fact that the yeast TBP functions well in a preinitiation complex in which all the other general transcription factors are mammalian.

Tjian's group demonstrated the importance of the carboxyl-terminal 180 amino acids of TBP when they showed by DNase I footprinting that a truncated form of human TBP containing only the carboxyl-terminal 180 amino acids of a human recombinant TBP is enough to bind to the TATA box region of a promoter, just as the native TFIID would.

How does the TBP in TFIID bind to the TATA box? The original assumption was that it acts like most other DNA-binding proteins (Chapter 9) and makes specific contacts with the base pairs in the major groove of the TATA box DNA. However, this assumption proved to be wrong.

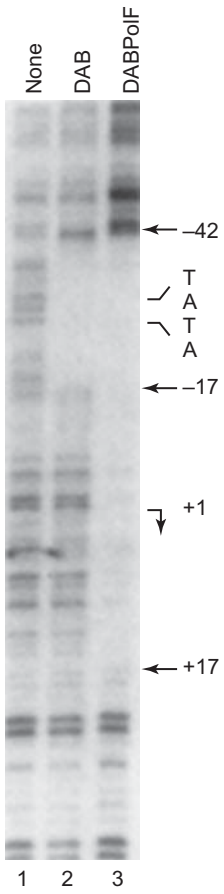


Figure 11.3 Footprinting the DABPolF complex. Reinberg and colleagues performed DNase footprinting with TFIID, A, and B (lane 2) and with TFIID, A, B, and F, and RNA polymerase II (lane 3). When RNA polymerase and TFIIF joined the complex, they caused a large extension of the footprint, to about position +17. This is consistent with the large size of RNA polymerase II. (*Source:* Flores O., H. Lu, M. Killeen, J. Greenblatt, Z.F. Burton, and D. Reinberg, The small subunit of transcription factor IIF recruits RNA polymerase II into the preinitiation complex. *Proceedings of the National Academy of Sciences USA* 88 (Nov 1991) p. 10001, f. 2b.)

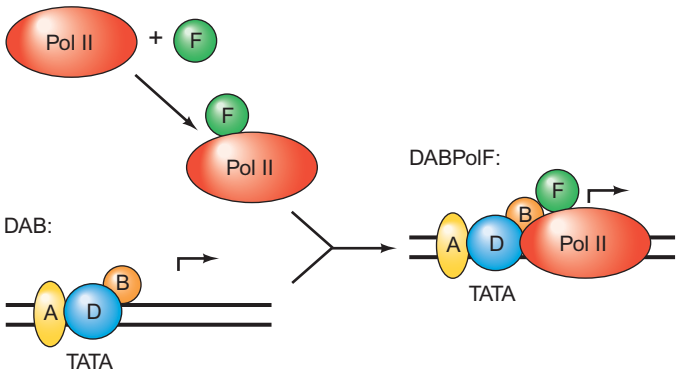


Figure 11.4 Model for formation of the DABPolF complex. TFIIF (green) binds to polymerase II (Pol II, red) and together they join the DAB complex. The result is the DABPolF complex. This model conveys the idea that polymerase II extends the DAB footprint in the downstream direction, and therefore binds to DNA downstream of the binding sites for TFIID, A, and B, which center on the TATA box.

Two research groups, headed by Diane Hawley and Robert Roeder, showed convincingly that the TBP in TFIID binds to the minor groove of the TATA box.

Barry Starr and Hawley changed all the bases of the TATA box, such that the major groove was changed, but the minor groove was not. This is possible because the hypoxanthine base in inosine (I) looks just like adenine (A) in the minor groove, but much different in the major groove (Figure 11.5a). Similarly, cytosine looks like thymine in the minor, but not the major, groove. Thus, Starr and Hawley made an adenovirus major late TATA box with all C's instead of T's, and all I's instead of A's (CICIII instead of TATAAA, Figure 11.5b). Then they measured TFIID binding to this CICI box and to the standard TATA box by a DNA mobility shift assay. As Figure 11.5c shows, the CICI box worked just as well as the TATA box, but a nonspecific

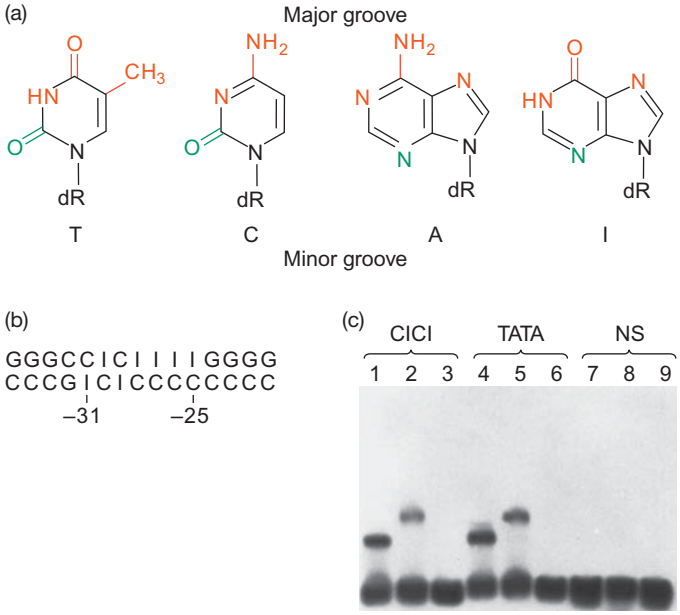


Figure 11.5 Effect of substituting C for T and I for A on TFIID binding to the TATA box. (a) Appearance of nucleosides as viewed from the major and minor grooves. Notice that thymidine and cytidine look identical from the minor groove (green, below), but quite different from the major groove (red, above). Similarly, adenosine and inosine look the same from the minor groove, but very different from the major groove. **(b)** Sequence of the adenovirus major late promoter (MLP) TATA box with C's substituted for T's and I's substituted for A's, yielding a CICI box. **(c)** Binding TBP to the CICI box. Starr and Hawley performed gel mobility shift assays using DNA fragments containing the MLP with a CICI box (lanes 1–3) or the normal TATA box (lanes 4–6), or a nonspecific DNA (NS) with no promoter elements (lanes 7–9). The first lane in each set (1, 4, and 7) contained yeast TBP; the second lane in each set (2, 5, and 8) contained human TBP; and the third lane in each set contained just buffer. The yeast and human TBPs gave rise to slightly different size protein–DNA complexes, but substituting a CICI box for the TATA box had little effect on the yield of the complexes. Thus, TBP binding to the TATA box was not significantly diminished by the substitutions. (*Source:* (*b–c*) Starr, D.B. and D.K. Hawley, TFIID binds in the minor groove of the TATA box. *Cell* 67 (20 Dec 1991) p. 1234, f. 2b. Reprinted by permission of Elsevier Science.)

DNA did not bind TFIID at all. Therefore, changing the bases in the TATA box did not affect TFIID binding as long as the minor groove was unaltered. This is strong evidence for binding of TFIID to the minor groove of the TATA box, and for no significant interaction in the major groove.

How does TFIID associate with the TATA box minor groove? Nam-Hai Chua, Roeder, and Stephen Burley and colleagues began to answer this question when they solved the crystal structure of the TBP of a plant, *Arabidopsis thaliana*. The structure they obtained was shaped like a saddle, complete with two "stirrups," which naturally suggested that TBP sits on DNA the way a saddle sits on a horse. The TBP structure has rough two-fold symmetry corresponding to the two sides of the saddle with their stirrups. Then, in 1993, Paul Sigler and colleagues and Stephen Burley and colleagues independently solved the crystal structure of TBP bound to a small synthetic piece of double-stranded DNA that contained a TATA box. That allowed them to see how TBP really interacts with the DNA, and it was not nearly as passive as a saddle sitting on a horse.

Figure 11.6 shows this structure. The curved undersurface of the saddle, instead of fitting neatly over the DNA, is roughly aligned with the long axis of the DNA, so its curvature forces the DNA to bend through an angle of 80 degrees. This bending is accomplished by a gross distortion in the DNA helix in which the minor groove is forced open. This opening is most pronounced at the first and last steps of the TATA box (between base pairs 1 and 2 and between base pairs 7 and 8). At each of those sites, two phenylalanine side chains from the stirrups of TBP intercalate, or insert, between base pairs, causing the DNA to kink. This distortion may help explain why the TATA sequence is so well conserved: The T–A step in a DNA double helix is relatively easy to distort, compared with any other dinucleotide step. This argument assumes that distortion of the TATA box is important to transcription initiation. Indeed, it is easy to imagine that peeling open the DNA minor groove aids the local DNA melting that is part of forming an open promoter complex.

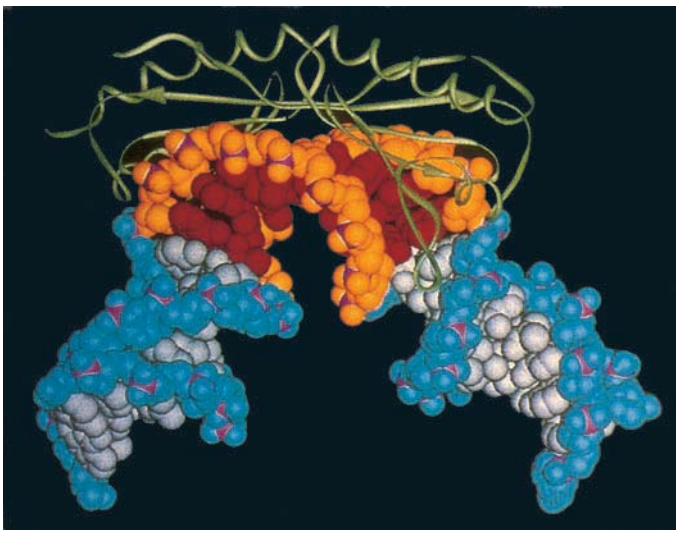


Figure 11.6 Structure of the TBP–TATA box complex. This diagram, based on Sigler and colleagues' crystal structure of the TBP–TATA box complex, shows the backbone of the TBP in olive at top. The long axis of the "saddle" is in the plane of the page. The DNA below the protein is in multiple colors. The backbones in the region that interacts with the protein are in orange, with the base pairs in red. Notice how the protein has opened up the narrow groove and almost straightened the helical twist in that region. One stirrup of the TBP is seen as an olive loop at right center, inserting into the minor groove. The other stirrup performs the same function, but it is out of view in back of the DNA. The two ends of the DNA, which do not interact with the TBP, are in blue and gray: blue for the backbones, and gray for the base pairs. The left end of the DNA sticks about 25 degrees out of the plane of the page, and the right end points inward by the same angle. The overall bend of about 80 degrees in the DNA, caused by TBP, is also apparent. (*Source:* Klug, A. Opening the gateway. *Nature* 365 (7 Oct 1993) p. 487, f. 2. © Macmillan Magazines Ltd.)

SUMMARY TFIID contains a 38-kD TATA-box-binding protein (TBP) plus several other polypeptides known as TBP-associated factors (TAFs). The C-terminal 180 amino acid fragment of the human TBP is the TATA-box-binding domain. The interaction between a TBP and a TATA box takes place in the DNA minor groove. The saddle-shaped TBP lines up with the DNA, and the underside of the saddle forces open the minor groove and bends the TATA box through an 80-degree curve angle.

The Versatility of TBP Molecular biology is full of wonderful surprises, and one of these is the versatility of TBP. This factor functions not only with polymerase II promoters that have a TATA box, but with TATA-less polymerase II promoters. Astonishingly, it also functions with TATA-less polymerase III promoters, *and* with TATA-less polymerase I promoters. In other words, TBP appears to be a universal eukaryotic transcription factor that operates at all promoters, regardless of their TATA content, and even regardless of the polymerase that recognizes them.

One indication of the widespread utility of TBP came from work by Ronald Reeder and Steven Hahn and colleagues on mutant yeasts with temperature-sensitive TBPs. We would have predicted that elevated temperature would block transcription by polymerase II in these mutants, but it also impaired transcription by polymerases I and III.

Figure 11.7 shows the evidence for this assertion. The investigators prepared cell-free extracts from wild-type and two different temperature-sensitive mutants, with lesions in TBP, as shown in Figure 11.7a. They made extracts from cells grown at 24°C and shocked for 1 h at 37°C, and from cells kept at the lower temperature. Then they added DNAs containing promoters recognized by all three polymerases and assayed transcription by S1 analysis. Figure 11.7b–e depicts the results. The heat shock had no effect on the wild-type extract, as expected (lanes 1 and 2). By contrast, the I143→N mutant extract could barely support transcription

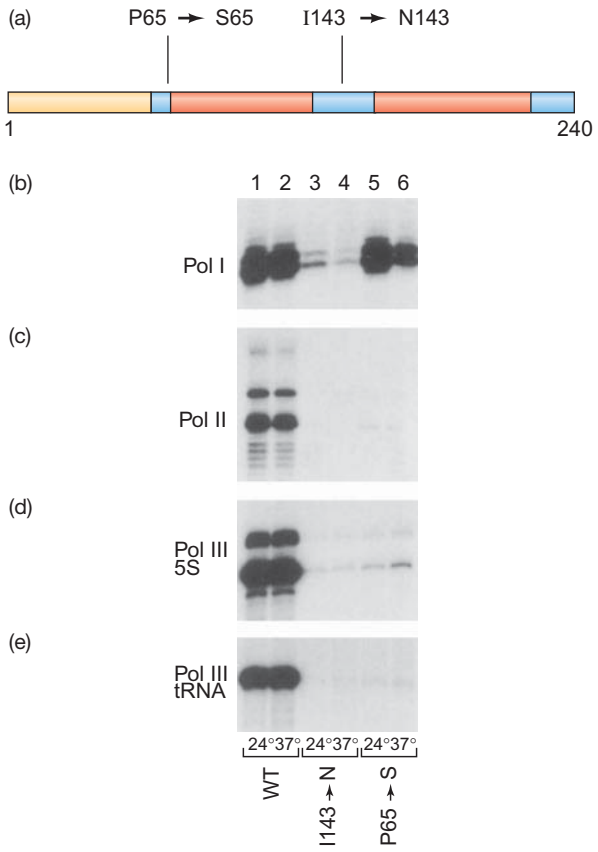


Figure 11.7 Effects of mutations in TBP on transcription by all three RNA polymerases. (a) Locations of the mutations. The blue and red regions indicate the conserved C-terminal domain of the TBP; red areas denote two repeated elements involved in DNA binding. The two mutations are: P65→S, in which proline 65 is changed to a serine; and I143→N, in which isoleucine 143 is changed to asparagine. **(b–e)** Effects of the mutations. Reeder and Hahn made extracts from wild-type or mutant yeasts, as indicated at bottom, and either heat-shocked them at 37°C or left them at 24°C, again as indicated at bottom. Then they tested these extracts by S1 analysis for ability to start transcription at promoters recognized by all three nuclear RNA polymerases: **(b)** the rRNA promoter (polymerase I); **(c)** the CYC1 promoter (polymerase II); **(d)** the 5S rRNA promoter (polymerase III); and **(e)** the tRNA promoter (also polymerase III). The I143→N extract was deficient in transcribing from all four promoters even when not heat-shocked. The P65→S extract was deficient in transcribing from polymerase II and III promoters, but could recognize the polymerase I promoter, even after heat shock. (*Source:* (*a*) Adapted from Schultz, M.C., R.H. Reeder, and S. Hahn. 1992. Variants of the TATA binding protein can distinguish subsets of RNA polymerase I, II, and III promoters. *Cell* 69:697–702.)

by any of the three polymerases, whether it was heat shocked or not (lanes 3 and 4). Clearly, the mutation in TBP was affecting not only polymerase II transcription, but transcription by the other two polymerases as well. The other mutant, P65→S, shows an interesting difference between the behavior of polymerase I and the other two polymerases. Whereas this mutant extract could barely support transcription of polymerase II and III genes, whether it had been heat shocked or not, it allowed wild-type levels of transcription by polymerase I if it was not heat-shocked, but heating reduced transcription by polymerase I by about twofold. Finally, wild-type TBP could restore transcription by all three polymerases in mutant extracts (data not shown).

Not only is TBP universally involved in eukaryotic transcription, it also seems to be involved in transcription in a whole different kingdom of organisms: the **archaea.** Archaea (formerly known as archaebacteria) are single-celled organisms that lack nuclei and usually live in extreme environments, such as hot springs or boiling hot deep ocean vents. They are as different from bacteria as they are from eukaryotes, and in several ways they resemble eukaryotes more than they do prokaryotes. In 1994, Stephen Jackson and colleagues reported that one of the archaea, *Pyrococcus woesei,* produces a protein that is structurally and functionally similar to eukaryotic TBP. This protein is presumably involved in recognizing the TATA boxes that frequently map to the 5′-flanking regions of archaeal genes. Moreover, a TFIIB-like protein has also been found in archaea. Thus, the transcription apparatus of the archaea bears at least some resemblance to that in eukaryotes, and suggests that the archaea and the eukaryotes diverged *after* their common ancestor diverged from the bacteria. This evolutionary scheme is also supported by the sequence of archaeal rRNA genes, which bear more resemblance to eukaryotic than to bacterial sequences.

SUMMARY Genetic studies have demonstrated that TBP mutant cell extracts are deficient, not only in transcription of class II genes, but also in transcription of class I and III genes. Thus, TBP is a universal transcription factor required by all three classes of genes. A similar factor has also been found in archaea.

The TBP-Associated Factors Many researchers have contributed to our knowledge of the TBP-associated factors (**TAFs**) in TFIIDs from several organisms. To identify TAFs from *Drosophila* cells, Tjian and his colleagues used an antibody specific for TBP to immunoprecipitate TFIID from a crude TFIID preparation. Then they treated the immunoprecipitate with 2.5 M urea to strip the TAFs off of the TBP–antibody precipitate and displayed the TAFs by SDS-PAGE. These and subsequent experiments have led to the identification of 13 TAFs associated with class II preinitiation complexes from a wide variety of organisms, from yeasts to humans.

These **core TAFs** were at first named according to their molecular masses, so the largest *Drosophila* TAF, with a molecular mass of 230 kD, was called $TAF_{II}230$, and the homologous human TAF was called $TAF_{II}250$. To avoid that kind of confusion, the core TAFs have been renamed according to their sizes, from largest to smallest, as TAF1 through TAF13. Thus, *Drosophila* $TAF_{II}230$, human $TAF_{II}250$, and fission yeast $TAF_{II}111$ are all now called TAF1. This nomenclature allows equivalent TAFs from different organisms to

be compared easily because they have the same names, regardless of their exact sizes. Note that the subscript II has been deleted. The context of the discussion should prevent confusion with class I and III TAFs. Some organisms encode TAF paralogs (homologous proteins in the same organism that have descended from a common ancestor protein). For example, we now know that human $TAF_{II}130/135$ and $TAF_{II}105$ are paralogs, so they are named TAF4 and TAF4b to indicate their homology. Some organisms encode TAF-like proteins that are similar, but not homologous to one of the core TAFs. These are given the designation L (for -like), as in TAF5L in humans and *Drosophila*. Some organisms (yeast and human, at least) have extra, non-core TAFs (TAF14 in yeast, and TAF15 in humans) that have no obvious homologs in other organisms.

Investigators have discovered several functions of the TAFs, but two that have received considerable attention are interaction with the promoter and interaction with gene-specific transcription factors. Let us consider the evidence for each of these functions and, where possible, the specific TAFs involved in each.

We have already seen the importance of the TBP in binding to the TATA box. But footprinting studies have indicated that the TAFs attached to TBP extend the binding of TFIID well beyond the TATA box in some promoters. In particular, Tjian and coworkers showed in 1994 that TBP protected the 20 bp or so around the TATA box in some promoters, but that TFIID protected a region extending to position +35, well beyond the transcription start site. This suggested that the TAFs in TFIID were contacting the initiator and downstream elements in these promoters.

To investigate this phenomenon in more detail, Tjian's group tested the abilities of TBP and TFIID to transcribe DNAs bearing two different classes of promoters in vitro. The first class (the adenovirus E1B and E4 promoters) contained a TATA box, but no initiator or downsteam promoter element (DPE). The second class (the adenovirus major late [AdML] promoter and the *Drosophila* heat shock protein [*hsp70*] promoter) contained a TATA box, an initiator, and a DPE. Figure 11.8 depicts the structures of these promoters, as well as the results of the in vitro transcription experiments. We can see that TBP and TFIID sponsored transcription equally well from the promoters that contained only the TATA box (compare lanes 1 and 2 and lanes 3 and 4). But TFIID had a decided advantage in sponsoring transcription from the promoters that also had an initiatior and DPE (compare lanes 5 and 6 and lanes 7 and 8). Thus, TAFs apparently help TBP facilitate transcription from promoters with initiators and DPEs.

Which TAFs are responsible for recognizing the initiator and DPE? To find out, Tjian and colleagues performed a photo-cross-linking experiment with *Drosophila* TFIID and a radioactively labeled DNA fragment containing the *hsp70* promoter. They incorporated bromodeoxyuridine (BrdU) into the promoter-containing DNA, then allowed TFIID to

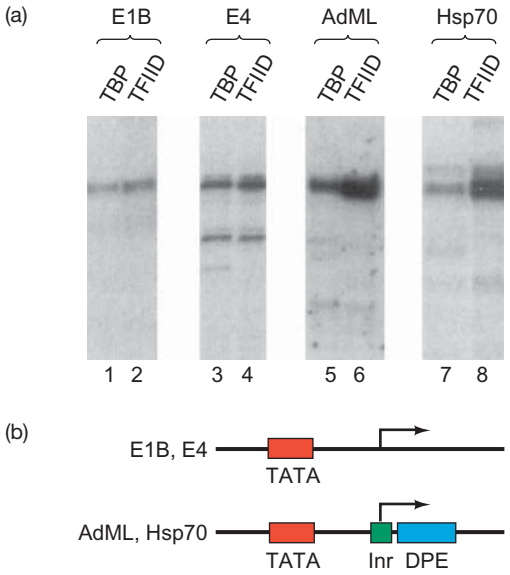


Figure 11.8 Activities of TBP and TFIID on four different promoters. (a) Experimental results. Tjian and colleagues tested a reconstituted *Drosophila* transcription system containing either TBP or TFIID (indicated at top) on templates bearing four different promoters (also as indicated at top). The promoters were of two types diagrammed in panel **(b).** The first type, represented by the adenovirus E1B and E4 promoters, contained a TATA box (red). The second type, represented by the adenovirus major late promoter (AdML) and the *Drosophila hsp70* promoter, contained a TATA box plus an initiator (Inr, green) and a DPE (blue). After transcription in vitro, Tjian and coworkers assayed the RNA products by primer extension (top). The autoradiographs show that TBP and TFIID fostered transcription equally well from the first type of promoter (TATA box only), but that TFIID worked much better than TBP in supporting transcription from the second type of promoter (TATA box plus Inr plus DPE). (*Source:* Verrijzer, C.P., J.-L. Chen, K. Yokomari, and R. Tjian, Promoter recognition by TAFs. *Cell* 81 (30 June 1995) p. 1116, f. 1. Reprinted with permission of Elsevier Science.)

bind to the promoter, then irradiated the complexes with UV light to cross-link the protein to the BrdU in the DNA. After washing away unbound protein, the investigators digested the DNA with nuclease to release the proteins, then subjected the labeled proteins to SDS-PAGE. Figure 11.9, lane 1, shows that two TAFs (TAF1 and TAF2) bound to the *hsp70* promoter and thereby became labeled. When TFIID was omitted (lane 2), no proteins became labeled. Following up on these findings, Tjian and coworkers reconstituted a ternary complex containing only TBP, TAF1, and TAF2 and tested it in the same photo-cross-linking assay. Lane 3 shows that this experiment also yielded labeled TAF1 and TAF2, and lane 4 shows that TBP did not become labeled when it was bound to the DNA by itself. We know that TBP binds to this TATA-box-containing DNA, but it does not become cross-linked to BrdU and therefore does not become labeled. Why not? Probably because this kind of photo-cross-linking works well only with proteins that bind in the major groove, and TBP binds in the minor groove of DNA.

To double-check the binding specificity of the ternary complex (TBP–TAF1–TAF2), Tjian and colleagues performed

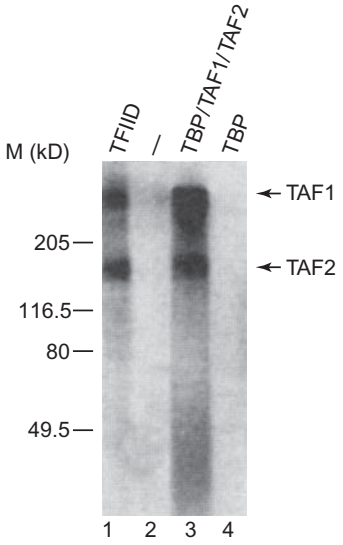


Figure 11.9 Identifying the TAFs that bind to the *hsp70* promoter. Tjian and colleagues photo-cross-linked TFIID to a ^{32}P-labeled template containing the *hsp70* promoter as follows: First, they bound the TFIID to the labeled template, which had also been substituted with the photosensitive nucleoside bromodeoxyuridine (BrdU). Next, these investigators irradiated the TFIID–DNA complex with UV light to form covalent bonds between the DNA and any proteins in close contact with the major groove of the DNA. Next, they digested the DNA with nuclease and subjected the proteins to SDS-PAGE. Lane 1 of the autoradiograph shows the results when TFIID was the input protein. TAF1 and TAF2 became labeled, implying that these two proteins had been in close contact with the labeled DNA's major groove. Lane 2 is a control with no TFIID. Lane 3 shows the results when a ternary complex containing TBP, TAF1, and TAF2 was the input protein. Again, the two TAFs became labeled, suggesting that they bound to the DNA. Lane 4 shows the results when TBP was the input protein. It did not become labeled, which was expected because it does not bind in the DNA major groove. (*Source:* Verrijzer, C.P., J.-L. Chen, K. Yokomari, and R. Tjian, *Cell* 81 (30 June 1995) p. 1117, f. 2a. Reprinted with permission of Elsevier Science.)

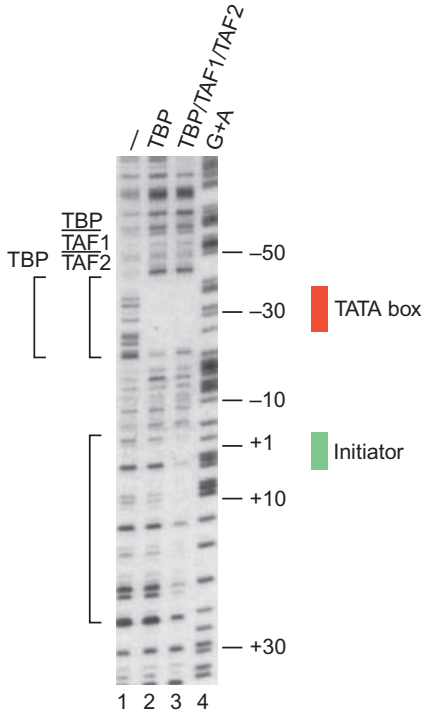


Figure 11.10 DNase footprinting the *hsp70* promoter with TBP and the ternary complex (TBP, TAF1, and TAF2). Lane 1, no protein; lane 2, TBP; lane 3, ternary complex. In both lanes 2 and 3, TFIIA was also added to stabilize the DNA–protein complexes, but separate experiments indicated that it did not affect the extent of the footprints. Lane 4 is a G+A sequencing lane used as a marker. The extents of the footprints caused by TBP and the ternary complex are indicated by brackets at left. The locations of the TATA box and initiator are indicated by boxes at right. (*Source:* Verrijzer, C.P., J.-L. Chen, K. Yokomori, and R. Tjian, *Cell* 81 (30 June 1995) p. 1117, f. 2c. Reprinted with permission of Elsevier Science.)

a DNase footprinting experiment with TBP or the ternary complex. Figure 11.10 shows that TBP caused a footprint only in the TATA box, whereas the ternary complex caused an additional footprint in the initiator and downstream sequences. This reinforced the hypothesis that the two TAFs bind at least to the initiator, and perhaps to the DPE.

Further experiments with binary complexes (TBP–TAF1 or TBP–TAF2) showed that these complexes were no better than TBP alone in recognizing initiators and DPEs. Thus, both TAFs seem to cooperate in enhancing binding to these promoter elements. Furthermore, the ternary complex (TBP–TAF1–TAF2) is almost as effective as TFIID in recognizing a synthetic promoter composed of the AdML TATA box and the TdT initiator. By contrast, neither binary complex functions any better than TBP in recognizing this promoter. These findings support the hypothesis that TAF1 and TAF2 cooperate in binding to the initiator alone, as well as to the initiator plus a DPE.

The TBP part of TFIID is of course important in recognizing the majority of the well-studied class II promoters, which contain TATA boxes (Figure 11.11a). But what about promoters that lack a TATA box? Even though these promoters cannot bind TBP directly, most still depend on this transcription factor for activity. The key to this apparent paradox is the fact that these TATA-less promoters contain other elements that ensure the binding of TBP. These other elements can be initiators and DPEs, to which TAF1 and TAF2 can bind and thereby secure the whole TFIID to the promoter (Figure 11.11b). Or they can be upstream elements that bind gene-specific transcription factors, which in turn interact with one or more TAFs to anchor TFIID to the promoter. For example, the activator Sp1 binds to proximal promoter elements (GC boxes) and also interacts with at least one TAF (TAF4). This bridging activity apparently helps TFIID bind to the promoter (Figure 11.11c).

The second major activity of the TAFs is to participate in the transcription stimulation provided by activators, some of which we will study in Chapter 12. Tjian and colleagues demonstrated in 1990 that TFIID is sufficient to

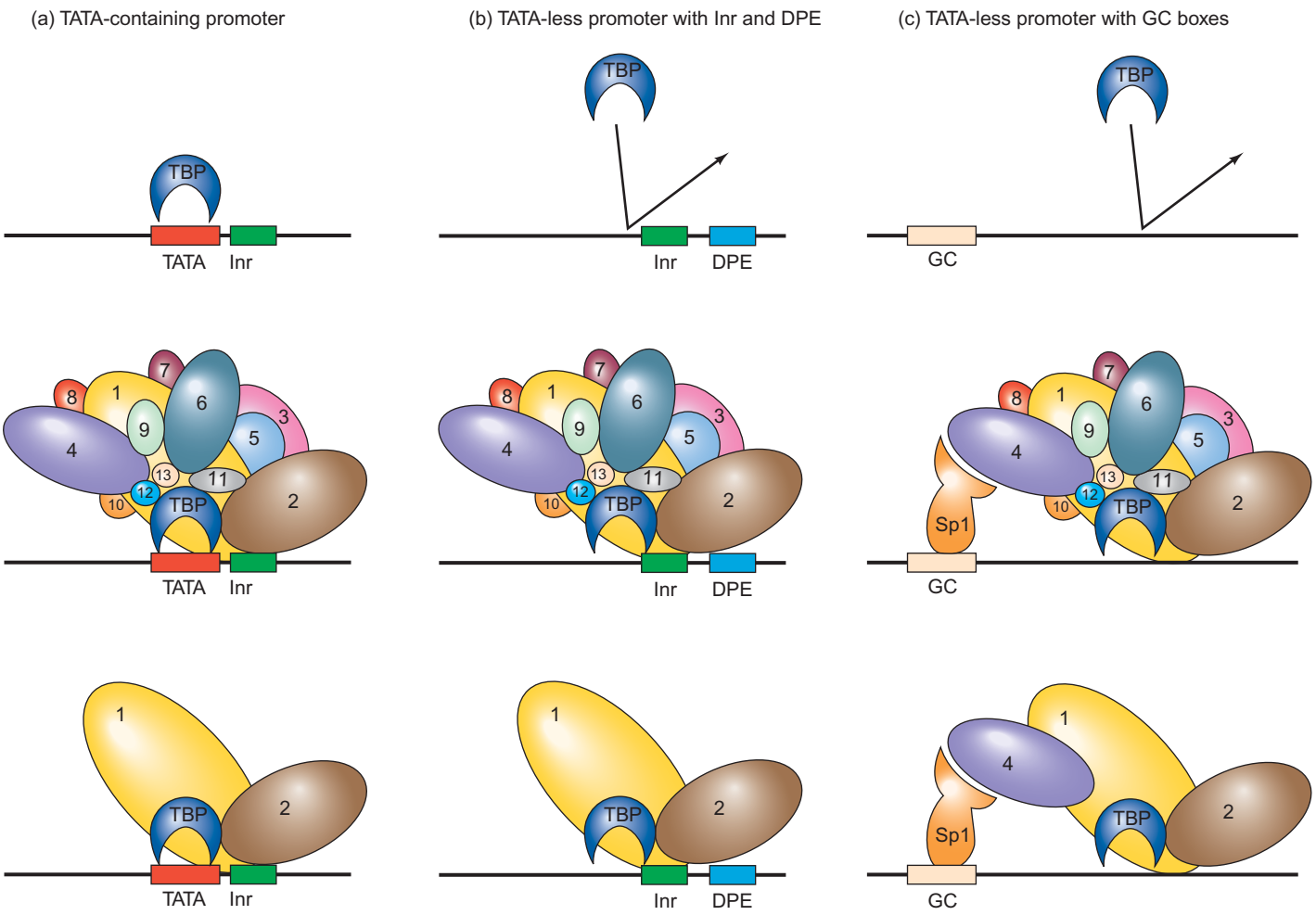


Figure 11.11 Model for the interaction between TBP and TATA-containing or TATA-less promoters. **(a)** TATA-containing promoter. TBP can bind by itself to the TATA box of this promoter (top). It can also bind in the company of all the TAFs in TFIID (middle). And it can bind with a subset of TAFs (bottom). **(b)** TATA-less promoter with initiator element and DPE. TBP cannot bind by itself to this promoter, which contains no TATA box (top). The whole TFIID is competent to bind to the TATA-less promoter through interactions between TAF1 (yellow) and TAF2 (brown, middle). TAF1 and TAF2 are sufficient to tether TBP to the initiator and DPE (bottom). **(c)** TATA-less promoter with GC boxes. TBP cannot bind to this promoter by itself (top). The whole TFIID can bind to this promoter through interactions with Sp1 bound at the GC boxes (middle). TAF1, TAF2, and TAF4 are sufficient to anchor TBP to the Sp1 bound to the GC boxes. (*Source:* Adapted from Goodrich, J.A., G. Cutter, and R. Tjian, Contacts in context: Promoter specificity and macromolecular interactions in transcription. *Cell* 84:826, 1996.)

participate in such stimulation by the factor Sp1, but TBP is not. These results suggest that some factors in TFIID are necessary for interaction with upstream-acting factors such as Sp1 and that these factors are missing from TBP. By definition, these factors are TAFs, and they are sometimes called **coactivators.**

We have seen that mixing TBP with subsets of TAFs can produce a complex with the ability to participate in transcription from certain promoters. For example, the TBP–TAF1–TAF2 complex functioned almost as well as the whole TFIID in recognizing a promoter composed of a TATA box and an initiator. Tjian and colleagues used a similar technique to discover which TAFs are involved in activation by Sp1. They found that activation by Sp1 in *Drosophila* or human extracts occurred only when TAF4 was present. Thus, TBP and TAF1 plus TAF2 were sufficient for basal transcription, but could not support activation by Sp1. Adding TAF4 in addition to the other two factors and TBP allowed Sp1 to activate transcription.

Tjian and colleagues also showed that Sp1 binds directly to TAF4, but not to TAF1 or TAF2. They built an affinity column containing GC boxes and Sp1 and tested it for the ability to retain the three TAFs. As predicted, only TAF4 was retained.

Using the same strategy, Tjian and colleagues demonstrated that another activator, NTF-1, binds to TAF2 and requires either TAF1 and TAF2 or TAF1 and TAF6 to activate transcription in vitro. Thus, different activators work with different combinations of TAFs to enhance transcription, and all of them seem to have TAF1 in common. This suggests that TAF1 serves as an assembly factor around which other TAFs can aggregate. These findings are compatible with the model in Figure 11.12: Each activator interacts with a particular subset of TAFs, so the holo-TFIID can

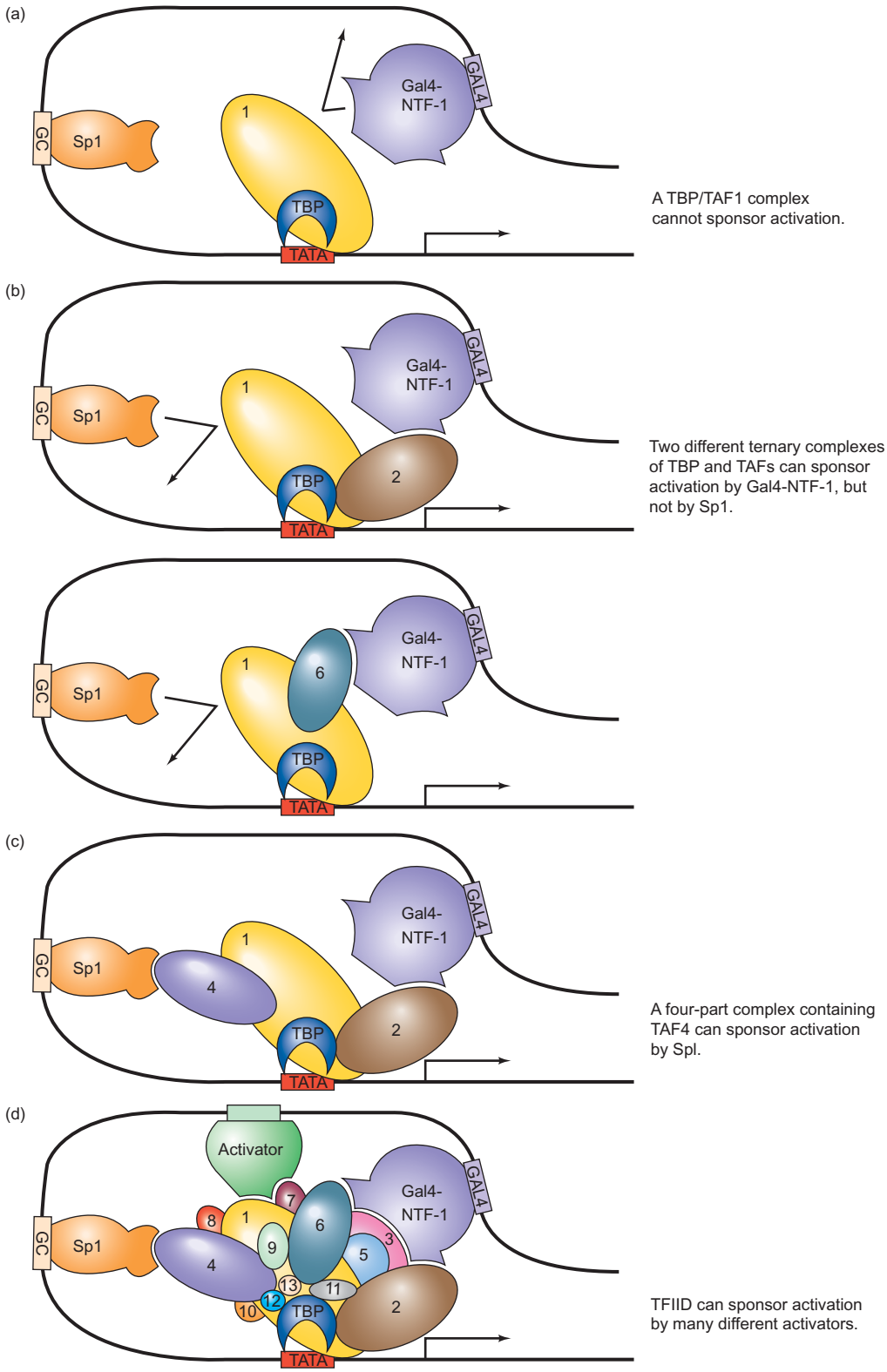


Figure 11.12 A model for transcription enhancement by activators. (a) TAF1 does not interact with either Sp1 or Gal4-NTF-1 (a hybrid activator with the transcription-activating domain of NTF-1), so no activation takes place. **(b)** Gal4-NTF-1 can interact with either TAF2 or TAF6 and activate transcription; Sp1 cannot interact with either of these TAFs or with TAF1 and does not activate transcription. **(c)** Gal4-NTF-1 interacts with TAF2 and Sp1 interacts with TAF4, so both factors activate transcription. **(d)** Holo-TFIID contains the complete assortment of TAFs, so it can respond to a wide variety of activators, represented here by Sp1, Gal4-NTF-1, and a generic activator (green) at top. (*Source:* Adapted from Chen, J.L., L.D. Attardi, C.P. Verrijzer, K. Yokomori, and R. Tjian, Assembly of recombinant TFIID reveals differential coactivator requirements for distinct transcriptional activators. *Cell* 79:101, 1994).

interact with several activators at once, magnifying their effect and producing strong enhancement of transcription.

In addition to their abilities to interact with promoter elements and activators, TAFs can have enzymatic activities. The best studied of these is TAF1, which has two known enzymatic activities. It is a histone acetyltransferase (HAT), which attaches acetyl groups to lysine residues of histones. Such acetylation is generally a transcription-activating event. We will study this process in greater detail in Chapter 13. TAF1 is also a protein kinase that can phosphorylate itself and TFIIF (and TFIIA and TFIIE, though to a lesser extent). These phosphorylation events may modulate the efficiency of assembly of the preinitiation complex.

Despite early indications that it was not required for preinitiation complex formation in vitro, TFIIA is essential for TBP (or TFIID) binding to promoters. Much evidence leads to this conclusion, but one experiment is particularly easy to describe: Mutations in either of the genes encoding the two subunits of TFIIA in yeast are lethal.

TFIIA not only stabilizes TBP-TATA box binding, it also stimulates TFIID-promoter binding by an *antirepression* mechanism, as follows: When TFIID is not bound to a promoter, the DNA-binding surface of TBP is covered by the N-terminal domain of TAF1, which inhibits TFIID binding to the promoter. But TFIIA can interfere with the interaction between the TAF1 N-terminal domain and the DNA-binding surface of TBP, freeing up TBP for binding to the promoter.

SUMMARY TFIID contains 13 TAFs, in addition to TBP. Most of these TAFs are evolutionarily conserved in the eukaryotes. The TAFs serve several functions, but two obvious ones are interacting with core promoter elements and interacting with activators. TAF1 and TAF2 help TFIID bind to the initiator and DPEs of promoters and therefore can enable TBP to bind to TATA-less promoters that contain such elements. TAF1 and TAF4 help TFIID interact with Sp1 that is bound to GC boxes upstream of the transcription start site. These TAFs therefore ensure that TBP can bind to TATA-less promoters that have GC boxes. Different combinations of TAFs are apparently required to respond to various activators, at least in higher eukaryotes. TAF1 also has two enzymatic activities. It is a histone acetyltransferase and a protein kinase.

Exceptions to the Universality of TAFs and TBP Genetic studies in yeast call into question the generality of the model in Figure 11.12. Michael Green and Kevin Struhl and their colleagues independently discovered that mutations in yeast TAF genes were lethal, but transcription activation was not affected, at least not in the first genes studied. For example, Green and colleagues made temperature-sensitive mutations in the gene encoding yeast TAF1. At the nonpermissive temperature, they found that there was a rapid decrease in the concentration of TAF1, and at least two other yeast TAFs. The loss of TAF1 apparently disrupted the TFIID enough to cause the degradation of other TAFs. However, in spite of these losses of TAFs, the in vivo transcription rates of five different yeast genes activated by a variety of activators were unaffected at the nonpermissive temperature. These workers obtained the same results with another mutant in which the TAF14 gene had been deleted. By contrast, when the genes encoding TBP or an RNA polymerase subunit were mutated, all transcription quickly ceased.

Green, Richard Young, and colleagues followed up these initial studies with a genome-wide analysis of the effects of mutations in two TAF genes, as well as several other yeast genes. They made temperature-sensitive mutations in TAF1 and in TAF9. Then they used high-density oligonucleotide arrays (such as those described in Chapter 25) to determine the extent of expression of each of 5460 yeast genes at an elevated temperature at which the mutant TAF was inactive and at a lower temperature at which the mutant TAF was active. These arrays contained oligonucleotides specific for each gene. Total yeast RNA can then be hybridized to these arrays, and the extent of hybridization to each oligonucleotide is a measure of the extent of expression of the corresponding gene. The investigators compared the hybridization of RNA to each oligonucleotide at low and high temperature and compared the response with the results of a similar analysis of a temperature-sensitive mutation in the largest subunit of RNA polymerase II (Rpb1). Because the latter mutation prevented transcription of *all* class II genes, it provided a baseline with which to compare the effects of mutations in other genes.

Table 11.1 presents the results of this analysis. It is striking that only 16% of the yeast genes analyzed were as dependent on TAF1 as they were on Rpb 1, indicating that TAF1 is required for transcription of only 16% of yeast genes. This is not what we would expect if the TAFs are essential parts of TFIID, and TFIID is an essential part of the preinitiation complexes formed at all class II genes.

Table 11.1 Whole Genome Analysis of Transcription Requirements in Yeast

General Transcription Factor (Subunit)	Fraction of Genes Dependent on Subunit Function (%)
TFIID (TAF1)	16
TFIID (TAF9)	67
TFIIE (Tfa1)	54
TFIIH (Kin28)	87

Indeed, TAF1, along with TBP, had been regarded as a keystone of TFIID, helping to assemble all the other TAFs in that factor, but this view is clearly not supported by the genome-wide expression analysis. Instead, TAF1 and its homolog in higher organisms appear to be required in the preinitiation complexes formed at only a subset of genes. In yeast, these genes tend to be ones governing progression through the cell cycle.

Mutation of the other yeast TAF (TAF9) had a more pronounced effect. Sixty-seven percent of the yeast genes analyzed were as dependent on this TAF as they were on Rpb1. But that does not mean that TFIID is required for transcription of all these genes, because TAF9 is also part of a transcription adapter complex known as **SAGA** (named for three classes of proteins it contains—SPTs, ADAs, and GCN5—and its enzymatic activity, histone acetyltransferase). Like TFIID, SAGA contains TBP, a number of TAFs, and histone acetyltransferase activity, and appears to mediate the effects of certain transcription activator proteins. So the effect of mutating TAF9 may be due to its role in SAGA or perhaps in other protein complexes yet to be discovered, rather than in TFIID.

Not only are some TAFs not universally required for transcription, the TFIIDs appear to be heterogenous in their TAF compositions. For example, TAF10 is found in only a fraction of human TFIIDs, and its presence correlates with responsiveness to estrogen.

Even more surprisingly, TBP is not universally found in preinitiation complexes in higher eukaryotes. The most celebrated example of an alternative TBP is **TRF1** (**TBP-related factor 1**) in *Drosophila melanogaster*. This protein is expressed in developing neural tissue, binds to TFIIA and TFIIB, and stimulates transcription just as TBP does, and it has its own group of TRF-associated factors called **nTAFs** (for neural TAFs). In 2000, Michael Holmes and Robert Tjian used primer extension analysis in vivo and in vitro to show that TRF1 stimulates transcription of the *Drosophilia tudor* gene. Furthermore, this analysis revealed that the *tudor* gene has two distinct promoters. The first is a downstream promoter with a TATA box recognized by a complex including TBP. The second promoter lies about 77 bp upstream of the first and has a TC box recognized by a complex including TRF1 (Figure 11.13). The TC box extends from position −22 to −33 with respect to the start of transcription and has the sequence ATTGCTTTTCTT in the nontemplate strand. It is protected by a complex of TRF1, TFIIA, and TFIIB in DNase footprinting experiments. However, none of these proteins alone make a footprint in this region, and neither does TBP, or TBP plus TFIIA and TFIIB.

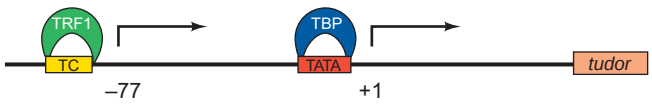


Figure 11.13 The *Drosophila tudor* control region. This gene has two promoters about 77 bp apart. The downstream promoter has a TATA box that attracts a preinitiation complex based on TBP. The upstream promoter has a TC box that attracts a preinitiation complex based on TRF1.

Thus, TRF appears to be a cell type-specific variant of TBP. The presence of alternative TBPs and TAFs raises the possibility that gene expression in higher eukaryotes could be controlled in part by the availability of the appropriate TBP and TAFs, as well as by the activator proteins we will study in Chapter 12. Indeed, the recognition of two different *tudor* promoters by two different TBPs is reminiscent of the recognition of two different prokaryotic promoters for the same gene by RNA polymerases bearing different σ-factors, as we saw in Chapter 8.

Actually, TRF appears to be unique to *Drosophila*. But another **TBP-like factor** (**TLF**) has been found in all multicellular animals investigated to date. TLF differs from TBP in lacking the pairs of phenylalanines that intercalate between base pairs in TATA boxes and help bend the DNA at the promoter. Accordingly, TLF appears not to bind to TATA boxes and may direct transcription at other, TATA-less promoters.

The central role of TBP in forming preinitiation complexes has been further challenged by the discovery of a **TBP-free TAF-containing complex** (**TFTC**) that is able to sponsor preinitiation complex formation without any help from TFIID or TBP. Structural studies by Patrick Schultz and colleagues have provided some insight into how TFTC can substitute for TFIID. They have performed electron microscopy and digital image analysis on both TFTC and TFIID and found that they have strikingly similar three-dimensional structures. Figure 11.14 shows three-dimensional models of the two protein complexes in three different orientations. The most obvious characteristics of both complexes is a groove large enough to accept a double-stranded DNA. In fact, it appears that the protein of both complexes would encircle the DNA and hold it like a clamp. The only major difference between the two complexes is the projection at the top of TFTC due to domain 5. TFIID lacks both the projection and domain 5.

In Chapter 10 we learned that many promoters in *Drosophila* lack a TATA box; instead, they have a DPE, usually coupled with an initiator element (Inr). We also learned that the DPE can attract TFIID through one or more of its TAFs. In 2000, James Kadonaga and colleagues also discovered a factor in *Drosophila* (dNC2) that is homologous to a factor from other organisms known as **NC2** (**negative cofactor 2**) or Dr1-Drap1. For simplicity's sake, we can refer to all such factors as NC2. Kadonaga and colleagues also made the interesting discovery that NC2 can discriminate between TATA box-containing promoters and DPE-containing promoters. In fact, NC2 stimulates transcription from DPE-containing promoters and represses transcription from TATA

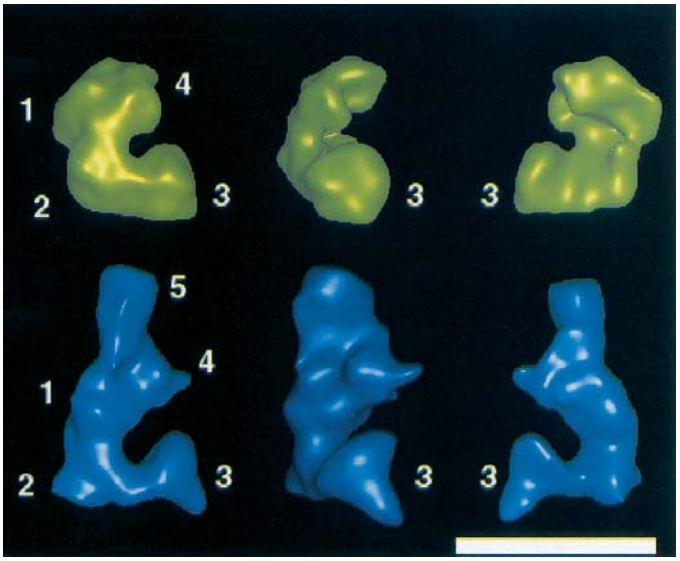


Figure 11.14 Three-dimensional models of TFIID and TFTC. Schultz and colleagues made negatively stained electron micrographs (see Chapter 19, for method) of TFIID and TFTC, then digitally combined images to arrive at an average. Then they tilted the grid in the microscope and analyzed the resulting micrographs to glean three-dimensional information for both proteins. The resulting models for TFIID (green) and TFTC (blue) are shown. (*Source:* Brand, M., C. Leurent, V. Mallouh, L. Tora, and P. Schuttz, Three-dimensional structures of the TAF_{II}-containing complexes TFDIID and TFTC. *Science* 286 (10 Dec 1999) f. 3, p. 2152. Copyright © AAAS.)

box-containing promoters. Thus, NC2 may be a focal point of gene regulation.

The crystal structure of an NC2–TATA box–TBP complex, determined by Stephen Burley and colleagues in 2001, shows how NC2 can inhibit transcription from TATA box-containing promoters. It binds to the underside of the DNA that has been bent by the saddle-shaped TBP. Once NC2 has bound to the promoter, one of its α-helices blocks TFIIB from joining the complex, and another part of NC2 interferes with TFIIA binding. Without TFIIA or TFIIB, the preinitiation complex cannot form and transcription cannot initiate.

SUMMARY The TAFs do not appear to be universally required for transcription of class II genes. Even TAF1 is not required for transcription of the great majority of yeast class II genes. Even TBP is not universally required. Some promoters in higher eukaryotes respond to an alternative protein such as TRF1 and not to TBP. Some promoters can be stimulated by a TBP-free TAF-containing complex (TFTC), rather than by TFIID. The general transcription factor NC2 stimulates transcription from DPE-containing promoters but represses transcription from TATA-containing promoters.

Structure and Function of TFIIB

Danny Reinberg and his coworkers cloned and expressed the gene for human TFIIB. This cloned TFIIB product can substitute for the authentic human protein in all in vitro assays, including response to activators such as Sp1. This suggests that TFIIB is a single-subunit factor (M_r = 35 kD) that requires no auxiliary polypeptides such as the TAFs. As we have already discovered, TFIIB is the third general transcription factor to join the preinitiation complex in vitro (after TFIID and A), or the second if TFIIA has not yet bound. It is essential for binding RNA polymerase because the polymerase–TFIIF complex will bind to the DAB complex, but not to the DA complex.

The position of TFIIB between TFIID and TFIIF/RNA polymerase II in the assembly of the preinitiation complex suggests that TFIIB is part of the measuring device that places RNA polymerase II in the proper position to initiate transcription. If so, TFIIB should have two domains: one to bind to each of these proteins. Indeed, TFIIB does have two domains: an N-terminal domain (**$TFIIB_N$**), and a C-terminal domain (**$TFIIB_C$**). Subsequent structural work in 2004 by Roger Kornberg and colleagues revealed that these two domains really do function to bridge between TFIID at the TATA box and RNA polymerase II so as to position the active center of the polymerase about 26–31 bp downstream of the TATA box, just where transcription should begin. In particular, this work showed that TBP, by bending the DNA at the TATA box, wraps the DNA around $TFIIB_C$, and that $TFIIB_N$ binds to a site on the polymerase that positions the enzyme correctly at the transcription initiation site.

Kornberg and colleagues crystallized a complex of RNA polymerase II and TFIIB from budding yeast *(Saccharomyces cerevisiae)*. Figure 11.15 shows two views of the structure of this complex, along with the positions of TBP and promoter DNA inferred from previous work. We can see the two domains of TFIIB in this complex. $TFIIB_C$ (magenta) appears to interact with TBP and DNA at the TATA box. Indeed, the DNA bent by TBP at the TATA box appears to wrap around $TFIIB_C$ and the polymerase. After the bend, the DNA extends straight toward $TFIIB_N$, which lies near the active site of the polymerase.

Previous studies had shown that mutations in $TFIIB_N$ altered the start site of transcription, and the present work provides a rationale for those findings. In particular, it was known that mutations in residues 62–66 cause changes in the initiation site. These amino acids lie on the side of a finger domain in $TFIIB_N$ that appears to contact bases −6 to −8, relative to the start site at +1, in the DNA template strand (top left in Figure 11.16). Moreover, the tip of the finger approaches the active center of the polymerase, and lies near the initiator region of the promoter (Chapter 10), which surrounds the transcription start site.

In the human TFIIB, the fingertip contains two basic residues (lysine), which could bind well to the DNA at the

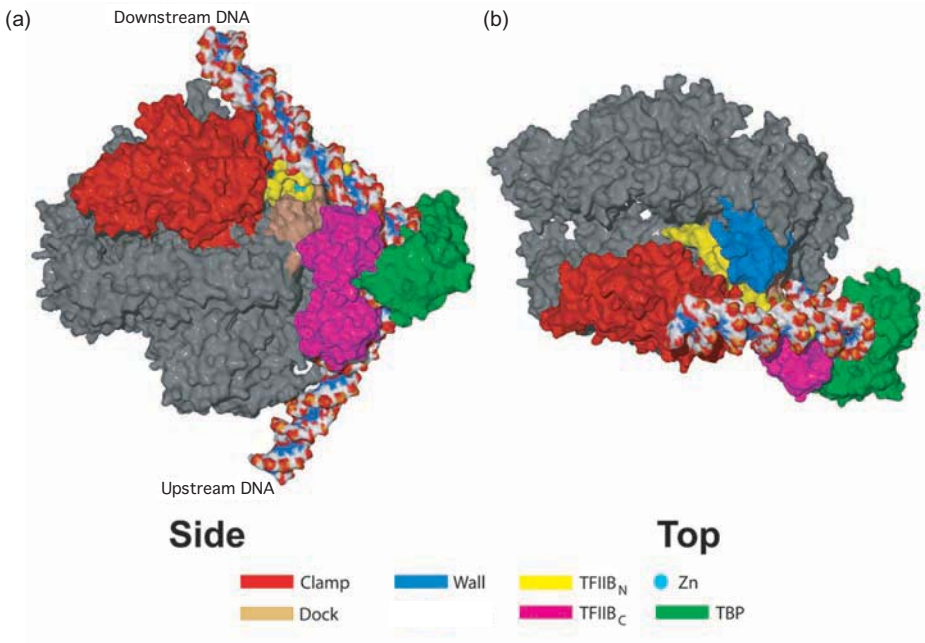


Figure 11.15 A model for the TFIIB–TBP–polymerase II-DNA structure. (a) and **(b)** show two different views of the structure, which Kornberg and colleagues inferred from separate structures of TFIIB$_C$–TBP–TATA box DNA and RNA polymerase II-TFIIB. The color key at bottom identifies TBP, the domains of TFIIB, and domains of the polymerase that interact with TFIIB. Other regions of the polymerase are in gray. The bent TATA box DNA, with 20-bp B-form DNA extensions, is in red, white, and blue. (*Source: (a–b)* Reprinted with permission from *Science,* Vol. 303, David A. Bushnell, Kenneth D. Westover, Ralph E. Davis, Roger D. Kornberg, "Structural Basis of Transcription: An RNA Polymerase II-TFIIB Cocrystal at 4.5 Angstroms" Fig. 3 c&d, p. 986. Copyright 2004, AAAS.)

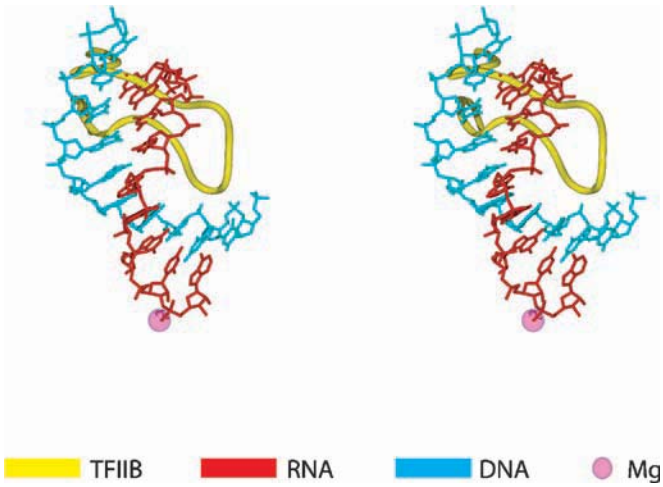


Figure 11.16 Stereo view of the interaction between the B finger of TFIIB$_N$, the DNA template strand, and the RNA product. The elements of the structure are identified by the color key at bottom. (*Source:* Reprinted with permission from *Science,* Vol 303, David A. Bushnell, Kenneth D. Westover, Ralph E. Davis, Roger D. Kornberg, "Structural Basis of Transcription: An RNA Polymerase II-TFIIB Cocrystal at 4.5 Angstroms" Fig. 4, p. 987. Copyright 2004, AAAS.)

initiator, thus positioning the start of transcription there. However, these two basic amino acids are replaced by acidic amino acids in yeast TFIIB, and initiator sequences do not exist in yeast promoters. These considerations may help explain why the human preinitiation complex can successfully position the start of transcription approximately 25–30 bp downstream of the TATA box, whereas transcription initiation is much more variable (40–120 bp downstream of the TATA box) in yeast.

Kornberg and colleagues concluded that TFIIB plays a dual role in positioning the transcription start site. First, it achieves coarse positioning by binding via its TFIIB$_C$ domain to TBP at the TATA box and binding to RNA polymerase via the finger and an adjacent zinc ribbon in the TFIIB$_N$ domain. In most eukaryotes, this places the polymerase in position to start about 25–30 bp downstream of the TATA box. Then, upon DNA unwinding, TFIIB achieves fine positioning by interacting with DNA at, and just upstream of, the initiator via the finger of TFIIB$_N$. Notice that TFIIB not only determines the start site of transcription, it also determines the *direction* of transcription. That is because its asymmetry of binding to the promoter—with its C-terminal domain upstream and its N-terminal domain downstream—establishes an asymmetry to the preinitiation complex, which in turn establishes the direction of transcription.

The importance of TFIIB and RNA polymerase II in establishing the transcription start site is underscored by the following experiment. In the budding yeast *Saccharomyces cerevisiae,* the start site is about 40 to 120 nt downstream of the TATA box, whereas in the fission yeast *Saccharomyces pombe,* it is about 25 to 30 nt downstream of the TATA box. However, when *S. pombe* TFIIB and RNA polymerase II

were mixed with the other general transcription factors from *S. cerevisiae,* initiation occurred 25 to 30 nt downstream of the TATA box. And the reverse experiment also worked: *S. cerevisiae* TFIIB and RNA polymerase II, mixed with the other factors from *S. pombe,* dictated transcription initiation 40 to 120 nt downstream of the TATA box.

A similar measuring mechanism appears to apply to the archaea. Transcription in archaea requires a basal transcription apparatus composed of a multisubunit RNA polymerase, an arachaeal TBP, and **transcription factor B (TFB)**, which is homologous to eukaryotic TFIIB. Stephen Bell and Stephen Jackson showed in 2000 that the transcription start site, relative to the TATA box in the archaeon *Sulfolobus acidocaldarius,* is determined by RNA polymerase and TFB.

The model presented in Figure 11.15 is appealing, but it is cobbled together from partial structures, so we are left wondering how closely it corresponds to the structure we would see in an intact preinitiation complex. To probe this question, Hung-Ta Chen and Steven Hahn used a combination of photo-cross-linking and **hydroxyl radical probing** to map the interactions between domains of yeast TFIIB and domains of yeast RNA polymerase II.

Hydroxyl radical probing uses the following strategy: The experimenters introduce cysteine residues into one protein by site-directed mutagenesis (Chapter 5). To each cysteine in turn, they attach an iron-EDTA (ethylenediamine tetraacetate) complex known as Fe-BABE, which can generate hydroxyl radicals that can cleave protein chains within about 15 Å. After cleavage, the protein fragments can be displayed by gel electrophoresis and detected by Western blotting. This procedure identifies any regions of a second protein lying within 15 Å of a given cysteine on the first protein.

In their first experiment, Chen and Hahn changed several amino acids in the finger and linker regions of TFIIB to cysteines, which were then linked to Fe-BABE. After assembling preinitiation complexes with these modified TFIIB molecules, they activated hydroxyl radical formation to cleave proteins in close proximity to the cysteines in the finger and linker regions of TFIIB. To facilitate Western blotting, they attached an epitope (FLAG) to the end of either Rpb1 or Rpb2, so they could use anti-FLAG antibodies to probe their Western blots. Figure 11.17a–c shows the results of the Western blots probed with anti-FLAG antibody when the FLAG epitope was placed at the N- or C- terminus of Rpb2, or the C-terminus of Rpb1. The novel bands created by hydroxyl radical cleavage (not found in lanes with no substituted cysteines [wt] or no Fe-BABE [–]) are marked with brackets.

These bands contain protein fragments of known length, and we know that they include either the protein's N-terminus or C-terminus because they are detected by an anti-FLAG antibody, and the FLAG epitope is attached to a protein terminus. Thus, the cleavage sites could be mapped to locations on the known crystal structure of the protein. Figure 11.17d presents a similar experiment, except that no FLAG epitope was used, and the blot was probed with an antibody against a natural epitope in the N-terminal 200 residues of Rpbl.

Using this information, Chen and Hahn mapped the parts of Rpb1 and Rpb2 that were in close contact with the cysteine attached to the Fe-BABE in each case. Figure 11.17e and f depict the maps of cleavages caused by TFIIB variants with cysteines introduced into the finger and linker regions, respectively. Dark blue and light blue regions denote strong and moderate-to-weak cleavage, respectively. These are the regions of Rpb1 and Rpb2 that are in close contact with the finger and linker regions of TFIIB. The similarities of these maps suggests that the finger and linker regions of TFIIB are close together in the preinitiation complex. Furthermore, as predicted, this part of TFIIB ($TFIIB_N$) does indeed contact RNA polymerase II. In particular, it contacts sites in the protrusion, wall, clamp, and fork regions of the polymerase, which are near the active center.

In their photo-cross-linking experiments, Chen and Hahn linked an ^{125}I-tagged photo-cross-linking reagent called PEAS to the cysteines in the same TFIIB cysteine variants used in the hydroxyl radical probing. After assembling preinitiation complexes with these derivatized TFIIBs, they irradiated the complexes to form covalent cross-links, then observed the cross-links by SDS-PAGE and autoradiography to detect the ^{125}I tags. As expected, they found that the TFIIB finger and linker domains cross-linked to RNA polymerase II. However, they also discovered something unexpected: The TFIIB finger and linker domains also cross-linked to the largest subunit of TFIIF, placing this polypeptide close to the active center of polymerase II.

SUMMARY Structural studies on a TFIIB-polymerase II complex show that TFIIB binds to TBP at the TATA box via its C-terminal domain, and to polymerase II via its N-terminal domain. This bridging action effects a coarse positioning of the polymerase active center about 25–30 bp downstream of the TATA box. In mammals, a loop motif of the N-terminal domain of TFIIB effects a fine positioning of the start of transcription by interacting with the single-stranded template DNA strand very near the active center. Biochemical studies confirm that the TFIIB N-terminal domain (the finger and linker domains, in particular) lies close to the RNA polymerase II active center, and to the largest subunit of TFIIF, in the preinitiation complex.

Structure and Function of TFIIH

TFIIH is the last general transcription factor to join the preinitiation complex. It appears to play two major roles in transcription initiation; one of these is to phosphorylate

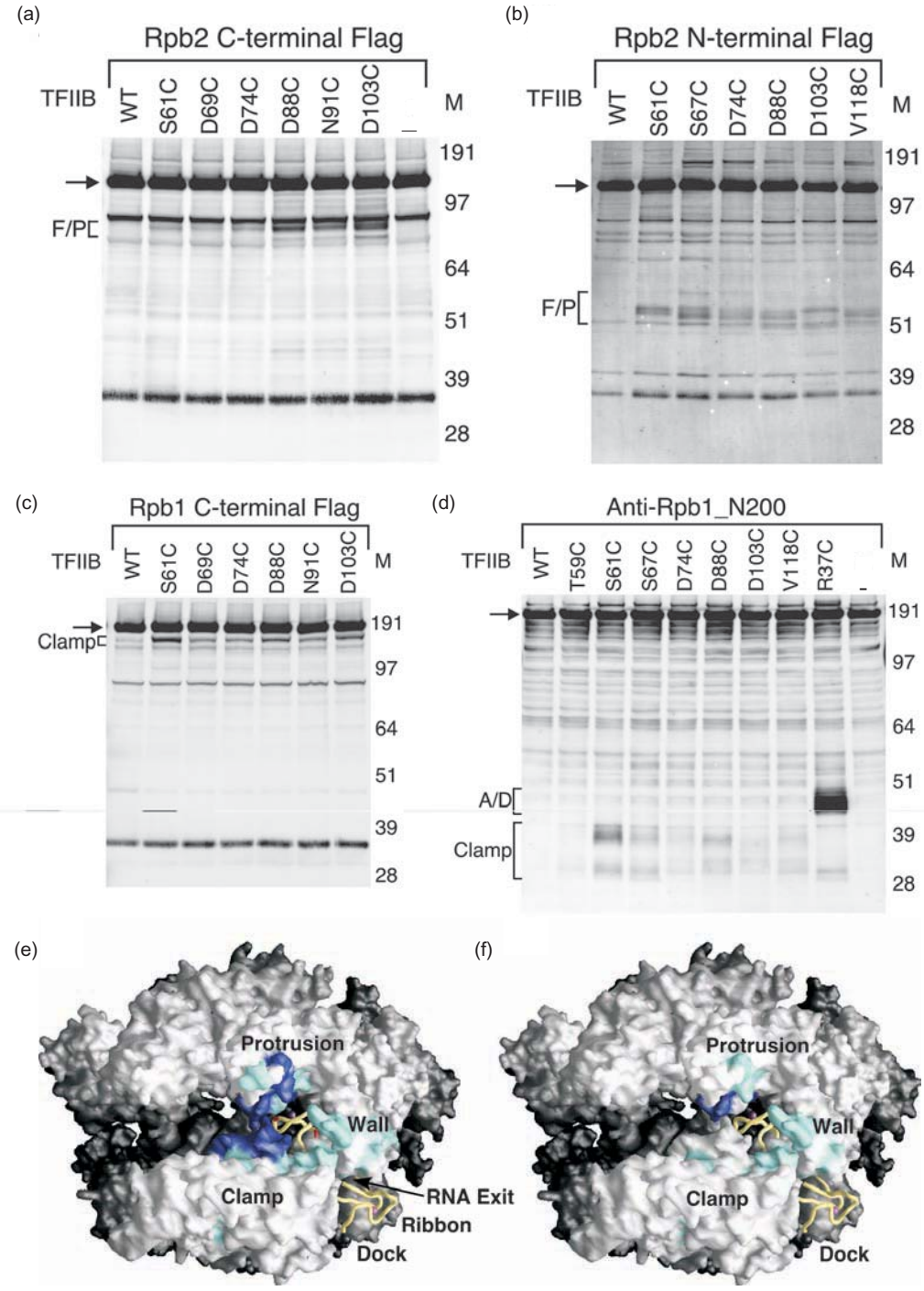


Figure 11.17 Mapping contacts between TFIIB and RNA polymerase II in the yeast preinitiation complex. (a–d) Chen and Hahn attached Fe-BABE hydroxyl-radical-generating reagents to cysteines that had been substituted for other amino acids (positions indicated at tops of lanes) in the finger and linker domains of TFIIB. Then they formed preinitiation complexes that included these substituted TFIIBs and RNA polymerases whose Rpb2 C-terminus (a) or N-terminus (b), or whose Rpbl C-terminus (c) had been tagged with the FLAG epitope, as indicated at the top of each gel. Then they activated hydroxyl radical formation to cleave proteins within about 15 Å of the cysteine in the TFIIB. Then they performed SDS-PAGE on the preinitiation complex proteins and protein fragments, and on proteins from complexes that did not contain substituted cysteines (wt), or did not contain TFIIB complexed with Fe-BABE (−). They blotted the protein bands and visualized them by probing the blots with an anti-FLAG antibody (a–c) or with an antibody against a natural epitope in the terminal 200 amino acids of Rpbl. The novel bands (brackets) that do not appear in the control lanes (wt and −) represent polypeptide fragments generated by hydroxyl radical cleavage. The lengths of these fragments, compared to markers (M), together with the knowledge that they contain one of the ends of either Rpbl or 2, allows the cleavage site to be determined to within four amino acids on either side. The locations of these cleavage sites are identified beside each bracket: clamp; F/P (fork and protrusion); or A/D (active site and dock regions). **(e)** and **(f)** Mapping the cleavage sites to the known crystal structure of the yeast RNA polymerase II when TFIIB contained substituted cysteines in the finger domain (e), or the linker domain (f). Dark blue represents strong cleavages, and light blue represents weak to moderate cleavages. To take account of the error inherent in the method, the color was spread out over nine amino acids, centered on the apparent cleavage site. (*Source: (a–f)* Reprinted from *Cell,* Vol. 119, Hung-Ta Chen and Steven Hahn, "Mapping the Location of TFIIB within the RNA Polymerase II Transcription Preinitiation Complex: A Model for the Structure of the PIC," pp. 169–180, fig 2, p. 172. Copyright 2004 with permission from Elsevier.)

the CTD of RNA polymerase II. The other is to unwind DNA at the transcription start site to create the "transcription bubble."

Phosphorylation of the CTD of RNA Polymerase II As we have already seen in Chapter 10, RNA polymerase II exists in two physiologically meaningful forms: IIA (unphosphorylated) and IIO (with many phosphorylated serines in the carboxyl-terminal domain [CTD]). The unphosphorylated enzyme, polymerase IIA, is the form that joins the preinitiation complex. But the phosphorylated enzyme, polymerase IIO, carries out RNA chain elongation. This behavior suggests that phosphorylation of the polymerase occurs between the time it joins the preinitiation complex and the time promoter clearance occurs. In other words, phosphorylation of the polymerase could be the trigger that allows the polymerase to shift from initiation to elongation mode. This hypothesis receives support from the fact that the unphosphorylated CTD in polymerase IIA binds much more tightly to TBP than does the phosphorylated form in polymerase IIO. Thus, phosphorylation of the CTD could break the tether that binds the polymerase to the TBP at the promoter and thereby permit transcription elongation to begin. On the other hand, this hypothesis is damaged somewhat by the finding that transcription can sometimes occur in vitro without phosphorylation of the CTD.

Whatever the importance of CTD phosphorylation, Reinberg and his colleagues have demonstrated that TFIIH was a good candidate for the protein kinase that catalyzes this process. First, these workers showed that the purified transcription factors, by themselves, are capable of phosphorylating the CTD of polymerase II, converting polymerase IIA to IIO. The evidence, shown in Figure 11.18 came from a gel mobility shift assay. Lanes 1–6 demonstrate that adding ATP had no effect on the mobility of the DAB, DABPolF, or DABPolFE complexes. On the other hand, after TFIIH was added to form the DABPolFEH complex, ATP produced a change to lower mobility. What accounted for this change? One possibility is that one of the transcription factors in the complex had phosphorylated the polymerase. Indeed, when Reinberg and colleagues isolated the polymerase from the lower mobility complex, it proved to be the phosphorylated form, polymerase IIO. But polymerase IIA had been added to the complex in the first place, so one of the transcription factors had apparently performed the phosphorylation.

Next Reinberg and colleagues demonstrated directly that the TFIIH preparation phosphorylates polymerase IIA. To do this, they incubated purified polymerase IIA and TFIIH together with [γ-^{32}P]ATP under DNA-binding conditions. A small amount of polymerase phosphorylation occurred, as shown in Figure 11.19a. Thus, this TFIIH preparation by itself is capable of carrying out the phosphorylation. By contrast, all the other factors together caused no such phosphorylation. However, these factors

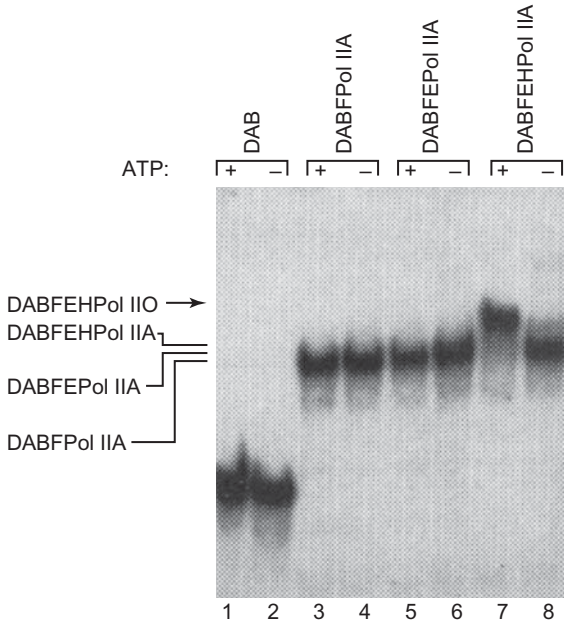


Figure 11.18 Phosphorylation of preinitiation complexes. Reinberg and colleagues performed gel mobility shift assays with preinitiation complexes DAB through DABPolFEH, in the presence and absence of ATP, as indicated at top. Only when TFIIH was present did ATP shift the mobility of the complex (compare lanes 7 and 8). The simplest explanation is that TFIIH promotes phosphorylation of the input polymerase (polymerase IIA) to polymerase IIO. (*Source:* Lu, H., I. Zawel, L. Fisher, J.M. Egly, and D. Reinberg, Human general transcription factor IIH phosphorylates the C-terminal domain of RNA polymerase II. *Nature* 358 (20 Aug 1992) p. 642, f. 1. Copyright © Macmillan Magazines Ltd.)

could greatly stimulate the phosphorylating capability of TFIIH. Lanes 6–9 show the results with TFIIH plus an increasing set of the other factors. As Reinberg and associates added each new factor, they noticed an increasing efficiency of phosphorylation of the polymerase and accumulation of polymerase IIO. Because the biggest increase in polymerase IIO labeling came with the addition of TFIIE, these workers performed a time-course study in the presence of TFIIH or TFIIH plus TFIIE. Figure 11.19b shows that the conversion of the IIa subunit to the IIo subunit was much more efficient when TFIIE was present. Figure 11.19c shows the same results graphically.

We know that the CTD of the polymerase IIa subunit is the site of the phosphorylation because polymerase IIB, which lacks the CTD, is not phosphorylated by the TFIIDBFEH complex, while polymerase IIA, and to a lesser extent, polymerase IIO, are phosphorylated (Figure 11.20a). Also, as we have seen, phosphorylation produces a polypeptide that coelectrophoreses with the IIo subunit, which does have a phosphorylated CTD. To demonstrate directly the phosphorylation of the CTD, Reinberg and colleagues cleaved the phosphorylated enzyme with chymotrypsin, which cuts off the CTD, and electrophoresed the products. The autoradiograph of the chymotrypsin

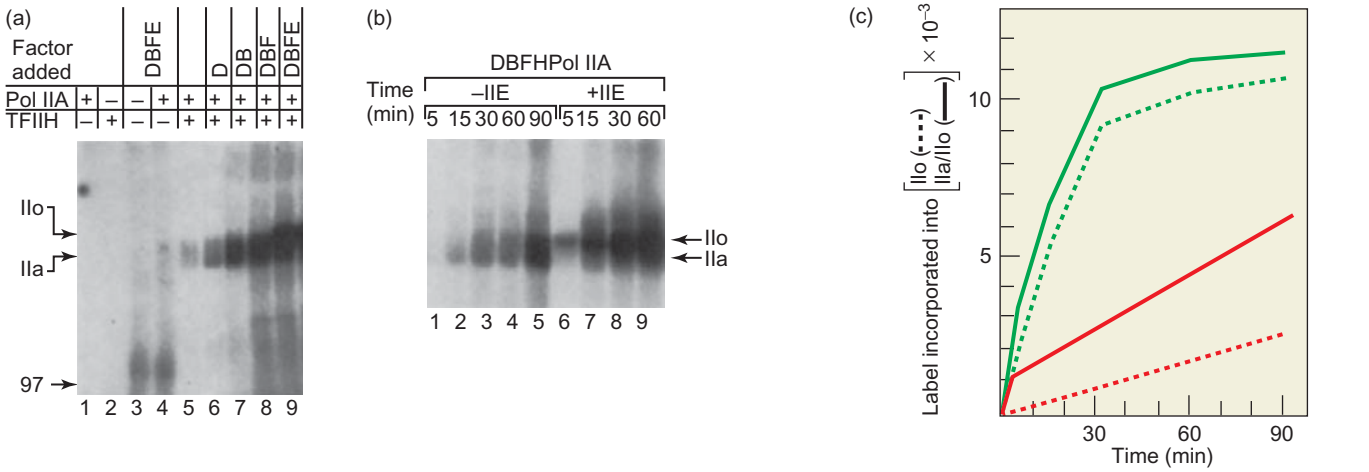


Figure 11.19 TFIIH phosphorylates RNA polymerase II. **(a)** Reinberg and colleagues incubated polymerase IIA (containing the hypophosphorylated subunit IIa) with various mixtures of transcription factors, as shown at top. They included [γ-^{32}P]ATP in all reactions to allow phosphorylation of the polymerase, then electrophoresed the proteins and performed autoradiography to visualize the phosphorylated polymerase. Lane 4 shows that TFIID, B, F, and E, were insufficient to cause phosphorylation. Lanes 5–9 demonstrate that TFIIH alone is sufficient to cause some polymerase phosphorylation, but that the other factors enhance the phosphorylation. TFIIE provides particularly strong stimulation of phosphorylation of the polymerase IIa subunit to IIo. **(b)** Time course of polymerase phosphorylation. Reinberg and colleagues performed the same assay for polymerase phosphorylation with TFIID, B, F, and H in the presence or absence of TFIIE, as indicated at top. They carried out the reactions for 60 or 90 min, sampling at various intermediate times, as shown at top. Arrows at right mark the positions of the two polymerase subunit forms. Note that polymerase phosphorylation is more rapid in the presence of TFIIE. **(c)** Graphic presentation of the data from panel (b). Green and red curves represent phosphorylation in the presence and absence, respectively, of TFIIE. Solid lines and dotted lines correspond to appearance of phosphorylated polymerase subunits IIa and IIo, or just IIo, respectively. (*Source:* Adapted from Lu, H., I. Zawel, L. Fisher, J.-M. Egly, and D. Reinberg, Human general transcription factor IIH phosphorylates the C-terminal domain of RNA polymerase II. *Nature* 358 (20 Aug 1992) p. 642, f. 2. Copyright © Macmillan Magazines Ltd.)

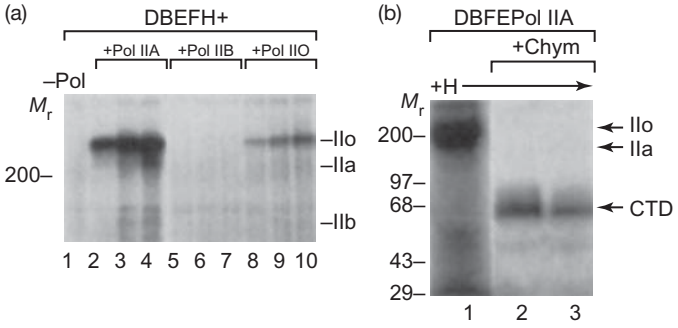


Figure 11.20 TFIIH phosphorylates the CTD of polymerase II. **(a)** Reinberg and colleagues phosphorylated increasing amounts of polymerases IIA, IIB, or IIO, as indicated at top, with TFIID, B, F, E, and H and radioactive ATP as described in Figure 11.19. Polymerase IIB, lacking the CTD, could not be phosphorylated. The unphosphorylated polymerase IIA was a much better phosphorylation substrate than IIO, as expected. **(b)** Purification of the phosphorylated CTD. Reinberg and colleagues cleaved the CTD from the phosphorylated polymerase IIa subunit with the protease chymotrypsin (Chym), electrophoresed the products, and visualized them by autoradiography. Lane 1, reaction products before chymotrypsin cleavage; lanes 2 and 3, reaction products after chymotrypsin cleavage. The position of the CTD had been identified in a separate experiment. (*Source:* Lu, H., L. Zawel, L. Fisher, J.-M. Egly, and D. Reinberg, Human general transcription factor IIH phosphorylates the C-terminal domain of RNA polymerase II. *Nature* 358 (20 Aug 1992) p. 642, f. 3. Copyright © Macmillan Magazines Ltd.)

products (Figure 11.20b) shows a labeled CTD fragment, indicating that labeled phosphate has been incorporated into the CTD part of the large polymerase II subunit. The rest of the subunit was not labeled.

To prove that none of the subunits of RNA polymerase II was helping in the kinase reaction, Reinberg and coworkers cloned a chimeric gene that codes for the CTD as a fusion protein that also includes the DNA-binding domain from the transcription factor GAL4 and the enzyme glutathione-*S*-transferase. It appeared that TFIIH, all by itself, was capable of phosphorylating the CTD domain of this fusion protein. Thus, this TFIIH preparation had the appropriate kinase activity, even in the absence of other polymerase II subunits.

All of the experiments described so far were done under conditions in which the polymerase (or polymerase domain) was bound to DNA. Is this important? To find out, Reinberg's group tried the kinase assay with polymerase II in the presence of DNA that had a complete promoter, or merely the TATA box or the initiator regions of the promoter, or even no promoter at all. The result was that the TFIIH preparation performed the phosphorylation quite well in the presence of a TATA box, or an initiator, but did very poorly with a synthetic DNA (poly [dI-dC]) that contained neither.

Thus, TFIIH appears to phosphorylate polymerase II only when it is bound to DNA. We now know that the kinase activity is provided by two subunits of TFIIH.

Ordinarily, two serines (serine 2 and serine 5) of the CTD are phosphorylated, and sometimes serine 7 is phosphorylated as well. In Chapter 15, we will see evidence that transcription complexes near the promoter have CTDs in which serine 5 is phosphorylated, but that this phosphorylation shifts to serine 2 as transcription progresses. That is, serine 5 loses phosphates as serine 2 gains them during transcription. It is important to note that the protein kinase of TFIIH phosphorylates only serine 5 of the CTD. Another kinase, called **CTDK-1** in yeast and **CDK9 kinase** in metazoans, phosphorylates serine 2.

Sometimes, phosphorylation on serine 2 of the CTD is also lost during elongation, and that can cause pausing of the polymerase. In order for elongation to begin again, re-phosphorylation of serine 2 of the CTD must occur.

SUMMARY The preinitiation complex forms with the hypophosphorylated form of RNA polymerase II (IIA). Then, TFIIH phosphorylates serine 5 in the heptad repeat in the carboxyl-terminal domain (CTD) of the largest RNA polymerase II subunit, creating the phosphorylated form of the enzyme (IIO). TFIIE greatly stimulates this process in vitro. This phosphorylation is essential for initiation of transcription. During the shift from initiation to elongation, phosphorylation shifts from serine 5 to serine 2. If phosphorylation of serine 2 is also lost, the polymerase pauses until re-phosphorylation by a non-TFIIH kinase occurs.

Creation of the Transcription Bubble TFIIH is a complex protein, both structurally and functionally. It contains nine subunits and can be separated into two complexes: a protein kinase complex composed of four subunits, and a five-subunit core TFIIH complex with two separate DNA helicase/ATPase activities. One of these, contained in the largest subunit of TFIIH, is essential for viability: When its gene in yeast *(RAD25)* is mutated, the organism cannot survive. Satya Prakash and colleagues demonstrated that this helicase is essential for transcription. First they overproduced the **RAD25** protein in yeast cells, purified it almost to homogeneity, and showed that this product had helicase activity. For a helicase substrate, they used a partial duplex DNA composed of a ^{32}P-labeled synthetic 41-base DNA hybridized to single-stranded M13 DNA (Figure 11.21a). They mixed RAD25 with this substrate in the presence and absence of ATP and electrophoresed the products. Helicase activity released the short, labeled DNA from its much longer partner, so it had a much higher electrophoretic mobility and was found at the bottom of the gel. As Figure 11.21b demonstrates, RAD25 has an ATP-dependent helicase activity.

Next, Prakash and colleagues showed that transcription was temperature-sensitive in cells bearing a temperature-sensitive *RAD25* gene (*rad25*-ts$_{24}$). Figure 11.22 shows the results of an in vitro transcription assay using a G-less cassette (Chapter 5) as template. This template had a yeast TATA box upstream of a 400-bp region with no G's in the nontemplate strand. Transcription in the presence of ATP, CTP, and UTP (but no GTP) apparently initiated (or terminated) at two sites within this G-less region and gave rise to two transcripts, 375 and 350 nt in length, respectively. Transcription must terminate at the end of the G-less cassette because G's are required at that point to extend the RNA chain, and they are not available.

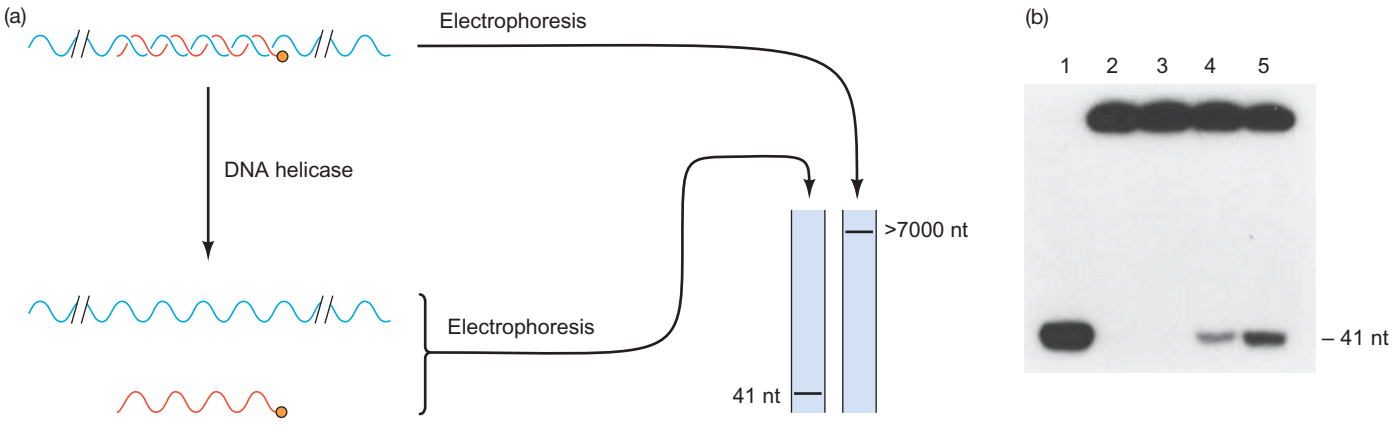


Figure 11.21 Helicase activity of TFIIH. (a) The helicase assay. The substrate consisted of a labeled 41-nt piece of DNA (red) hybridized to its complementary region in a much larger, unlabeled, single-stranded M13 phage DNA (blue). DNA helicase unwinds this short helix and releases the labeled 41-nt DNA from its larger partner. The short DNA is easily distinguished from the hybrid by electrophoresis. **(b)** Results of the helicase assay. Lane 1, heat-denatured substrate; lane 2, no protein; lane 3, 20 ng of RAD25 with no ATP; lane 4, 10 ng of RAD25 plus ATP; lane 5, 20 ng of RAD25 plus ATP. (*Source:* (*b*) Gudzer, S.N., P. Sung, V. Bailly, L. Prakash, and S. Prakash, RAD25 is a DNA helicase required for DNA repair and RNA polymerase II transcription. *Nature* 369 (16 June 1994) p. 579, f. 2c. Copyright © Macmillan Magazines Ltd.)

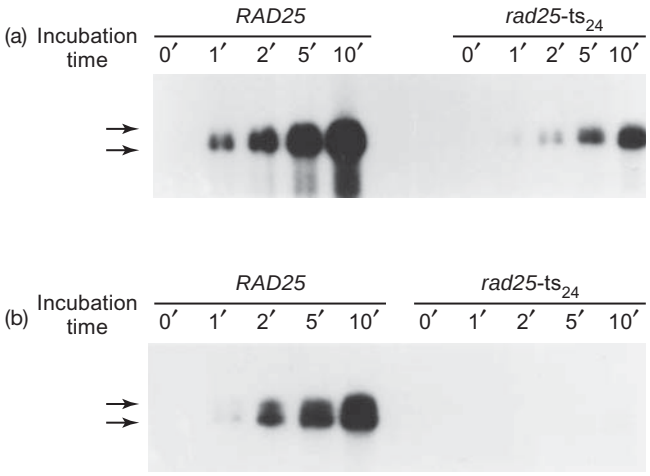


Figure 11.22 The TFIIH DNA helicase gene product (RAD25) is required for transcription in yeast: Prakash and colleagues tested extracts from wild-type (*RAD25*) and temperature-sensitive mutant (*rad25*-ts_{24}) cells for transcription of a G-less cassette template at the **(a)** permissive and **(b)** nonpermissive temperatures. After allowing transcription for 0–10 min in the presence of ATP, CTP, and UTP (but no GTP), with one ^{32}P-labeled nucleotide, they electrophoresed the labeled products and detected the bands by autoradiography. The origin of the extract (*RAD25* or *rad25*-ts_{24} cells), as well as the time of incubation in minutes, is given at top. Arrows at left denote the positions of the two G-less transcripts. We can see that transcription is temperature-sensitive when the TFIIH DNA helicase (RAD25) is temperature-sensitive. (*Source:* Gudzer, S.N., P. Sung, V. Bailly, L. Prakash, and S. Prakash, RAD25 is a DNA helicase required for DNA repair and RNA polymerase II transcription. *Nature* 369 (16 June 1994) p. 580, f. 3 b–c. Copyright © Macmillan Magazines Ltd.)

(The shorter transcript may have come from premature termination within the G-less cassette, rather than from a different initiation site.) Panel (a) shows the results of transcription for 0–10 min at the permissive temperature (22°C). It is clear that the *rad25*-ts_{24} mutant extract gave weaker transcription than the wild-type *(RAD25)* extract even at low temperature. Panel (b) shows the results of transcription at the nonpermissive temperature (37°C). The elevated temperature completely inactivated transcription in the *rad25*-ts_{24} mutant extract. Thus, the *RAD25* product (the TFIIH DNA helicase) is required for transcription.

What step in transcription requires DNA helicase activity? The chain of evidence leading to the answer begins with the following consideration: Transcription of class II genes, unlike transcription of class I and III genes, requires ATP (or dATP) hydrolysis. Of course, the α-β-bonds of all four nucleotides, including ATP, are hydrolyzed during all transcription, but class II transcription requires hydrolysis of the β-γ-bond of ATP. The question arises: What step requires ATP hydrolysis? We would naturally be tempted to look at TFIIH for the answer to this question because it has two activities (CTD kinase and DNA helicase) that involve hydrolysis of ATP. The answer appears to be that the helicase activity of TFIIH is the ATP-requiring step. The main evidence in favor of this hypothesis is that GTP can substitute for ATP in CTD phosphorylation, but GTP cannot satisfy the ATP hydrolysis requirement for transcription. Thus, transcription requires ATP hydrolysis for some process besides CTD phosphorylation, and the best remaining candidate is DNA helicase.

Now let us return to the main question: What transcription step requires DNA helicase activity? The most likely answer is promoter clearance. In Chapter 6 we defined transcription initiation to include promoter clearance, but promoter clearance can also be considered a separate event that serves as the boundary between initiation and elongation. James Goodrich and Tjian asked this question: Are TFIIE and TFIIH required for initiation or for promoter clearance? To find the answer, they devised an assay that measures the production of abortive transcripts (trinucleotides). The appearance of abortive transcripts indicates that a productive transcription initiation complex has formed, including local DNA melting and synthesis of the first phosphodiester bond. Goodrich and Tjian found that TFIIE and TFIIH were not required for production of abortive transcripts, but TBP, TFIIB, TFIIF, and RNA polymerase II were required. Thus, TFIIE and TFIIH are not required for transcription initiation, at least up to the promoter clearance step. However, TFIIH is required for full DNA melting at promoters. If the largest subunit of human TFIIH is mutated, the DNA helicase of that subunit is defective, and the DNA at the promoter does not open completely. This could block promoter clearance, as explained later in this section.

These findings left open the possiblitity that TFIIE and TFIIH are required for either promoter clearance or RNA elongation, or both. To distinguish among these possibilities, Goodrich and Tjian assayed for elongation and measured the effect of TFIIE and TFIIH on that process. By leaving out the nucleotide required in the 17th position, but not before, they allowed transcription to initiate (without TFIIE and TFIIH) on a supercoiled template and proceed to the 16-nt stage. (They used a supercoiled template because transcription on such templates in vitro does not require TFIIE and TFIIH, nor does it require ATP.) Then they linearized the template by cutting it with a restriction enzyme and added ATP to allow transcription to continue in the presence or absence of TFIIE and TFIIH. They found that TFIIE and TFIIH made no difference in this elongation reaction. Thus, because TFIIE and TFIIH appear to have no effect on initiation or elongation, Goodrich and Tjian concluded that TFIIE and TFIIH are required in the promoter clearance step. Figure 11.23 summarizes these findings and more recent data discussed in the next paragraphs.

Tjian and others assumed that the DNA helicase activity of TFIIH acted directly on the DNA at the initiator to melt it. But cross-linking studies performed in 2000 by Tae-Kyung Kim, Richard Ebright, and Danny Reinberg showed that TFIIH (in particular, the subunit bearing the promoter-melting DNA helicase) forms cross-links with DNA between positions +3 and +25, and perhaps farther downstream. This site of interaction for TFIIH is downstream of

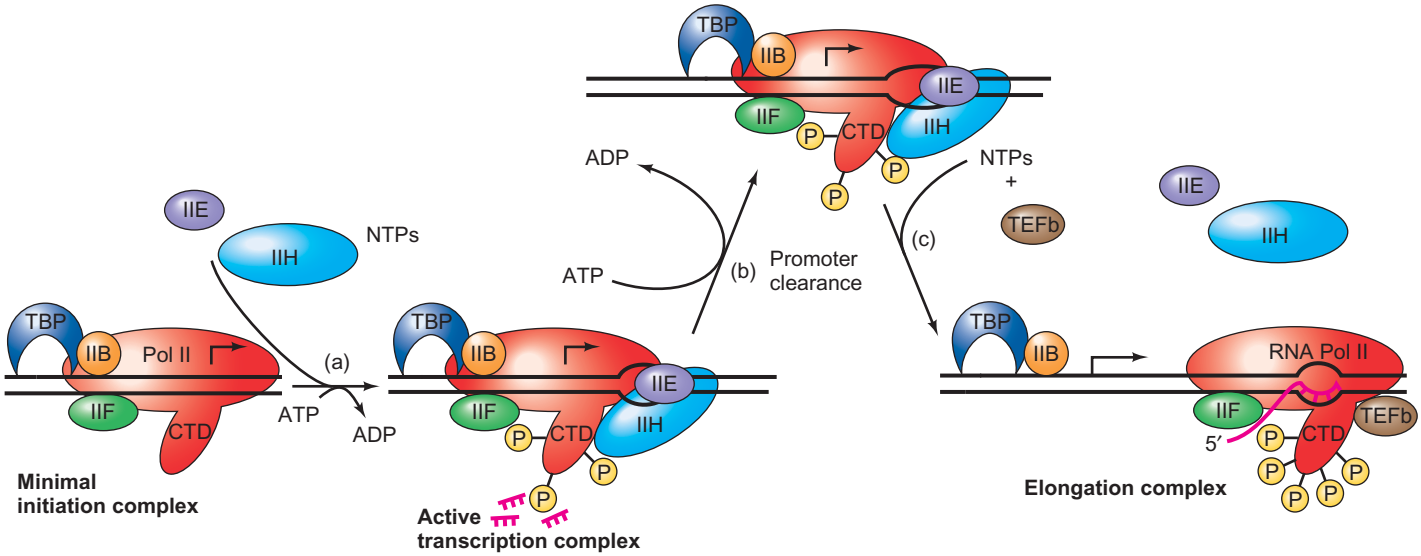


Figure 11.23 A model for the participation of general transcription factors in initiation, promoter clearance, and elongation. (a) TBP (or TFIID), along with TFIIB, TFIIF, and RNA polymerase II form a minimal initiation complex at the initiator. Addition of TFIIH, TFIIE, and ATP allows DNA melting at the initiator region and partial phosphorylation of the CTD of the largest subunit of RNA polymerase. These events allow production of abortive transcripts (magenta), but the polymerase stalls at position +10 to +12. **(b)** With energy provided by ATP, the DNA helicase of TFIIH causes further unwinding of the DNA, expanding the transcription bubble. This expansion releases the stalled polymerase and allows it to clear the promoter. **(c)** With further phosphorylation of the polymerase CTD by TEFb and with continuous addition of NTPs, the elongation complex continues elongating the RNA. TBP and TFIIB remain at the promoter. TFIIE and TFIIH are not needed for elongation and dissociate from the elongation complex. (*Source:* Adapted from Goodrich, J.A. and T. Tjian. 1994. Transcription factors IIE and IIH and ATP hydrolysis direct promoter clearance by RNA polymerase II. *Cell* 77:145–56.)

the site of the first transcription bubble (position −9 to +2). On the other hand, TFIIE cross-links to the transcription bubble region; TFIIB, TFIID, and TFIIF cross-link to the region upstream of the bubble; and RNA polymerase cross-links to the entire region encompassing all the other factors. These findings imply that the DNA helicase of TFIIH is not in contact with the first transcription bubble, and therefore cannot create the bubble by directly unwinding DNA there. Addition of ATP has no effect on the interactions upstream of the transcription bubble, but it does perturb the interactions within and downstream of the bubble.

We know from previous work that the helicase of TFIIH is responsible for creating the transcription bubble, but the cross-linking work described here indicates that it cannot directly unwind the DNA at the transcription bubble. So how does it create the bubble? Kim and associates suggested that it acts like a molecular "wrench" by untwisting the downstream DNA. Because TFIID and TFIIB (and perhaps other proteins) hold the DNA upstream of the bubble tightly, and this binding persists after addition of ATP, untwisting the downstream DNA would create strain in between and open up the DNA at the transcription bubble. This would allow the polymerase to initiate transcription and move 10–12 bp downstream. But previous work has shown that the polymerase stalls at that point unless it gets further help from TFIIH, which apparently twists the downstream DNA further to lengthen the transcription bubble, releasing the stalled polymerase to clear the promoter.

Figure 11.23 is drawn schematically so the effects of TFIIH on CTD phosphorylation and DNA unwinding are easy to see. But the real structure of the preinitiation complex is more complicated. Kornberg and colleagues modeled the positions of all the general transcription factors (except TFIIA) in the preinitiation complex, based on previous structural studies of TFIIE-polymerase II, TFIIF-polymerase II, and TFIIE-TFIIH complexes (Figure 11.24). The second-largest subunit of TFIIF (Tfg2) is homologous to the bacterial σ-factor, and lies at approximately the same position relative to the promoter as σ. In fact, two domains of Tfg2 that are homologous to domains 2 and 3 of *E. coli* σ-factor are labeled "2" and "3" in the figure. TFIIE lies about 25 bp downstream of the polymerase active center, in position to fulfill its role in recruiting TFIIH. And TFIIH is in position for its DNA helicase activity to act as a molecular wrench to open the promoter DNA, either directly, or indirectly by inducing negative supercoiling.

SUMMARY TFIIE and TFIIH are not essential for formation of an open promoter complex, or for elongation, but they are required for promoter clearance. TFIIH has a DNA helicase activity that is essential for transcription, presumably because it causes full melting of the DNA at the promoter and thereby facilitates promoter clearance.

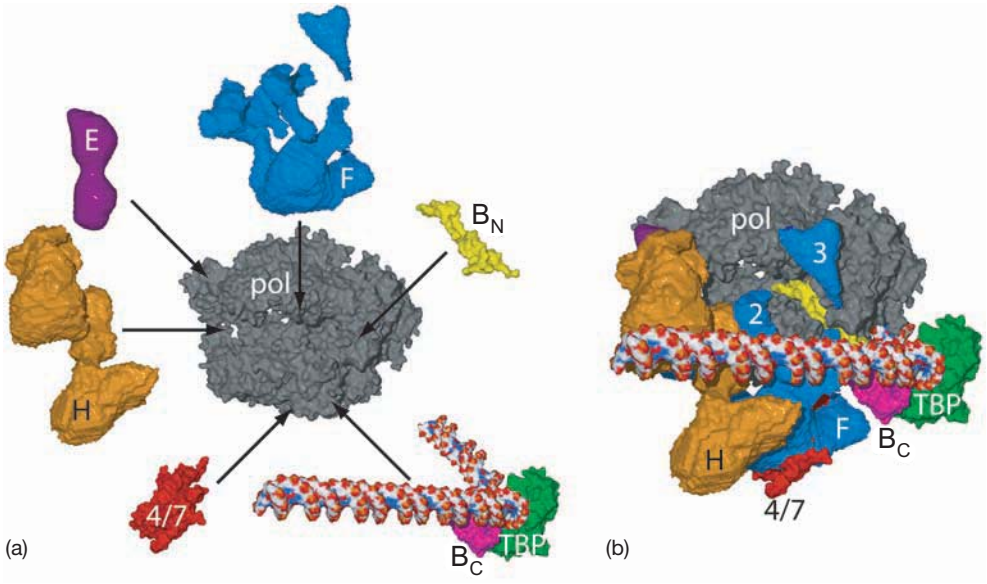


Figure 11.24 A model for the class II preinitiation complex. Kornberg and colleagues added previous structural information about the positions of promoter DNA, TFIIF, TFIIE, and TFIIH to their crystal structure of the TFIIB-RNA polymerase II complex to generate this composite model. **(a)** A blow-up to show the identities of all the components of the complex. The red component (4/7) represents Rpb4 and Rpb7, and pol (gray) denotes the rest of RNA polymerase II. B_N and B_C denote the N-terminal and C-terminal domains of TFIIB, respectively. The promoter DNA is represented by a red, white, and blue model, with a pronounced bend caused by binding of TBP. **(b)** Intact structure. Note that the transcription bubble has not yet formed. The direction of transcription is right to left. (*Source:* (*a–b*) Reprinted with permission from *Science,* Vol. 303, David A. Bushnell, Kenneth D. Westover, Ralph E. Davis, Roger D. Kornberg, "Structural Basis of Transcription: An RNA Polymerase II-TFIIB Cocrystal at 4.5 Angstroms" Fig. 6, p. 986. Copyright 2004, AAAS.)

The Mediator Complex and the RNA Polymerase II Holoenzyme

Another collection of proteins, known as **Mediator,** can also be considered a general transcription factor because it is part of most, if not all, class II preinitiation complexes. Unlike the other general transcription factors, Mediator is not required for initiation per se. But it is required for activated transcription, as we will see in Chapter 12. Mediator was first discovered in yeast, and found to contain about 20 polypeptides. A human Mediator was discovered later, and it is also a very large complex of over 20 polypeptides, only a minority of which have clear homology to those of yeast Mediator.

Our discussion so far has assumed that a preinitiation complex assembles at a class II promoter one protein at a time. This may indeed occur, but some evidence suggests that class II preinitiation complexes can assemble by binding a preformed **RNA polymerase II holoenzyme** to the promoter. The holoenzyme contains RNA polymerase, a subset of general transcription factors, and the Mediator complex.

Evidence for the holoenzyme concept came in 1994 with work from the laboratories of Roger Kornberg and Richard Young. Both groups isolated a complex protein from yeast cells, which contained RNA polymerase II and many other proteins. Kornberg and colleagues used immunoprecipitation with an antibody directed against one component of the holoenzyme to precipitate the whole complex. They recovered the subunits of RNA polymerase II, the subunits of TFIIF, and 17 other polypeptides. They could restore accurate transcription activity to this holoenzyme by adding TBP, TFIIB, E, and H. TFIIF was not required because it was already part of the holoenzyme.

Anthony Koleske and Young used a series of purification steps to isolate a holoenzyme from yeast that contained RNA polymerase II, TFIIB, TFIIF, and TFIIH. All this holenzyme needed for accurate transcription in vitro was TFIIE and TBP, so it contained more of the general transcription factors than the holoenzyme isolated by Kornberg and associates. Koleske and Young also identified some of the Mediator polypeptides in their holoenzyme and named them *SRB proteins* (*SRB2, SRB4, SRB5,* and *SRB6*).

The SRB proteins were discovered by Young and colleagues in a genetic screen whose logic went like this: Deletion of part of the CTD of the largest polymerase II subunit led to ineffective stimulation of transcription by the GAL4 protein, a transcription activator we will study in greater detail in Chapter 12. Young and coworkers then screened for mutants that could suppress this weak stimulation by GAL4. They identified several suppressor mutations in genes they named *SRB*s, for "suppressor of RNA polymerase B." We will discuss the probable basis for this suppression in Chapter 12. For now, it is enough to stress that these SRB proteins are required, at least in yeast, for

optimal activation of transcription in vivo, and that they are part of the Mediator complex of the yeast polymerase II holoenzyme. Mammalian, including human, holoenzymes have also been isolated.

SUMMARY Yeast and mammalian cells have an RNA polymerase II holoenzyme that contains many polypeptides in addition to the subunits of the polymerase. The extra polypeptides include a subset of general transcription factors (not including TBP) and Mediator.

Elongation Factors

Eukaryotes control transcription primarily at the initiation step, but they also exert some control during elongation, at least in class II genes. This can involve overcoming transcription pausing or transcription arrest. A common characteristic of RNA polymerases is that they do not transcribe at a steady rate. Instead, they pause, sometimes for a long time, before resuming transcription. These pauses tend to occur at certain defined **pause sites,** because the DNA sequences at these sites destabilize the RNA–DNA hybrid and cause the polymerase to backtrack, probably extruding the free 3′-end of the nascent RNA into a pore in the enzyme, as we learned in Chapter 10. If the backtracking is limited to just a few nucleotides, the pause is relatively short, and the polymerase can resume transcribing on its own. On the other hand, if the backtracking goes too far, the polymerase cannot recover on its own, but needs help from an elongation factor. This more severe situation is termed a **transcription arrest** rather than a **transcription pause.**

Promoter Proximal Pausing Genome-wide analysis of the positions of RNA polymerase II on genes has shown that a sizable fraction of genes (perhaps 20–30%) contain polymerases paused at specific pause sites lying 20–50 bp downstream of the transcription start site. Some of the genes with such paused polymerases are those, such as the *Drosophila Hsp70* gene, that need to be activated quickly upon induction—in this case, by heat shock. These genes have polymerases poised to resume transcribing, as soon as they receive the signal to do so.

To understand this signal, it helps to understand how the polymerase became paused in the first place. Two protein factors are known to help stabilize RNA polymerase II in the paused state. These are **DRB sensitivity-inducing factor (DSIF)** and **negative elongation factor (NELF).** DSIF comprises two subunits, the elongation factors Spt4 and Spt5, which are found in eukaryotes from yeast to humans. NELF, on the other hand, is found in vertebrates, but not in all metazoans.

The signal to leave the paused state is delivered by **positive transcription elongation factor-b (P-TEFb).** This factor has a protein kinase that can phosphorylate polymerase II, DSIF, and NELF. Upon phosphorylation, NELF leaves the paused complex, but DSIF remains behind to stimulate, rather than inhibit, elongation.

SUMMARY RNA polymerases can be induced to pause at specific sites near promoters by proteins such as DSIF and NELF. This pausing can be reversed by P-TEFb, which phosphorylates the polymerase, as well as DSIF and NELF.

TFIIS Reverses Transcription Arrest In 1987, Reinberg and Roeder discovered a HeLa cell factor, which they named **TFIIS,** that specifically stimulates transcription elongation in vitro. This factor is homologous to IIS, which was originally found by Natori and colleagues in Ehrlich ascites tumor cells.

Reinberg and Roeder demonstrated that TFIIS affects elongation, but not initiation, by testing it on preinitiated complexes (Figure 11.25). They incubated polymerase II with a DNA template and nucleotides to allow initiation to occur, then added heparin (a polyanion that can bind to RNA polymerase as DNA would) to bind any free polymerase and block new initiation, then added either TFIIS or buffer and measured the rate of incorporation of labeled GMP into RNA. Figure 11.25 shows that TFIIS enhanced RNA synthesis considerably: the vertical dashed lines show that TFIIS

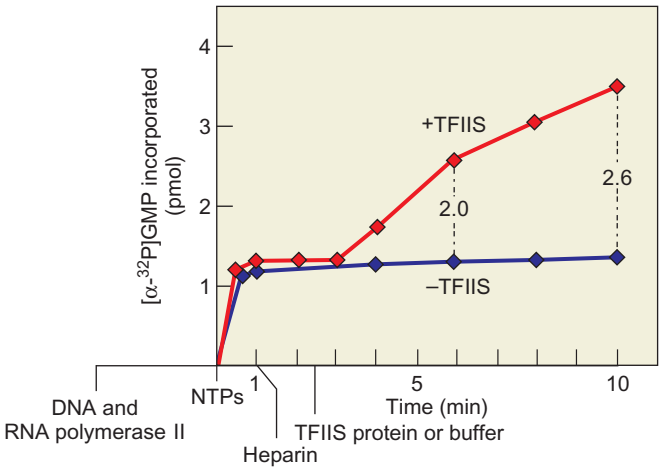


Figure 11.25 Effect of TFIIS on transcription elongation. Reinberg and Roeder formed elongation complexes as outlined in the time line at bottom. At time –3 min, they added DNA and RNA polymerase, then at time 0 they started the reaction by adding all four NTPs, one of which (GTP) was ^{32}P-labeled. At time +1 min, they added heparin to bind any free RNA polymerase, so all transcription complexes thereafter should be elongation complexes. Finally, at time +2.5 min, they added either TFIIS (red) or buffer (blue) as a negative control. They allowed labeled GMP incorporation to occur for various lengths of time, then took samples of the reaction mixture and measured the label incorporated into RNA. The dashed vertical lines indicate the fold stimulation of total RNA synthesis by TFIIS. (*Source:* Adapted from D. Reinberg and R.G. Roeder, Factors involved in specific transcription by mammalian RNA polymerase II. Transcription factor IIS stimulates elongation of RNA chains. *Journal of Biological Chemistry* 262:3333, 1987.)

stimulated GMP incorporation 2.0-fold by the 6-min mark, and 2.6-fold by the 10-min mark. Clearly, the *rate* of elongation increased even more dramatically—at least 10-fold.

It remained possible that TFIIS also stimulated transcription initiation. To investigate this possibility, Reinberg and Roeder repeated the experiment, but added TFIIS in the initial incubation, before they added heparin. If TFIIS really did stimulate initiation as well as elongation, then it should have produced a greater stimulation in this experiment than in the first. But the stimulations by TFIIS in the two experiments were almost identical. Thus, TFIIS appears to stimulate elongation only.

How does TFIIS enhance transcription elongation? Reinberg and Roeder performed an experiment that strongly suggested it does so by limiting transcription arrest.

One can detect pausing (or arresting) during in vitro transcription by electrophoresing the in vitro transcripts and finding discrete bands that are shorter than full-length transcripts. Reinberg and Roeder found that TFIIS minimized the appearance of these short transcripts, indicating that it minimized transcription arrest. Other workers have since confirmed this conclusion.

Daguang Wang and Diane Hawley demonstrated in 1993 that RNA polymerase II has an inherent, weak RNase activity that can be stimulated by TFIIS. This finding, and subsequent studies, led to a hypothesis to explain how TFIIS can restart arrested transcription (Figure 11.26). The arrested RNA polymerase has backtracked so far that the 3′-end of the nascent RNA is no longer in the enzyme's active site. Instead, it is extruded out through the pore and funnel that lead to the active site. With no 3′-terminal nucleotide to add to, the polymerase is stuck. So TFIIS activates the RNase activity in RNA polymerase II, which cleaves off the extruded part of the nascent RNA and creates a new 3′-terminus in the enzyme's active site.

How does TFIIS convert an enzyme that normally synthesizes RNA to one that breaks down RNA? Patrick Cramer and colleagues have obtained an x-ray crystal structure of an RNA polymerase II-TFIIS complex that sheds additional light on this question. Figure 11.27 shows a cutaway diagram of the complex, based on the crystal structure. TFIIS consists of three domains, including one that features a zinc ribbon. This zinc ribbon lies in the same pore and funnel of polymerase II as the extruded RNA. Just at the tip of the zinc ribbon are two acidic residues in very close proximity to metal A at the active site of the enzyme. In this position, the acidic side chains are ideally located to coordinate a second magnesium ion that would participate, along with the first, in ribonuclease activity.

Thus, TFIIS appears to change the activity of RNA polymerase, not by binding to the surface of the enzyme and effecting some conformational change within, but by getting right into the active site of the enzyme and actively participating in catalysis. This hypothesis receives strong support from the finding of a bacterial protein, called **GreB** in *E. coli,* that has the same function as TFIIS in restarting

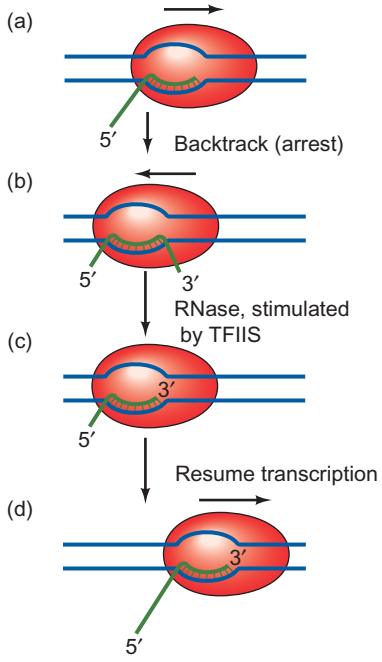


Figure 11.26 A model for reversal of transcription arrest by TFIIS. (a) RNA polymerase II, transcribing the DNA from left to right, has paused at a pause site. **(b)** The polymerase has backtracked to the left, extruding the 3′-end of the nascent RNA out of the enzyme's active site. This has caused a transcription arrest from which the polymerase cannot recover on its own. **(c)** A latent ribonuclease activity of the polymerase, stimulated by TFIIS, has cleaved off the extruded 3′-end of the nascent RNA. **(d)** With a free RNA 3′-end back in the active site, the polymerase can resume transcription.

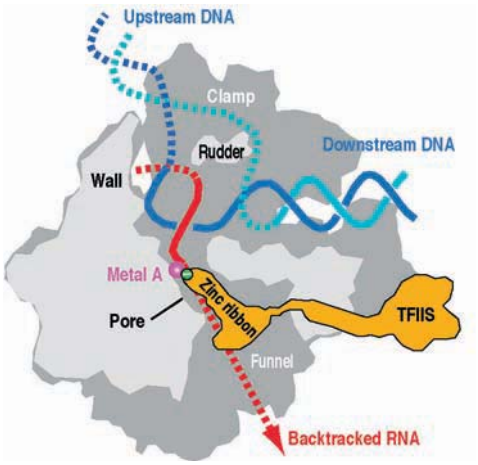


Figure 11.27 Cutaway view of the arrested yeast RNA polymerase II-TFIIS complex. The polymerase has backtracked, extruding the 3′-end of the nascent RNA (red) out of the enzyme's active site, into the pore and funnel. The zinc ribbon of TFIIS (orange) also lies in the pore and funnel, and its tip, containing two acidic residues, represented by the green circle and minus sign, approaches the metal A at the catalytic center of the polymerase, represented by the magenta circle. In this position, the two acidic residues can coordinate a second metal that collaborates with the first to constitute a ribonuclease activity that cleaves off the end of the extruded RNA. (*Source:* Reprinted from *Cell,* Vol 114, Conaway et al., "TFIIS and GreB: Two Like-Minded Transcription Elongation Factors with Sticky Fingers," fig. 1, pp. 272–274. Copyright 2003, with permission from Elsevier. Image courtesy of Joan Weliky Conaway and Patrick Cramer.)

arrested transcription. The two proteins are not **homologous**; that is, they share no sequence similarity, so they do not seem to have descended from a common evolutionary ancestor. However, GreB has a coiled-coil domain that extends into the exit channel for extruded RNA in the *E. coli* RNA polymerase in the same way the zinc ribbon in TFIIS does. Furthermore, located at the tip of the coiled-coil of GreB, adjacent to the metal ion at the polymerase active site, are two acidic residues that probably play the same role in ribonuclease catalysis as their counterparts in TFIIS appear to. This apparent convergent evolution of function argues for the validity of that proposed function.

It is interesting that an initiation factor (TFIIF) is also reported to play a role in elongation. It apparently does not limit arrests at defined DNA sites, as TFIIS does, but limits transient pausing at random DNA sites.

SUMMARY Polymerases that have backtracked and have become arrested can be rescued by TFIIS. This factor performs the rescue by inserting into the active site of RNA polymerase and stimulating an RNase that cleaves off the extruded 3′-end of the nascent RNA, which is causing transcription arrest. TFIIF also stimulates elongation, apparently by limiting transient pausing.

TFIIS Stimulates Proofreading of Transcripts Not only does TFIIS counteract pausing, it also contributes to proofreading of transcripts, presumably by a variation on the mechanism it uses to restart arrested transcription: stimulating an inherent RNase in the RNA polymerase to remove misincorporated nucleotides. Diane Hawley and her colleagues followed the procedure described in Figure 11.28a to measure the effect of TFIIS on proofreading. First, they isolated unlabeled elongation complexes that were paused at a variety of sites close to the promoter. Next, they walked the complexes to a defined position (Chapter 6) in the presence of radioactive UTP to label the RNA in the complexes. Next, they added ATP or GTP to extend the RNA by one more base, to position +43. The base that is called for at this position is A, but if G is all that is available, the polymerase will incorporate it, though at lower efficiency. Actually, Hawley and colleagues discovered that their ultrapure GTP contained a small amount of ATP, so AMP and GMP were incorporated in about equal quantities at position +43, even though ultrapure GTP was the only nucleotide they added. Next, they either cleaved the products with RNase T1, which cuts after G's, or chased with all four nucleotides to extend the labeled RNA to full length and then cut it with RNase T1. Finally, they subjected all RNase T1 products to electrophoresis and visualized the labeled products by autoradiography.

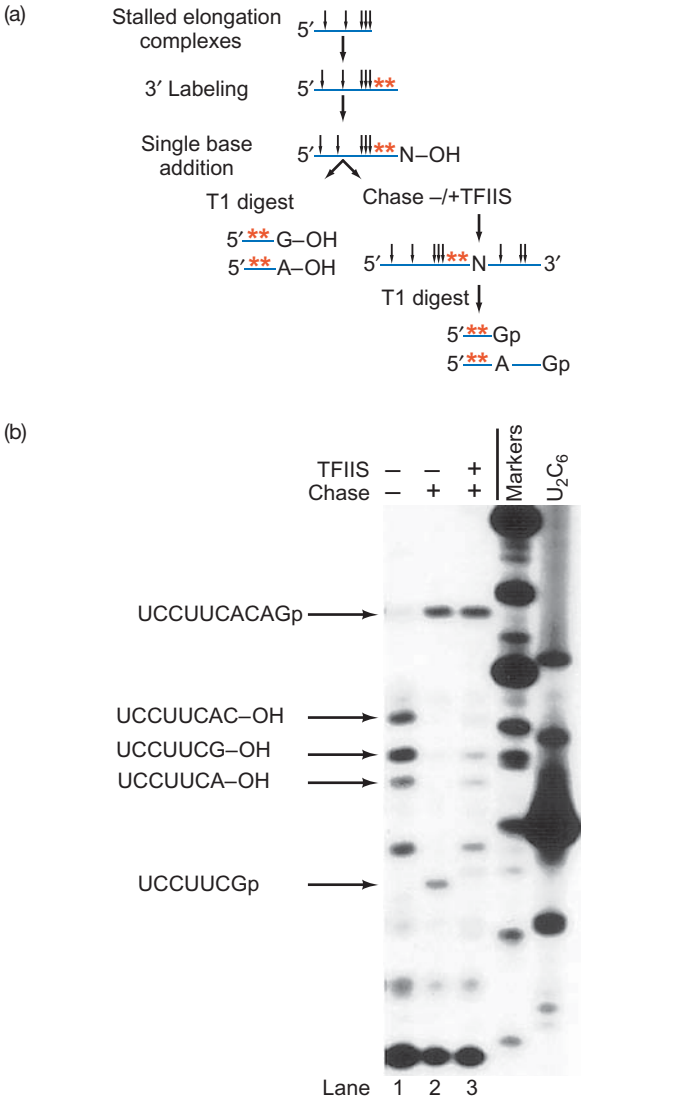


Figure 11.28 TFIIS stimulates proofreading by RNA polymerase II. **(a)** Experimental scheme. Hawley and colleagues started with short elongation complexes and 3′-end-labeled the short transcripts by walking the polymerase farther in the presence of [α–^{32}P]UTP. Then they added GTP to force misincorporation of G into position +43 where an A was called for. Then they digested the labeled transcripts with RNase T1 to measure the misincorporation of G (left), or chased the transcripts into full length with all four nucleotides, then cleaved the transcripts with RNase T1 to measure the loss of G from position +43 by proofreading. **(b)** Experimental results. Hawley and colleagues electrophoresed the RNase T1 products from part (a) and visualized them by autoradiography. Lane 1 contained unchased transcripts. The 7-mer resulting from misincorporation of G (UCCUUCG−OH), and the 7-mer (UCCUUCA) and 8-mer (UCCUUCAC) resulting from normal incorporation of A (or A and C) are indicated by arrows at left. Lanes 2 and 3 contained RNase T1 products of transcripts chased in the absence (lane 2) or presence (lane 3) of TFIIS. The 7-mer (UCCUUCGp) indicative of the misincorporated G that remained in the chased transcript is denoted by an arrow at left. The 10-mer (UCCUUCACAGp) indicative of incorporation of A in position +43, or G replaced by A at that position by proofreading, is also denoted by an arrow at left. TFIIS allowed removal of all detectable misincorporated G. (*Source:* (*b*) Thomas, M.J., A.A. Platas, and D.K. Hawley, Transcriptional fidelity and proofreading by RNA Polymerase II. *Cell* 93 (1998) f. 4, p. 631. Reprinted by permission of Elsevier Science.)

Simply cleaving the transcript with RNase T1 allowed Hawley and coworkers to measure the relative incorporations of AMP and GMP into position +43 because electrophoresis clearly separated the terminal 7-mers ending in A and G. Figure 11.28b, lane 1 shows the results of an experiment with no chasing. The 7-mer ending in G, the result of misincorporation of G, is about equally represented with the combination of a 7-mer ending in A, and an 8-mer ending in AC, which result from correct incorporation of A (or AC) from nucleotides contaminating the GTP substrate. Lanes 2 and 3 show the effects of chasing in the absence or presence, respectively, of TFIIS. The chased, full-length transcripts were cleaved with RNase T1, which yielded a 7-mer ending in Gp from a full-length transcript that still contained the misincorporated G, or a 10-mer ending in Gp from a full-length transcript in which proofreading had changed the misincorporated G to an A. When Hawley and colleagues did the chase in the absence of TFIIS, a significant amount of the misincorporated G remained in the RNA (see the band in lane 2 opposite the arrow indicating the 7-mer UCCUUCGp). However, most of the product appeared in the 10-mer (arrow labeled UCCUUCACAGp), which indicates that the polymerase was able to do some proofreading even without TFIIS. On the other hand, when they included TFIIS in the chase, Hawley and colleagues discovered that the 7-mer disappeared, and all of the labeled product was in the form of the 10-mer. Thus, TFIIS stimulates proofreading of the transcript.

The current model for proofreading (recall Figure 11.26) is that the polymerase not only pauses in response to a misincorporated nucleotide, it backtracks, extruding the 3′-end of the RNA out of the polymerase. This causes transcription to arrest. Then, TFIIS stimulates the latent RNase activity of the polymerase, which cuts off the extruded end of the RNA, including the misincorporated nucleotide, allowing the polymerase to resume transcribing.

Recall from Chapter 6 that the auxiliary factors that stimulate proofreading in bacteria are dispensable, but that the polymerase, with help from the mismatched end of a nascent RNA, can carry out proofreading in the absence of auxiliary factors. The strong conservation of the active site of RNA polymerases suggests that the same phenomenon will be observed in eukaryotic RNA polymerases, too. Indeed, this notion fits with the finding of Hawley and colleagues that polymerase II can carry out proofreading without any help from TFIIS.

SUMMARY TFIIS stimulates proofreading—the correction of misincorporated nucleotides—presumably by stimulating the RNase activity of the RNA polymerase, allowing it to cleave off a misincorporated nucleotide (with a few other nucleotides) and replace it with the correct one.

11.2 Class I Factors

The preinitiation complex that forms at rRNA promoters is much simpler than the polymerase II preinitiation complex we have just discussed. It involves polymerase I, of course, in addition to just two transcription factors. The first is a core-binding factor called **SL1** in humans, and **TIF-IB** in some other organisms; the second is a UPE-binding factor called **upstream-binding factor** (**UBF**) in mammals and **upstream activating factor** (**UAF**) in yeast. SL1 (or TIF-IB) is the core-binding factor. Along with RNA polymerase I, it is required for basal transcription activity. In fact, the core-binding factor is necessary to recruit polymerase I to the promoter. UBF (or UAF) is the factor that binds to the UPE. It is an **assembly factor** that helps the core-binding factor bind to the core promoter element. It does so by bending the DNA dramatically, so it can also be called an **architectural transcription factor** (Chapter 12). Humans and *Xenopus laevis* exhibit an almost absolute reliance on UBF for transcription of class I genes, whereas other organisms, including yeast, rats, and mice, can carry out some transcription without the help of the assembly factor. Still other organisms, such as the amoeba *Acanthamoeba castellanii,* show relatively little need for the assembly factor.

The Core-Binding Factor

Tjian and his colleagues discovered SL1 in 1985, when they separated a HeLa cell extract into two functional fractions. One fraction had RNA polymerase I activity, but no ability to initiate accurate transcription of a human rRNA gene in vitro. Another fraction had no polymerase activity of its own, but could direct the polymerase fraction to initiate accurately on a human rRNA template. Furthermore, this transcription factor, SL1, showed species specificity. That is, it could distinguish between the human and mouse rRNA promoter.

The experiments described so far used impure polymerase I and SL1. Further experiments with highly purified components revealed that human SL1 by itself cannot stimulate human polymerase I to bind to class I promoters and begin transcribing. It requires the UBF to assist its binding, as we will see in the next section.

Because human class I transcription works so poorly with the core-binding factor SL1 in the absence of UBF, the human system is not well suited to studies of the role of the core-binding factor in recruiting polymerase I to the promoter. On the other hand, *A. castellanii,* which exhibits little dependence on a UPE-binding protein, is a better choice because the effect of the core-binding factor can be studied by itself. Marvin Paule and Robert White exploited this system to show that the core-binding factor (TIF-IB) can recruit polymerase I to the promoter and stimulate initiation in the proper place. The actual DNA sequence where the polymerase binds appears to be irrelevant.

Paule and colleagues created mutant templates with various numbers of base pairs inserted or deleted between the TIF-IB-binding site and the normal transcription initiation site. This is reminiscent of the experiment performed by Benoist and Chambon with a class II promoter, reported in Chapter 10. In that experiment, deleting base pairs between the TATA box and the normal transcription initiation site did not alter the strength of transcription and did not change the transcription initiation site relative to the TATA box. In all cases, transcription began about 30 bp downstream of the TATA box.

With the class I promoter, Paule and colleagues reached a similar conclusion. They found that adding or subtracting up to 5 base pairs between the TIF-IB binding site and the normal transcription start site still allowed transcription to occur. Furthermore, the initiation site moved upstream or downstream according to the number of base pairs added or deleted (Figure 11.29). Adding or subtracting more than 5 bp blocked transcription activity (data not shown). Paule and colleagues concluded that TIF-IB contacts polymerase I and positions it for initiation a set number of base pairs downstream.

The exact base sequence contacted by the polymerase must not matter, because it is different in each mutant DNA. To confirm that the polymerase is contacting DNA in the same place relative to the TIF-IB-binding site in each mutant, Paule and colleagues performed DNase footprinting with a wild-type template and with each mutant template. The footprints were essentially indistinguishable, reinforcing the conclusion that the polymerase binds in the same spot regardless of the DNA sequence there. This is consistent with the hypothesis that TIF-IB binds to its DNA target and positions the polymerase I by direct protein–protein contact. The polymerase appears to contact the DNA because it extends the footprint caused by TIF-IB, but this contact appears to be nonspecific.

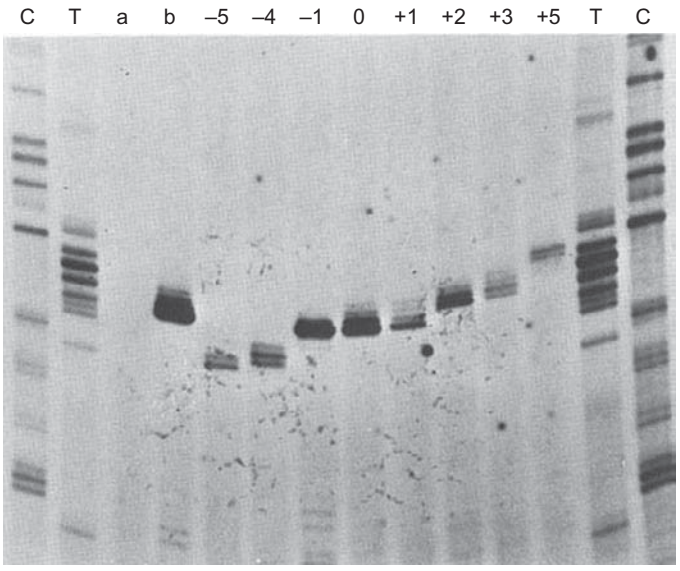


Figure 11.29 Effect of insertions and deletions on polymerase I transcription initiation site. Paule and colleagues made insertions and deletions of up to 5 bp, as indicated at top, between the TIF-IB binding site and the normal transcription start site in an *A. castellanii* rRNA promoter. Then they transcribed these templates in vitro and performed primer extension analysis (Chapter 5) with a ^{32}P-labeled 17-nt sequencing primer. They electrophoresed the labeled extended primers alongside C and T sequencing lanes using the same primer (lanes C and T). Lane a is a negative control run with vector DNA, but no rRNA promoter, lane b is a positive control containing a wild-type rRNA promoter. Lane 0 also contained the extended primer generated from the transcript of wild-type DNA with no deletion. (*Source:* Reprinted from *Cell* v. 50, Kownin et al., p. 695 © 2001, with permission from Elsevier Science.)

SUMMARY Class I promoters are recognized by two transcription factors, a core-binding factor and a UPE-binding factor. The human core-binding factor is called SL1; in some other organisms, such as *A. castellanii,* the homologous factor is known as TIF-IB. The core-binding factor is the fundamental transcription factor required to recruit RNA polymerase I. This factor also determines species specificity, at least in animals. The factor that binds the UPE is called UBF in mammals and most other organisms, but UAF in yeast. It is an assembly factor that helps the core-binding factor bind to the core promoter element. The degree of reliance on the UPE-binding factor varies considerably from one organism to another. In *A. castellanii,* TIF-IB alone suffices to recruit the RNA polymerase I and position it correctly for initiation of transcription.

The UPE-Binding Factor

Because human SL1 by itself did not appear to bind directly to the rRNA promoter, but a partially purified RNA polymerase I preparation did, Tjian and his coworkers began a search for DNA-binding proteins in the polymerase preparation. This led to the purification of human UBF in 1988. The factor as purified was composed of two polypeptides, of 97 and 94 kD. However, the 97-kD polypeptide alone is sufficient for UBF activity. When Tjian and colleagues performed footprint analysis with this highly purified UBF, they found that it had the same behavior as observed previously with partially purified polymerase I. That is, it gave the same footprint in the core element and a section of the UPE called *site A,* and SL1 intensified this footprint and extended it to a part of the UPE called *site B* (Figure 11.30). Thus, UBF, not polymerase I, was the agent that bound to the promoter in the previous experiments, and SL1 facilitates this binding. These studies did not reveal whether SL1 actually contacts the DNA in a complex with UBF, or whether it merely changes the conformation of UBF so it can contact a longer stretch of DNA that

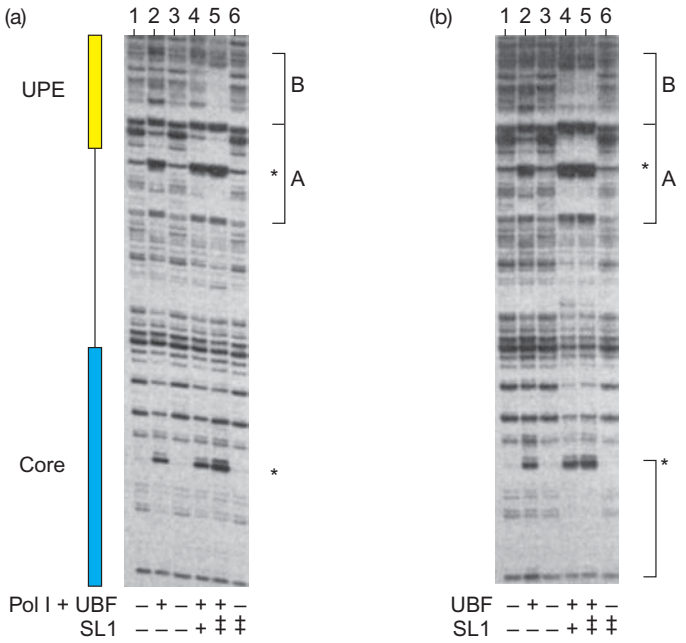


Figure 11.30 Interaction of UBF and SL1 with the rRNA promoter. Tjian and colleagues performed DNase footprinting with the human rRNA promoter and various combinations of **(a)** polymerase I + UBF and SL1 or **(b)** UBF and SL1. The proteins used in each lane are indicated at bottom. The positions of the UPE and core elements are shown at left, and the locations of the A and B sites are illustrated with brackets at right. Asterisks mark the positions of enhanced DNase sensitivity. SL1 caused no footprint on its own, but enhanced and extended the footprints of UBF in both the UPE and the core element. This enhancement is especially evident in the absence of polymerase I (panel b). (*Source:* Adapted from Bell S.P., R.M. Learned, H.-M. Jantzen, and R. Tjian, Functional cooperativity between transcription factors UBF1 and SL1 mediates human ribosomal RNA synthesis. *Science* 241 (2 Sept 1988) p. 1194, f. 3 a–b.)

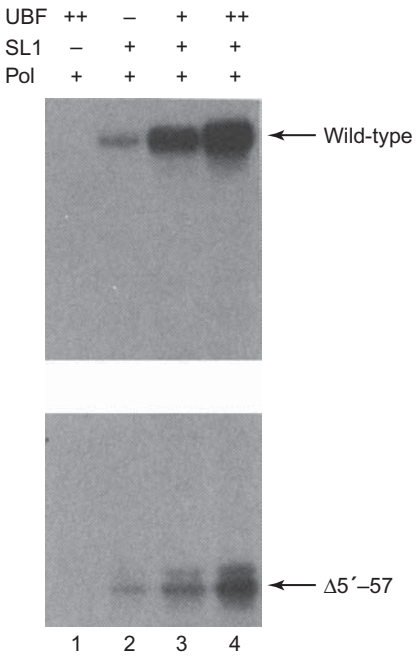


Figure 11.31 Activation of transcription from the rRNA promoter by UBF and SL1. Tjian and colleagues used an S1 assay to measure transcription from the human rRNA promoter in the presence of RNA polymerase I and various combinations of UBF and SL1, as indicated at top. The top panel shows transcription from the wild-type promoter; the bottom panel shows transcription from a mutant promoter (Δ5′–57) lacking UPE function. SL1 was required for at least basal activity, but UBF enhanced this activity on both templates. (*Source:* Bell S.P., R.M. Learned, H.-M. Jantzen, and R. Tjian, Functional cooperativity between transcription factors UBF1 and SL1 mediates human ribosomal RNA synthesis, *Science* 241 (2 Sept 1988) p. 1194, f. 4. Copyright © AAAS.)

extends into site B. Based on this and other data, we can conclude that SL1 cannot bind by itself, while UBF can. However, SL1 and UBF appear to bind cooperatively to give more extensive binding together than either could accomplish on its own.

Tjian and associates also found that UBF stimulates transcription of the rRNA gene in vitro. Figure 11.31 depicts the results of a transcription experiment using the wild-type human rRNA promoter and the mutant promoter (Δ5′–57) that lacks the UPE, and including various combinations of SL1 and UBF. Polymerase I was present in all reactions, and transcription efficiency was assayed by the S1 technique (Chapter 5). Lane 1 contained UBF, but no SL1, and showed no transcription of either template. This reaffirms that SL1 is absolutely required for transcription. Lane 2 had SL1, but no UBF, and showed a basal level of transcription. This demonstrates again that SL1 by itself is capable of stimulating basal transcription. Moreover, about as much transcription occurred on the mutant template that lacks the UPE as on the wild-type template. Thus, UBF is required for stimulation of transcription through the UPE. Lanes 3 and 4 contained both SL1 and an increasing amount of UBF. Significantly enhanced transcription occurred on both templates, but especially on the template containing the UPE. Tjian and colleagues concluded that UBF is a transcription factor that can stimulate transcription by binding to the UPE, but it can also exert an effect in the absence of the UPE, presumably by binding to the core element.

SUMMARY Human UBF is a transcription factor that stimulates transcription by polymerase I. It can activate the intact promoter, or the core element alone, and it mediates activation by the UPE. UBF and SL1 act synergistically to stimulate transcription.

Structure and Function of SL1

We have been discussing just two human factors, UBF and SL1, that are involved in transcription by polymerase I, and one of these, UBF, is probably just a single 97-kD polypeptide. But work presented earlier in this chapter showed that TATA box-binding protein (TBP) is essential

for class I transcription. Where then does TBP fit in? Tjian and coworkers demonstrated in 1992 that SL1 is composed of TBP and three TAFs. First, they purified human (HeLa cell) SL1 by several different procedures. After each step, they used an S1 assay to locate SL1 activity. Then they assayed these same fractions for TBP by Western blotting. Figure 11.32 shows the striking correspondence they found between SL1 activity and TBP content.

If SL1 really does contain TBP, then it should be possible to inhibit SL1 activity with an anti-TBP antibody. Tjian and colleagues confirmed that this worked as predicted. A nuclear extract was depleted of SL1 activity with an anti-TBP antibody. Activity could then be restored by adding back SL1, but not just by adding back TBP. Something besides TBP must have been removed.

What other factors are removed along with TBP by immunoprecipitation? To find out, Tjian and colleagues subjected the immunoprecipitate to SDS-PAGE. Figure 11.33 depicts the results. In addition to TBP and antibody (IgG), we see three polypeptides, with molecular masses of 110, 63, and 48 kD (although the 48-kD polypeptide is partially obscured by TBP). Because these were immunoprecipitated along with TBP, they must bind tightly to TBP and are therefore TBP-associated factors, or **TAF$_I$s**, by definition. Hence, Tjian called them TAF$_I$110, TAF$_I$63, and TAF$_I$48. These are completely different from the TAFs found in TFIID (compare lanes 4 and 5). The TAFs could be stripped off of the TBP and antibody in the immunoprecipitate by treating the precipitate with 1 M guanidine-HCl and reprecipitating. The antibody and TBP remained together in the

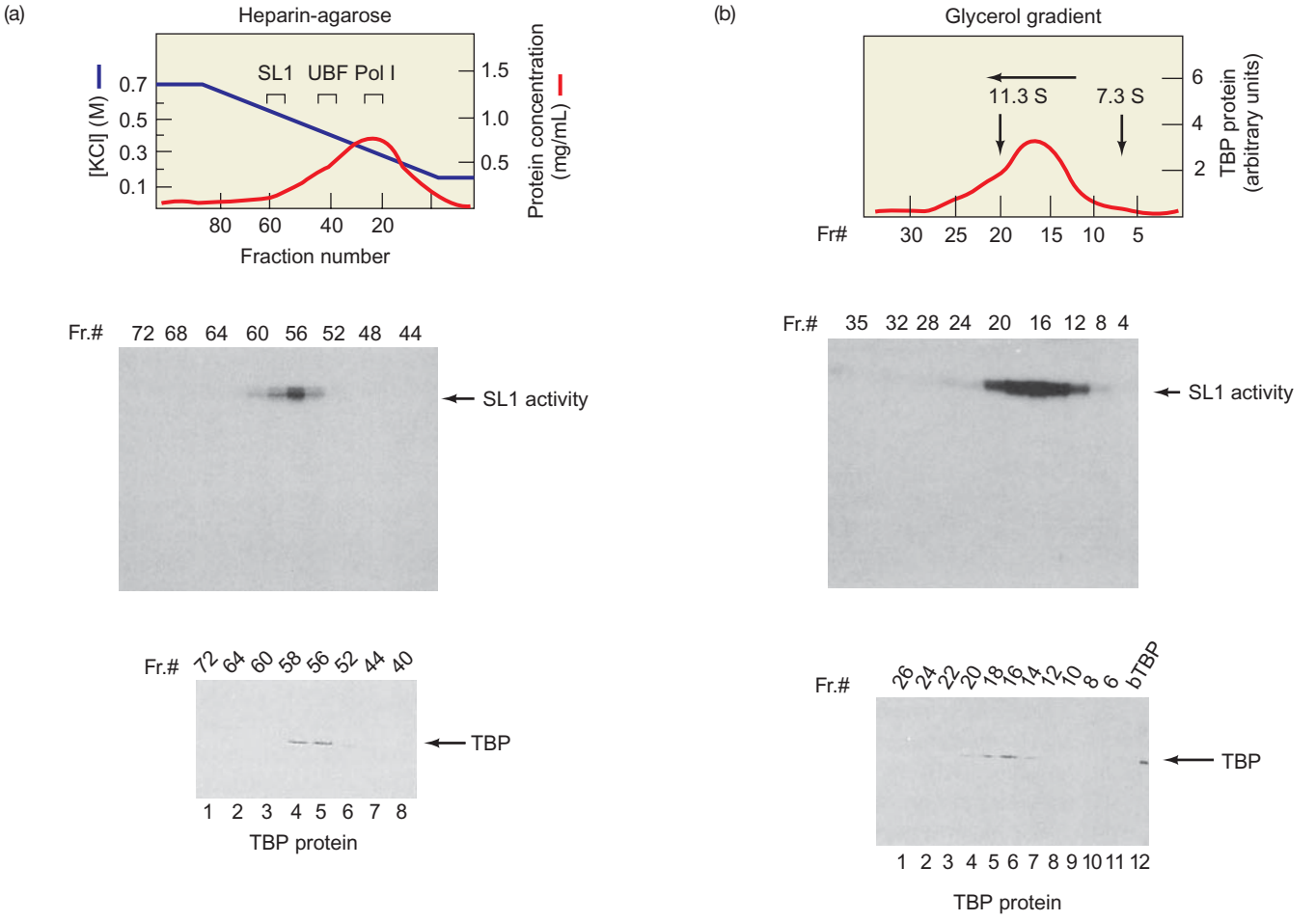


Figure 11.32 Co-purification of SL1 and TBP. (a) Heparin–agarose column chromatography (see Chapter 5 for column chromatography methods.) Top: Pattern of elution from the column of total protein (red) and salt concentration (blue), as well as three specific proteins (brackets). Middle: SL1 activity, measured by S1 protection analysis, in selected fractions. Bottom: TBP protein, detected by Western blotting, in selected fractions. Both SL1 and TBP were centered on fraction 56. **(b)** Glycerol gradient ultracentrifugation. Top: Sedimentation profile of TBP. Two other proteins, catalase and aldolase, with sedimentation coefficients of 11.3 S and 7.3 S, respectively, were run in a parallel centrifuge tube as markers. Middle and bottom panels, as in panel (a). Both SL1 and TBP sedimented to a position centered around fraction 16. (*Source:* Comai, L., N. Tanese, and R. Tjian, The TATA-binding protein and associated factors are integral components of the RNA polymerase I transcription factor, SL1. *Cell* 68 (6 Mar 1992) p. 968, f. 2a–b. Reprinted by permission of Elsevier Science.)

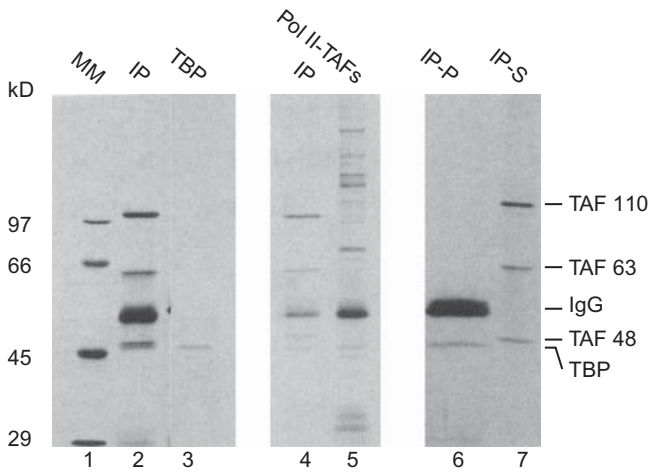


Figure 11.33 The TAFs in SL1. Tjian and colleagues immunoprecipitated SL1 with an anti-TBP antibody and subjected the polypeptides in the immunoprecipitate to SDS-PAGE. Lane 1, molecular weight markers; lane 2, immunoprecipitate (IP); lane 3, purified TBP for comparison; lane 4, another sample of immunoprecipitate; lane 5, TFIID TAFs (Pol II-TAFs) for comparison; lane 6, pellet after treating immunoprecipitate with 1 M guanidine–HCl and reprecipitating, showing TBP and antibody (IgG); lane 7, supernatant after treating immunoprecipitate with 1 M guanidine–HCl and reprecipitating, showing the three TAFs (labeled at right). (*Source:* Comai, L., N. Tanese, and R. Tjian, The TATA-binding protein and associated factors are integral components of the RNA polymerase I transcription factor, SL1. *Cell* 68 (6 Mar 1992) p. 971, f. 5. Reprinted by permission of Elsevier Science.)

precipitate (lane 6) and the TAFs stayed in the supernatant (lane 7). Tjian and colleagues could reconstitute SL1 activity by adding together purified TBP and the three TAFs, and this activity was species-specific, as one would expect. In later work, Tjian and coworkers showed that the TAF_Is and TAF_{II}s could compete with each other for binding to TBP. This finding suggested that binding of one set of TAFs to TBP is mutually exclusive of binding of the other set.

Thus, both polymerase I and polymerase II rely on transcription factors (SL1 and TFIID, respectively) composed of TBP and several TAFs. The TBP is identical in the two factors but the TAFs are completely different.

A unifying theme for all class I core-binding factors, except in yeast, is TBP. Yeast TBP binds to the core-binding factor, but not stably, the way other TBPs bind to their corresponding TAF_Is. The number and sizes of the TAF_Is we have discussed are typical of human cells. Other organisms have their own spectrum of TAF_Is.

SUMMARY Human-SL1 is composed of TBP and three TAFs: TAF_I110, TAF_I63, and TAF_I48. Fully functional and species-specific SL1 can be reconstituted from these purified components, and binding of TBP to the TAF_Is precludes binding to the TAF_{II}s. Other organisms have their own groups of TAF_Is.

11.3 Class III Factors

In 1980, Roeder and his colleagues discovered a factor that bound to the internal promoter of the 5S rRNA gene and stimulated its transcription. They named the factor **TFIIIA.** Since then, two other factors, TFIIIB and C, have been discovered. These two factors participate, not only in 5S rRNA gene transcription, but in all transcription by polymerase III.

Barry Honda and Robert Roeder demonstrated the importance of the TFIIIA factor in 5S rRNA gene transcription when they developed the first eukaryotic in vitro transcription system, from *Xenopus laevis,* and found that it could make no 5S rRNA unless they added TFIIIA. Donald Brown and colleagues went on to show that similar cell-free extracts provided with a 5S rRNA gene and a tRNA gene could make both 5S rRNA and tRNA simultaneously. Furthermore, an antibody against TFIIIA could effectively halt the production of 5S rRNA, but had no effect on tRNA synthesis (Figure 11.34). Thus, TFIIIA is required for transcription of the 5S rRNA genes, but not the tRNA genes.

If transcription of the tRNA genes does not require TFIIIA, what factors *are* involved? In 1982, Roeder and colleagues separated two new factors they called **TFIIIB** and **TFIIIC** and found that they are necessary and sufficient for transcription of the tRNA genes. We have subsequently learned that these two factors govern transcription of all classical polymerase III genes, including the 5S rRNA genes. That means that the original extracts that needed to be supplemented only with TFIIIA to make 5S rRNA must have contained TFIIIB and C.

SUMMARY Transcription of all classical class III genes requires TFIIIB and C, and transcription of the 5S rRNA genes requires these two plus TFIIIA.

TFIIIA

As the very first eukaryotic transcription factor to be discovered, TFIIIA received a considerable amount of attention. It was the first member of a large group of DNA-binding proteins that feature a so-called **zinc finger.** We will discuss the zinc finger proteins in detail in Chapter 12. Here, let us concentrate on the zinc fingers of TFIIIA. The essence of a zinc finger is a roughly finger-shaped protein domain containing four amino acids that bind a single zinc ion. In TFIIIA, and in other typical zinc finger proteins, these four amino acids are two cysteines, followed by two histidines. However, some other zinc finger-like proteins have four cysteines and no histidines. TFIIIA has nine zinc fingers in a row, and these appear to insert into the DNA major groove on either side of the internal promoter of the 5S rRNA gene. This allows specific amino acids to make contact with specific base pairs, forming a tight protein–DNA complex.

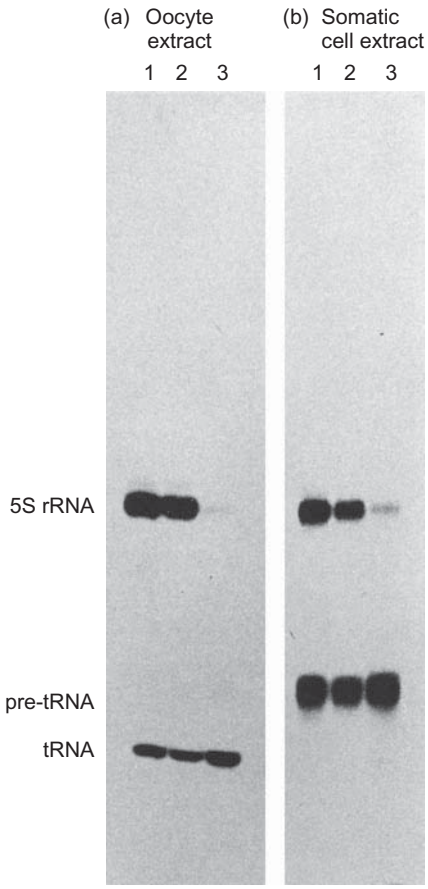


Figure 11.34 Effect of anti-TFIIIA antibody on transcription by polymerase III. Brown and colleagues added cloned 5S rRNA and tRNA genes to **(a)** an oocyte extract, or **(b)** a somatic cell extract in the presence of labeled nucleotide and: no antibody (lanes 1), an irrelevant antibody (lanes 2), or an anti-TFIIIA antibody (lanes 3). After transcription, these workers electrophoresed the labeled RNAs. The anti-TFIIIA antibody blocked 5S rRNA gene transcription in both extracts, but did not inhibit tRNA gene transcription in either extract. The oocyte extract could process the pre-tRNA product to the mature tRNA form, but the somatic cell extract could not. Nevertheless, transcription occurred in both cases. (*Source:* Pelham, H.B., W.M. Washington, and D.D. Brown, Related 5S rRNA transcription factors in *Xenopus* oocytes and somatic cells. *Proceedings of The National Academy of Sciences USA* 78 (Mar 1981) p. 1762, f. 3.)

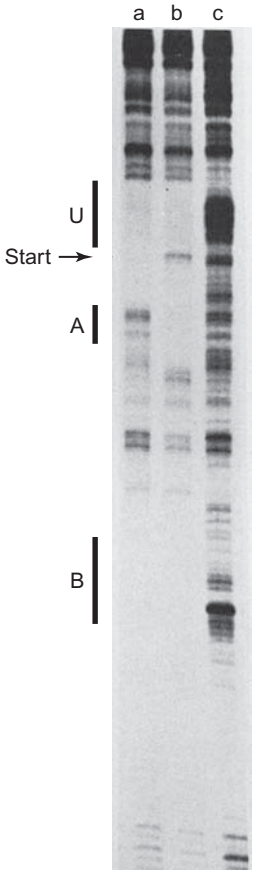


Figure 11.35 Effect of transcription on DNA binding between a tRNA gene and transcription factors. Geiduschek and colleagues performed DNase footprinting with a tRNA gene and an extract containing polymerase III, TFIIIB, and TFIIIC. Lane a contained transcription factors, but no nucleotides. Lane b had factors plus three of the four nucleotides (all but GTP), so transcription could progress for 17 nt, until GTP was needed. Lane c was a control with no added protein. The 17-bp migration of the polymerase in lane b relative to lane a caused a corresponding downstream shift in the footprint around the transcription start site, to a position extending upstream and downstream of the A box. On the other hand, the footprint in the region just upstream of the start of transcription remained unchanged. (*Source:* Kassavetis, G.A., D.L. Riggs, R. Negri, L.H. Nguyen, and E.P. Geiduschek, Transcription factor III B generates extended DNA interactions in RNA polymerase III transcription complexes on tRNA genes. *Molecular and Cellular Biology.* 9, no.171 (June 1989) p. 2555, f. 3. Copyright © 1989 American Society for Microbiology, Washington, DC. Reprinted with permission.)

TFIIIB and C

TFIIIB and C are both required for transcription of the classical polymerase III genes, and it is difficult to separate the discussion of these two factors because they depend on each other for their activities. Peter Geiduschek and coworkers established in 1989 that a crude transcription factor preparation bound both the internal promoter and an upstream region in a tRNA gene. Figure 11.35 contains DNase footprinting data that led to this conclusion. Lane c is the digestion pattern with no added protein, lane a is the result with factors and polymerase III, and lane b has all this plus three nucleoside triphosphates (ATP, CTP, and UTP), which allowed transcription for just 17 nt, until the first GTP was needed. Notice in lane a that the factors and polymerase strongly protected box B of the internal promoter and the upstream region (U) and weakly protected box A of the internal promoter. Lane b shows that the polymerase shifted downstream and a new region overlapping box A was protected. However, the protection of the upstream region persisted even after the polymerase moved away.

What accounts for the persistent binding to the upstream region? To find out, Geiduschek and colleagues partially purified TFIIIB and C and performed footprinting studies with these separated factors. Figure 11.36

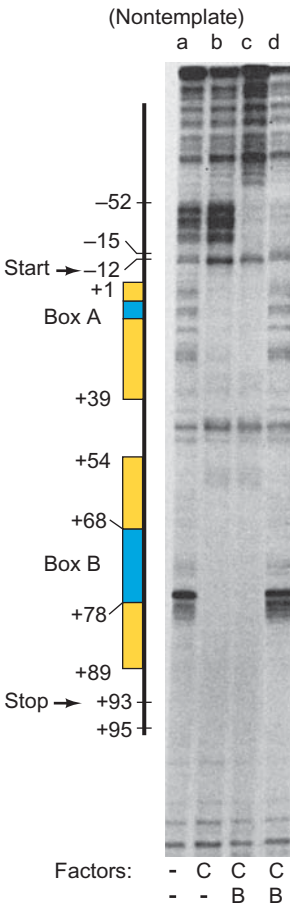


Figure 11.36 Binding of TFIIIB and C to a tRNA gene. Geiduschek and coworkers performed DNase footprinting with a labeled tRNA gene (all lanes), and combinations of purified TFIIIB and C. Lane a, negative control with no factors; lane b, TFIIIC only; lane c, TFIIIB plus TFIIIC; lane d, TFIIIB plus TFIIIC added, then heparin added to strip off any loosely bound protein. Note the added protection in the upstream region afforded by TFIIIB in addition to TFIIIC (lane c). Note also that this upstream protection provided by TFIIIB survives heparin treatment, but the protection of boxes A and B does not. Yellow boxes represent coding regions for mature tRNA. Boxes A and B within these regions are indicated in blue. (*Source:* From Kassavetis, G.A., D.L. Riggs, R. Negri, L.H. Nguyen, and E.P. Geiduschek, Transcription factor III B generates extended DNA interactions in RNA polymerase III transcription complexes on tRNA genes. *Molecular and Cellular Biology* 9:2558, 1989. Copyright © 1989 American Society for Microbiology, Washington, DC. Reprinted by permission.)

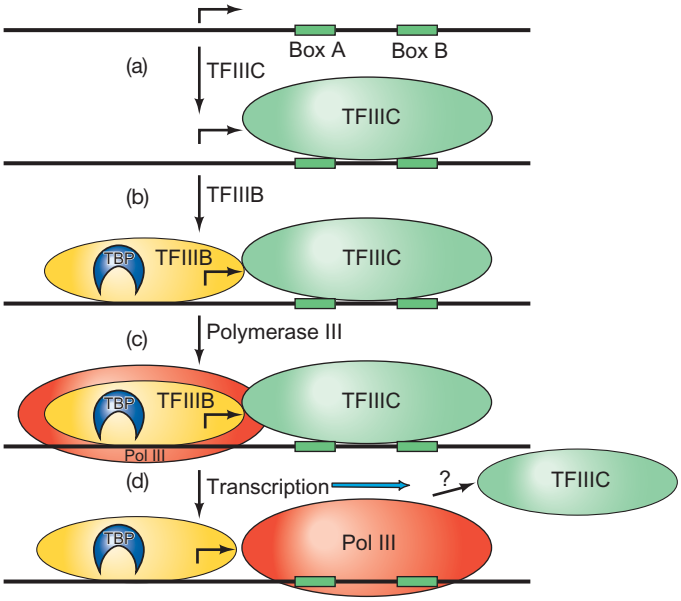


Figure 11.37 Hypothetical scheme for assembly of the preinitiation complex on a classical polymerase III promoter (tRNA), and start of transcription. (a) TFIIIC (light green) binds to the internal promoter's A and B blocks (green). **(b)** TFIIIC promotes binding of TFIIIB (yellow), with its TBP (blue), to the region upstream of the transcription start site. **(c)** TFIIIB promotes polymerase III (red) binding at the start site, ready to begin transcribing. **(d)** Transcription begins. As the polymerase moves to the right, making RNA (not shown), it may or may not remove TFIIIC from the internal promoter. But TFIIIB remains in place, ready to sponsor a new round of polymerase binding and transcription.

shows the results of one such experiment. Lane b, with TFIIIC alone, reveals that this factor protects the internal promoter, especially box B, but does not bind to the upstream region. When both factors are present, the upstream region is also protected (lane c). Similar DNase footprinting experiments made it clear that TFIIIB by itself does not bind to any of these regions. Its binding is totally dependent on TFIIIC. However, once TFIIIC has sponsored the binding of TFIIIB to the upstream region, TFIIIB appears to remain there, even after polymerase has moved on (recall Figure 11.35). Moreover, Figure 11.36, lane d, shows that TFIIIB binding persists even after heparin has stripped TFIIIC away from the internal promoter, as the upstream region is still protected from DNase, even though boxes A and B are not.

The evidence we have seen so far suggests the following model for involvement of transcription factors in polymerase III transcription (Figure 11.37): First, TFIIIC (or TFIIIA and C, in the case of the 5S rRNA genes) binds to the internal promoter; then these **assembly factors** allow TFIIIB to bind to the upstream region; then TFIIIB helps polymerase III bind at the transcription start site; finally, the polymerase transcribes the gene, perhaps removing TFIIIC (or A and C) in the process, but TFIIIB remains bound, so it can continue to promote further rounds of transcription.

Geiduschek and colleagues have provided further evidence to bolster this hypothesis. They bound TFIIIC and B to a tRNA gene (or TFIIIA, C, and B to a 5S rRNA gene), then removed (stripped) the assembly factors, TFIIIC (or A and C) with either heparin or high salt, then separated the remaining TFIIIB–DNA complex from the other factors. Finally, they demonstrated that this TFIIIB–DNA complex was still capable of supporting one round, or even multiple rounds, of transcription by polymerase III (Figure 11.38). How does TFIIIB remain so tightly bound to its DNA

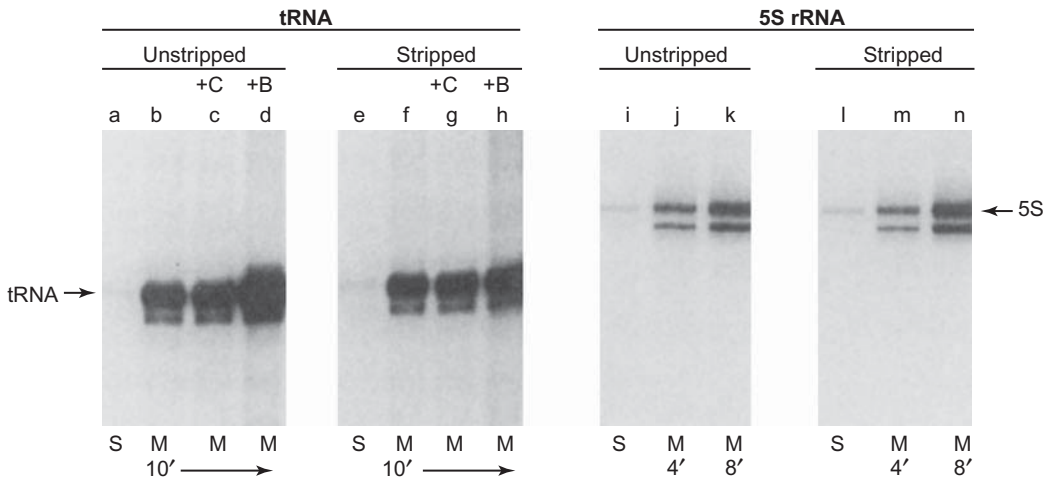


Figure 11.38 Transcription of polymerase III genes complexed only with TFIIIB. Geiduschek and coworkers made complexes containing a tRNA gene and TFIIIB and C (two panels at left), or a 5S rRNA gene and TFIIIA, B, and C (two panels at right), then stripped off TFIIIC with heparin (lanes e–h), or TFIIIA and C with a high ionic strength buffer (lanes l–n). They passed the stripped templates through gel filtration columns to remove any unbound factors, and demonstrated by gel mobility shift and DNase footprinting (not shown) that the purified complexes contained only TFIIIB bound to the upstream regions of the respective genes. Next, they tested these stripped complexes alongside unstripped complexes for ability to support single-round transcription (S; lanes a, e, i, and l), or multiple-round transcription (M; all other lanes) for the times indicated at bottom. (The single-round signals are faint, but visible.) They added extra TFIIIC in lanes c and g, and extra TFIIIB in lanes d and h as indicated at top. They confined transcription to a single round in lanes a, e, i, and l by including a relatively low concentration of heparin, which allowed elongation of RNA to be completed, but then bound up the released polymerase so it could not reinitiate. Notice that the stripped template, containing only TFIIIB, supported just as much transcription as the unstripped template in both single-round and multiple-round experiments, even when the experimenters added extra TFIIIC (compare lanes c and g, and lanes k and n). The only case in which the unstripped template performed better was in lane d, which was the result of adding extra TFIIIB. This presumably resulted from some remaining free TFIIIC that helped the extra TFIIIB bind, thus allowing more preinitiation complexes to form. (*Source:* Kassavetis, G.A., B.R. Brawn, L.H. Nguyen, and E.P. Geiduschek, *S. cerevisiae* TFIIIB is the transcription initiation factor proper of RNA polymerase III, while TFIIIA and TFIIIC are assembly factors. *Cell* 60 (26 Jan 1990) p. 237, f. 3. Reprinted by permission of Elsevier Science.)

target when it has no affinity for this DNA on its own? The answer may be that TFIIIC (or TFIIIA and TFIIIC) can cause a conformational shift in TFIIIB, revealing a site that can bind tenaciously to DNA.

TFIIIC is a remarkable protein. It can bind to both box A and box B of tRNA genes, as demonstrated by DNase footprinting and protein–DNA cross-linking studies. In some tRNA genes there is an intron between boxes A and B, and TFIIIC still manages to contact both promoter elements. How can it do that? It helps that TFIIIC is one of the largest and most complex of all the known transcription factors. The yeast TFIIIC contains six subunits with a combined molecular mass of about 600 kD. Furthermore, electron microscopic studies have shown that TFIIIC has a dumbbell shape with two globular regions separated by a stretchable linker region that allows the protein to span a surprisingly long distance.

In these studies, André Sentenac and colleagues bound yeast TFIIIC (which they called τ factor) to cloned tRNA genes having variable distances between their boxes A and B. Then they visualized the complexes by scanning transmission electron microscopy. Figure 11.39 shows the results: When the distance between boxes A and B was zero, TFIIIC appeared as a large blob on the DNA. However, with increasing distance between boxes A and B, TFIIIC appeared as two globular domains separated by a linker of increasing length between them. Thus, the combination of large size and stretchability allows TFIIIC to contact two widely separated promoter regions with its two globular domains.

SUMMARY Classical class III genes require two factors, TFIIIB and C, in order to form a preinitiation complex with the polymerase. The 5S rRNA genes also require TFIIIA. TFIIIC and A are assembly factors that bind to the internal promoter and help TFIIIB bind to a region just upstream of the transcription start site. TFIIIB then remains bound and can sponsor the initiation of repeated rounds of transcription. TFIIIC is a very large protein. The yeast protein has six subunits that are arranged into two globular regions joined through a flexible linker. The stretchability of this linker allows the protein to cover the long distance between boxes A and B of the internal promoter.

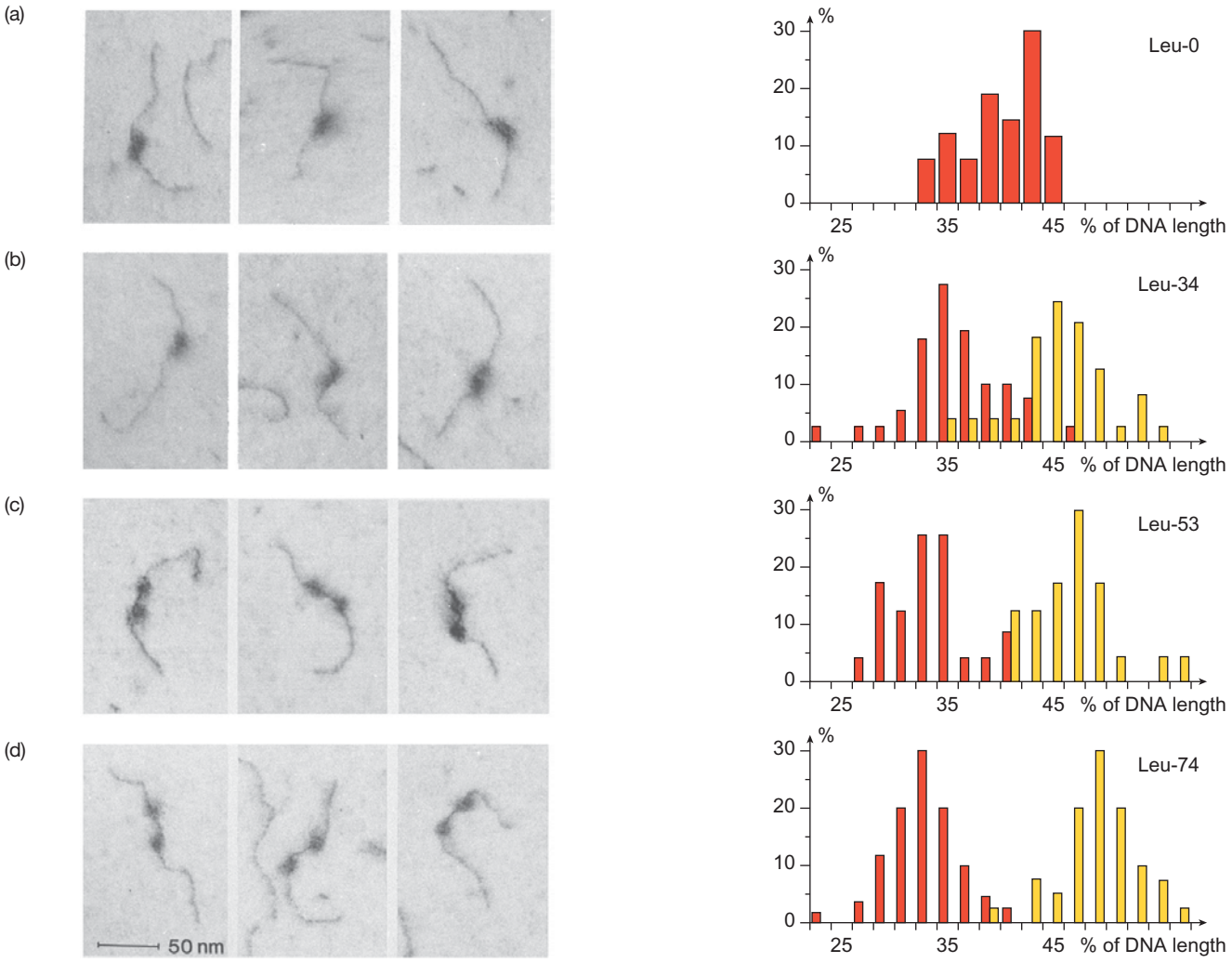


Figure 11.39 Yeast TFIIIC contains two globular domains connected by a flexible linker. Sentenac and colleagues bound yeast TFIIIC to cloned tRNA genes with variable distances between their boxes A and B. Next, they subjected the complexes to negative staining with uranyl acetate, then submitted them for scanning transmission electron microscopy. The distances between boxes A and B are given at right: **(a)** 0 bp; **(b)** 34 bp; **(c)** 53 bp; and **(d)** 74 bp, which is the wild-type distance. Three examples of micrographs with each DNA are presented at left. The histograms at right display the positions of the globular domains of TFIIIC on the DNA, determined from many different micrographs. The bars show the percentages of DNAs with globular domains at each location along the DNA. The red bars show the locations of the globular domain closest to the end of the DNA, and the yellow bars show the locations of the other globular domain. (*Source:* Schultz et al *EMBO Journal* 8: p. 3817 © 1989.)

The Role of TBP

If TFIIIC is necessary for TFIIIB binding in classical class III genes, what about nonclassical genes that have no boxes A or B to which TFIIIC can bind? What stimulates TFIIIB binding to these genes? Because the promoters of these genes have TATA boxes (Chapter 10), and we have already seen that TBP is required for their transcription, it makes sense to propose that the TBP binds to the TATA box and anchors TFIIIB to its upstream binding site.

But what about classical polymerase III genes? These have no TATA box, and yet we have seen that TBP is required for transcription of classical class III genes such as the tRNA and 5S rRNA genes in yeast and human cells. Where does TBP fit into this scheme? It has now become clear that TFIIIB contains TBP along with a small number of TAFs. In mammals, these TAFs are called Brf1 and Bdp1. Geiduschek and coworkers showed that TBP was present even in the purest preparations of TFIIIB. Further studies on yeast TFIIIB, including reconstitution from cloned components, have revealed that the factor is composed of three subunits: TBP and two TAF_{III}s. These two proteins have different names in different organisms. The yeast versions are called *B″* and *TFIIB-related factor,* or *BRF,* because of its homology to TFIIB.

Subsequently, Tjian and coworkers have shown by adding factors back to immunodepleted nuclear extracts that TRFI, not TBP, is essential for transcribing *Drosophila* tRNA, 5S rRNA and U6 snRNA genes. Thus, transcription by polymerase III in the fruit fly is another exception to the generality of dependence on TBP.

A unifying principle that emerges from the studies on transcription factors for all three RNA polymerases is that the assembly of a preinitiation complex starts with an assembly factor that recognizes a specific binding site in the promoter. This protein then recruits the other components of the preinitiation complex. For TATA-containing class II promoters, the assembly factor is usually TBP, and its binding site is the TATA box. This presumably applies to TATA-containing class III promoters as well, at least in yeast and human cells. We have already seen a model for how this process begins in TATA-containing class II promoters (Figure 11.4). Figure 11.40 shows, in highly schematic form, the nature of these preinitiation complexes for all kinds of TATA-less promoters. In class I promoters, the assembly factor is UBF, which binds to the UPE and then attracts the TBP-containing SL1 to the core element. TATA-less class II promoters can attract TBP in at least two ways. TAFs in TFIID can bind to core promoter elements, or they can bind to activators, such as Sp1 bound to proximal promoter elements, such as GC boxes. Both methods anchor TFIID to the TATA-less promoter. Classical class III promoters, at least in yeast and human cells, follow the same general scheme. TFIIIC, or in the case of the 5S rRNA genes, TFIIIA plus TFIIIC, play the role of assembly factor, binding to the internal promoter and attracting the TBP-containing TFIIIB to a site upstream of the start point. In *Drosophila* cells, TRFI appears to substitute for TBP in these preinitiation complexes.

Just because TBP does not always bind first, we should not discount its importance in organizing the preinitiation complex on these TATA-less promoters. Once TBP binds, it helps bring the remaining factors, including RNA polymerase, to the complex. This is a second unifying principle: TBP plays an organizing role in preinitiation complexes on most types of eukaryotic promoters. A third unifying principle is that the specificity of TBP is governed by the TAFs with which it associates; thus, TBP affiliates with different TAFs when it binds to each of the various kinds of promoter.

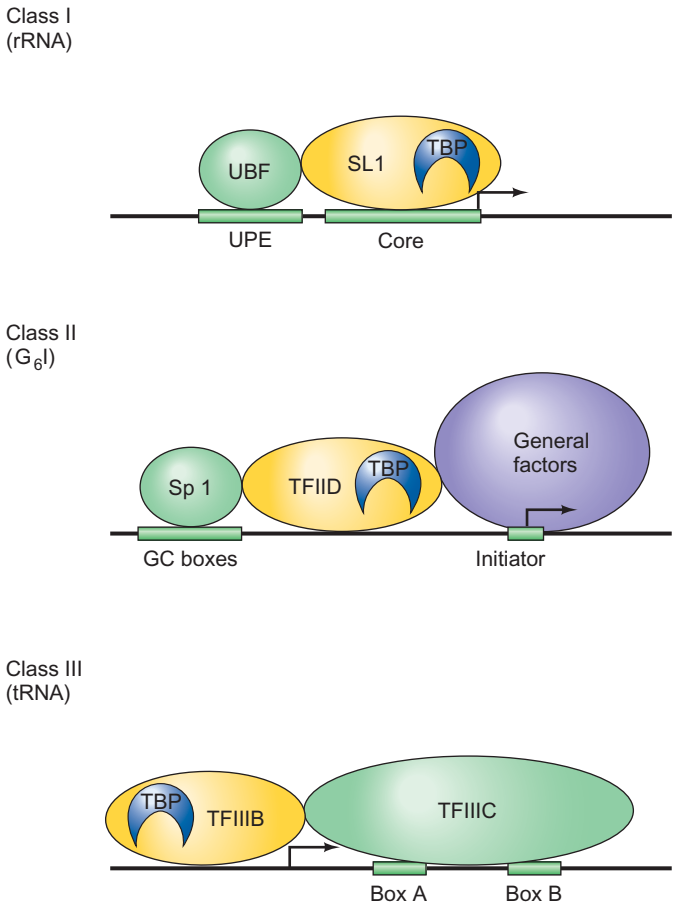


Figure 11.40 Model of preinitiation complexes on TATA-less promoters recognized by all three polymerases. In each case, an assembly factor (green) binds first (UBF, Sp1, and TFIIIC in class I, II, and III promoters, respectively). This in turn attracts another factor (yellow), which contains TBP (blue); this second factor is SL1, TFIID, or TFIIIB in class I, II, or III promoters, respectively. These complexes are sufficient to recruit polymerase for transcription of class I and III promoters, but in class II promoters more general factors (purple) besides polymerase II must bind before transcription can begin. (*Source:* Adapted from White, R.J. and S.P. Jackson, Mechanism of TATA-binding protein recruitment to a TATA-less class III promoter. *Cell* 71:1051, 1992.)

SUMMARY The assembly of the preinitiation complex on each kind of eukaryotic promoter begins with the binding of an assembly factor to the promoter. With TATA-containing class II (and presumably class III) promoters, this factor is TBP, but other promoters have their own assembly factors. Even if TBP is not the first-bound assembly factor at a given promoter, it becomes part of the growing preinitiation complex on most known promoters and serves an organizing function in building the complex. The specificity of the TBP—which kind of promoter it will bind to—depends on its associated TAFs. TRFI substitutes for TBP, at least in some preinitiation complexes in *Drosophila* class III genes.

SUMMARY

Transcription factors bind to class II promoters in the following order in vitro: (1) TFIID, apparently with help from TFIIA, binds to the TATA box. (2) TFIIB binds next. (3) TFIIF helps RNA polymerase II bind. The remaining factors bind in this order: TFIIE and TFIIH, forming the DABPolFEH preinitiation complex. The participation of TFIIA seems to be optional in vitro.

TFIID contains a TATA-box-binding protein (TBP) plus 13 other polypeptides known as TBP-associated factors (TAFs). The C-terminal 180 amino acid fragment of the human TBP is the TATA-box-binding domain. The interaction between a TBP and a TATA box takes place in the DNA minor groove. The saddle-shaped TBP lines up with the DNA, and the underside of the saddle forces open the minor groove and bends the TATA box through an 80-degree angle. TBP is required for transciption of most members of all three classes of genes, not just class II genes.

Most of the TAFs are evolutionarily conserved in the eukaryotes. They serve several functions, but two obvious ones are interacting with core promoter elements and interacting with gene-specific transcription factors. TAF1 and TAF2 help TFIID bind to the initiator and DPEs of promoters and therefore can enable TBP to bind to certain TATA-less promoters that contain such elements. TAF1 and TAF4 help TFIID interact with Sp1 that is bound to GC boxes upstream of the transcription start site. These TAFs therefore ensure that TBP can bind to TATA-less promoters that have GC boxes. Different combinations of TAFs are apparently required to respond to various transcription activators, at least in higher eukaryotes. TAF1 also has two enzymatic activities. It is a histone acetyltransferase and a protein kinase. TFIID is not universally required, at least in higher eukaryotes. Some promoters in *Drosophila* require an alternative factor, TRF1, and some promoters require a TBP-free TAF-containing complex.

Structural studies on a TFIIB-polymerase II complex show that TFIIB binds to TBP at the TATA box via its C-terminal domain, and to polymerase II via its N-terminal domain. This bridging action effects a coarse positioning of the polymerase active center about 25–30 bp downstream of the TATA box. In mammals, a loop motif of the N-terminal domain of TFIIB effects a fine positioning of the start of transcription by interacting with the single-stranded template DNA strand very near the active center. Biochemical studies confirm that the TFIIB N-terminal domain (the finger and linker domains, in particular) lies close to the RNA polymerase II active center, and to the largest subunit of TFIIF, in the preinitiation complex.

The preinitiation complex forms with the hypophosphorylated form to RNA polymerase II (IIA). Then, a subunit of TFIIH phosphorylates serine 5 in the heptad repeat in the carboxyl-terminal domain (CTD) of the largest RNA polymerase II subunit, creating the phosphorylated form of the enzyme (IIO). TFIIE greatly stimulates this process in vitro. This phosphorylation is essential for initiation of transcription. During the shift from initiation to elongation, phosphorylation shifts from serine 5 to serine 2. If phosphorylation of serine 2 is also lost, the polymerase pauses until re-phosphorylation by a non-TFIIH kinase occurs.

TFIIE and TFIIH are not essential for formation of an open promoter complex, or for elongation, but they are required for promoter clearance. TFIIH has a DNA helicase activity that is essential for transcription, presumably because it facilitates promoter clearance by fully melting the DNA at the promoter.

RNA polymerases can be induced to pause at specific sites near promoters by proteins such as DSIF and NELF. This pausing can be reversed by P-TEFb, which phosphorylates the polymerase, as well as DSIF and NELF. Polymerases that have backtracked and have become arrested can be rescued by TFIIS. This factor inserts into the active site of the polymerase, stimulates an RNase activity inherent in the polymerase, which cleaves off the 3′-end of the nascent RNA, extruded during backtracking. This allows resumption of elongation. TFIIS also stimulates proofreading, presumably by stimulating the RNase activity of RNA polymerase II, allowing it to remove misincorporated nucleotides.

Yeast and mammalian cells have been shown to contain an RNA polymerase II holoenzyme with many polypeptides in addition to the subunits of the polymerase.

Class I promoters are recognized by two transcription factors, a core-binding factor and a UPE-binding factor. The human core-binding factor is called SL1; in some other organisms, such as *A. castellanii,* the homologous factor is known as TIF-IB. The core-binding factor is the fundamental transcription factor required to recruit RNA polymerase I. This factor also determines species specificity, at least in animals. The factor that binds the UPE is called UBF in mammals and most other organisms, but UAF in yeast. It is an assembly factor that helps the core-binding factor bind to the core promoter element. The degree of reliance on the UPE-binding factor varies considerably from one organism to another. In *A. castellanii,* TIF-IB alone suffices to recruit the RNA polymerase I and position it correctly for initiation of transcription. Human UBF is a transcription factor that stimulates transcription by polymerase I. It can activate the intact promoter, or the core element alone, and it mediates activation by the UCE. UBF and SL1 act synergistically to stimulate transcription.

Human SL1 is composed of TBP and three TAFs, TAF_I110, TAF_I63, and TAF_I48. Fully functional and species-specific SL1 can be reconstituted from these purified components, and binding of TBP to the TAF_Is precludes binding to the TAF_{II}s. Other organisms have their own groups of TAF_Is.

Classical class III genes require two factors, TFIIIB and C, to form a preinitiation complex with the

polymerase. The 5S rRNA genes also require TFIIIA. TFIIIC and A are assembly factors that bind to the internal promoter and help TFIIIB bind to a region just upstream of the transcription start site. TFIIIB then remains bound and can sponsor the initiation of repeated rounds of transcription.

The assembly of the preinitiation complex on each kind of eukaryotic promoter begins with the binding of an assembly factor to the promoter. With TATA-containing class II (and presumably class III) promoters, this factor is usually TBP, but other promoters have their own assembly factors. Even if TBP is not the first-bound assembly factor at a given promoter, it becomes part of the growing preinitiation complex on most known promoters and serves an organizing function in building the complex. The specificity of the TBP—which kind of promoter it will bind to—depends on its associated TAFs, and there are TAFs specific for each of the promoter classes.

REVIEW QUESTIONS

1. List in order the proteins that assemble in vitro to form a class II preinitiation complex.
2. Describe and give the results of an experiment that shows that TFIID is the fundamental building block of the class II preinitiation complex.
3. Describe and give the results of an experiment that shows that TFIIF and polymerase II bind together, but neither can bind independently to the preinitiation complex.
4. Describe and give the results of an experiment that shows where TFIID binds.
5. Show the difference between the footprints caused by the DAB and the DABPolF complexes. What conclusion can you reach, based on this difference?
6. Present a hypothesis that explains the fact that substitution of dCs for dTs and dIs for dAs, in the TATA box (making a CICI box) has no effect on TFIID binding. Provide the rationale for your hypothesis.
7. What shape does TBP have? What is the geometry of interaction between TBP and the TATA box?
8. Describe and give the results of an experiment that shows TBP is required for transcription from all three classes of promoters.
9. Describe and give the results of an experiment that shows that a class II promoter is more active in vitro with TFIID than with TBP.
10. Describe and give the results of an experiment that identifies the TAFs that bind to a class II promoter containing a TATA box, an initiator, and a downstream promoter element.
11. Describe and give the results of a DNase footprinting experiment that shows how the footprint is expanded by TAF1 and TAF2 compared with TBP alone.
12. Draw a diagram of a model for the interaction of TBP (and other factors) with a TATA-less class II promoter.
13. Whole genome expression analysis indicates that yeast TAF1 is required for transcription of only 16% of yeast genes, and TAF9 is required for transcription of 67% of yeast genes. Provide a rationale for these results.
14. Present examples of class II preinitiation complexes with:
    a. An alternative TBP
    b. A missing TAF
    c. No TBP or TBP-like protein
15. What are the apparent roles of TFIIA and TFIIB in transcription?
16. Draw a rough sketch of the TBP–TFIIB–RNA polymerase II complex bound to DNA, showing the relative positions of the proteins. How do these positions correlate with the apparent roles of the proteins? Include an explanation of how TFIIB determines the direction of transcription.
17. Describe and give the results of an experiment that mapped the sites on Rpb1 and Rpb2 that are in close contact with the finger and linker regions of TFIIB.
18. Describe and give the results of an experiment that shows that TFIIH, but not the other general transcription factors, phosphorylates the IIA form of RNA polymerase II to the IIO form. In addition, include data that show that the other general transcription factors help TFIIH in this task.
19. Describe and give the results of an experiment that shows that TFIIH phosphorylates the CTD of polymerase II.
20. Describe an assay for DNA helicase and show how it can be used to demonstrate that TFIIH is associated with helicase activity.
21. Describe a G-less cassette transcription assay and show how it can be used to demonstrate that the RAD25 DNA helicase activity associated with TFIIH is required for transcription in vitro.
22. Draw a rough diagram of the class II preinitiation complex, showing the relative positions of the polymerase, the promoter DNA, TBP, and TFIIB, E, F, and H. Show the direction of transcription.
23. Describe and give the results of an experiment that shows that TFIIS stimulates transcription elongation by RNA polymerase II.
24. Present a model for reversal of transcription arrest by TFIIS. What part of TFIIS participates most directly? How?
25. Describe and give the results of an experiment that shows that TFIIS stimulates proofreading by RNA polymerase II.
26. What is the meaning of the term *RNA polymerase II holoenzyme?* How does the holoenzyme differ from the core polymerase II?
27. Describe and give the results of an experiment that shows the effect of adding or removing a few base pairs between the core element and the transcription start site in a class I promoter.
28. Which general transcription factor is the assembly factor in class I promoters? In other words, which binds first

and helps the other bind? Describe a DNase footprinting experiment you would perform to prove this, and show idealized results, not necessarily those that Tjian and colleagues actually obtained. Make sure your diagrams indicate an effect of both transcription factors on the footprints.

29. Describe and give the results of copurification and immunoprecipitation experiments that show that SL1 contains TBP.
30. Describe and give the results of an experiment that identified the TAFs in SL1.
31. How do we know that TFIIIA is necessary for transcription of 5S rRNA, but not tRNA, genes?
32. Geiduschek and colleagues performed DNase footprinting with polymerase III plus TFIIIB and C and a tRNA gene. Show the results they obtained with: No added protein; polymerase and factors; and polymerase, factors and three of the four NTPs. What can you conclude from these results?
33. The classical class III genes have internal promoters. Nevertheless, TFIIIB and C together cause a footprint in a region upstream of the gene's coding region. Draw a diagram of the binding of these two factors that explains these observations.
34. Draw a diagram of what happens to TFIIIB and C after polymerase III has begun transcribing a classical class III gene such as a tRNA gene. How does this explain how new polymerase III molecules can continue to transcribe the gene, even though factors may not remain bound to the internal promoter?
35. Describe and give the results of a DNase footprint experiment that shows that TFIIIB + C, but not TFIIIC alone, can protect a region upstream of the transcription start site in a tRNA gene. Show also what happens to the footprint when you strip off TFIIIC with heparin.
36. Describe and give the results of an experiment that shows the following: Once TFIIIB binds to a classical class III gene, it can support multiple rounds of transcription, even after TFIIIC (or C and A) are stripped off the promoter.
37. Describe and give the results of an experiment that demonstrates the flexibility of TFIIIC in binding to boxes A and B that are close together or far apart in a class III promoter.
38. Diagram the preinitiation complexes with all three classes of TATA-less promoters. Identify the assembly factors in each case.

ANALYTICAL QUESTIONS

1. You are studying a new class of eukaryotic promoters (class IV) recognized by a novel RNA polymerase IV. You discover two general transcription factors that are required for transcription from these promoters. Describe experiments you would perform to determine which, if any, is an assembly factor, and which is required to recruit the RNA polymerase to the promoter. Provide sample results of your experiments.
2. You discover that one of your novel class IV transcription factors contains TBP. Describe an experiment you would perform to identify the TAFs in this factor.
3. Some of the class IV promoters contain two DNA elements (boxes X and Y), others contain just one (box X). Describe experiments you would perform to identify the TAFs that bind to each of these two types of promoters.
4. You incubate cells with an inhibitor of the protein kinase activity of TFIIH and then perform in vitro transcription and DNase footprinting experiments. What step in transcription would you expect to see blocked? What kind of assay would reveal such a blockage? Would you still expect to see a footprint at the promoter? Why or why not? If so, how large would the footprint be, compared to the footprint in the absence of the inhibitor?
5. You know that protein X and protein Y interact, but you want to know whether a particular domain of protein X interacts with protein Y, and if so, where. Design a hydroxyl radical cleavage analysis experiment to answer this question.

SUGGESTED READINGS

General References and Reviews

Asturias, F.J. and J.L. Craighead. 2003. RNA polymerase II at initiation. *Proceedings of the National Academy of Sciences USA*. 100:6893–95.

Berk, A.J. 2000. TBP-like factors come into focus. *Cell* 103:5–8.

Buratowski, S. 1997. Multiple TATA-binding factors come back into style. *Cell* 91:13–15.

Burley, S.K. and R.G. Roeder. 1996. Biochemistry and structural biology of transcription factor IID (TFIID). *Annual Review of Biochemistry* 65:769–99.

Chao, D.M. and R.A. Young. 1996. Activation without a vital ingredient. *Nature* 383:119–20.

Conaway, R.C., S.E. Kong, and J.W. Conaway. 2003. TFIIS and GreB: Two like-minded transcription elongation factors with sticky fingers. *Cell* 114:272–74.

Goodrich, J.A., G. Cutler, and R. Tjian. 1996. Contacts in context: Promoter specificity and macromolecular interactions in transcription. *Cell* 84:825–30.

Grant, P. and J.L. Workman. 1998. A lesson in sharing? *Nature* 396:410–11.

Green, M.A. 1992. Transcriptional transgressions. *Nature* 357:364–65.

Hahn, S. 1998. The role of TAFs in RNA polymerase II transcription. *Cell* 95:579–82.

Hahn, S. 2004. Structure and mechanism of the RNA polymerase II transcription machinery. *Nature Structural & Molecular Biology* 11:394–403.

Klug, A. 1993. Opening the gateway. *Nature* 365:486–87.

Paule, M.R. and R.J. White. 2000. Transcription by RNA polymerases I and III. *Nucleic Acids Research* 28:1283–98.

Sharp, P.A. 1992. TATA-binding protein is a classless factor. *Cell* 68:819–21.
White, R.J. and S.P. Jackson. 1992. The TATA-binding protein: A central role in transcription by RNA polymerases I, II, and III. *Trends in Genetics* 8:284–88.

Research Articles

Armache, K.-J., H. Kettenberger, and P. Cramer. 2003. Architecture of initiation-competent 12-subunit RNA polymerase II. *Proceedings of the National Academy of Sciences USA*. 100:6964–68.
Bell, S.P., R.M. Learned, H.-M. Jantzen, and R. Tjian. 1988. Functional cooperativity between transcription factors UBF1 and SL1 mediates human ribosomal RNA synthesis. *Science* 241:1192–97.
Brand, M., C. Leurent, V. Mallouh, L. Tora, and P. Schultz. 1999. Three-dimensional structures of the TAF_{II}-containing complexes TFIID and TFTC. *Science* 286:2151–53.
Bushnell, D.A. and R.D. Kornberg. 2003. Complete, 12-subunit RNA polymerase II at 4.1-Å resolution: Implications for the initiation of transcription. *Proceedings of the National Academy of Sciences USA*. 100:6969–73.
Bushnell, D.A., K.D. Westover, R.E. Davis, and R.D. Kornberg. 2004. Stuctural basis of transcription: An RNA polymerase II-TFIIB cocrystal at 4.5 angstroms. *Science* 303:983–88.
Chen, H.-T. and S. Hahn. 2004. Mapping the location of TFIIB within the RNA polymerase II transcription preinitiation complex: A model for the structure of the PIC. *Cell* 119: 169–80.
Dynlacht, B.D., T. Hoey, and R. Tjian. (1991). Isolation of coactivators associated with the TATA-binding protein that mediate transcriptional activation. *Cell* 66:563–76.
Flores, O., H. Lu, M. Killeen, J. Greenblatt, Z.F. Burton, and D. Reinberg. 1991. The small subunit of transcription factor IIF recruits RNA polymerase II into the preinitiation complex. *Proceedings of the National Academy of Sciences USA* 88:9999–10003.
Flores, O., E. Maldonado, and D. Reinberg. 1989. Factors involved in specific transcription by mammalian RNA polymerase II: Factors IIE and IIF independently interact with RNA polymerase II. *Journal of Biological Chemistry* 264:8913–21.
Guzder, S.N., P. Sung, V. Bailly, L. Prakash, and S. Prakash. 1994. RAD25 is a DNA helicase required for DNA repair and RNA polymerase II transcription. *Nature* 369:578–81.
Hansen, S.K., S. Takada, R.H. Jacobson, J.T. Lis, and R. Tjian. 1997. Transcription properties of a cell type-specific TATA binding protein, TRF. *Cell* 91:71–83.
Holmes, M.C. and R. Tjian. 2000. Promoter-selective properties of the TBP-related factor TRF1. *Science* 288:867–70.
Holstege, F.C.P., E.G. Jennings, J.J. Wyrick, T.I. Lee, C.J. Hengartner, M.R. Green, T.R. Golub, E.S. Lander, and R.A. Young. 1998. Dissecting the regulatory circuitry of a eukaryotic genome. *Cell* 95:717–28.
Honda, B.M. and R.G. Roeder. 1980. Association of a 5S gene transcription factor with 5S RNA and altered levels of the factor during cell differentiation. *Cell* 22:119–26.
Kassavetis, G.A., B.R. Braun, L.H. Nguyen, and E.P. Geiduschek. 1990. *S. cerevisiae* TFIIIB is the transcription initiation factor proper of RNA polymerase III, while TFIIIA and TFIIIC are assembly factors. *Cell* 60:235–45.
Kassavetis, G.A., D.L. Riggs, R. Negri, L.H. Nguyen, and E.P. Geiduschek. 1989. Transcription factor IIIB generates extended DNA interactions in RNA polymerase III transcription complexes on tRNA genes. *Molecular and Cellular Biology* 9:2551–66.
Kettenberger, H., K.-J. Armache, and P. Cramer. 2003. Architecture of the RNA polymerase II-TFIIS complex and implications for mRNA cleavage. *Cell* 114:347–57.
Kim, J.L., D.B. Nikolov, and S.K. Burley. 1993. Co-crystal structure of a TBP recognizing the minor groove of a TATA element. *Nature* 365:520–27.
Kim, T.-K., R.H. Ebright, and D. Reinberg. 2000. Mechanism of ATP-dependent promoter melting by transcription factor IIH. *Science* 288:1418–21.
Kim, Y.J., S. Björklund, Y. Li, M.H. Sayre, and R.D. Kornberg. 1994. A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II. *Cell* 77:599–608.
Koleske, A.J. and R.A. Young. 1994. An RNA polymerase II holoenzyme responsive to activators. *Nature* 368:466–69.
Kownin, P., E. Bateman, and M.R. Paule. 1987. Eukaryotic RNA polymerase I promoter binding is directed by protein contacts with transcription initiation factor and is DNA sequence-independent. *Cell* 50:693–99.
Learned, R.M., S. Cordes, and R. Tjian. 1985. Purification and characterization of a transcription factor that confers promoter specificity to human RNA polymerase I. *Molecular and Cellular Biology* 5:1358–69.
Lobo, S.L., M. Tanaka, M.L. Sullivan, and N. Hernandez. 1992. A TBP complex essential for transcription from TATA-less but not TATA-containing RNA polymerase III promoters is part of the TFIIIB fraction. *Cell* 71:1029–40.
Lu, H., L. Zawel, L. Fisher, J.-M. Egly, and D. Reinberg. 1992. Human general transcription factor IIH phosphorylates the C-terminal domain of RNA polymerase II. *Nature* 358:641–45.
Maldonado, E., I. Ha, P. Cortes, L. Weis, and D. Reinberg. 1990. Factors involved in specific transcription by mammalian RNA polymerase II: Role of transcription factors IIA, IID, and IIB during formation of a transcription-competent complex. *Molecular and Cellular Biology* 10:6335–47.
Ossipow, V., J.-P. Tassan, E.I. Nigg, and U. Schibler. 1995. A mammalian RNA polymerase II holoenzyme containing all components required for promoter-specific transcription initiation. *Cell* 83:137–46.
Pelham, H.B., Wormington, W.M., and D.D. Brown. 1981. Related 5S rRNA transcription factors in *Xenopus* oocytes and somatic cells. *Proceeding of the National Academy of Sciences USA* 78:1760–64.
Pugh, B.F. and R. Tjian. 1991. Transcription from a TATA-less promoter requires a multisubunit TFIID complex. *Genes and Development* 5:1935–45.
Rowlands, T., P. Baumann, and S.P. Jackson. 1994. The TATA-binding protein: A general transcription factor in eukaryotes and archaebacteria. *Science* 264:1326–29.
Sauer, F., D.A. Wassarman, G.M. Rubin, and R. Tjian. 1996. TAF_{II}s mediate activation of transcription in the *Drosophila* embryo. *Cell* 87:1271–84.

Schultz, M.C., R.H. Roeder, and S. Hahn. 1992. Variants of the TATA-binding protein can distinguish subsets of RNA polymerase I, II, and III promoters. *Cell* 69:697–702.

Setzer, D.R. and D.D. Brown. 1985. Formation and stability of the 5S RNA transcription complex. *Journal of Biological Chemistry* 260:2483–92.

Shastry, B.S., S.-Y. Ng, and R.G. Roeder. 1982. Multiple factors involved in the transcription of class III genes in *Xenopus laevis*. *Journal of Biological Chemistry* 257:12979–86.

Starr, D.B. and D.K. Hawley. 1991. TFIID binds in the minor groove of the TATA box. *Cell* 67:1231–40.

Taggart, K.P., J.S. Fisher, and B.F. Pugh. 1992. The TATA-binding protein and associated factors are components of Pol III transcription factor TFIIIB. *Cell* 71:1051–28.

Takada, S., J.T. Lis, S. Zhou, and R. Tjian. 2000. A TRF1:BRF complex directs *Drosophila* RNA polymerase III transcription. *Cell* 101:459–69.

Tanese, N. 1991. Coactivators for a proline-rich activator purified from the multisubunit human TFIID complex. *Genes and Development* 5:2212–24.

Thomas, M.J., A.A. Platas, and D.K. Hawley. 1998. Transcriptional fidelity and proofreading by RNA polymerase II. *Cell* 93:627–37.

Verrijzer, C.P., J.-L. Chen, K. Yokomori, and R. Tjian. 1995. Binding of TAFs to core elements directs promoter selectivity by RNA polymerase II. *Cell* 81:1115–25.

Walker, S.S., J.C. Reese, L.M. Apone, and M.R. Green. 1996. Transcription activation in cells lacking TAF_{II}s. *Nature* 383:185–88.

Wieczorek, E., M. Brand, X. Jacq, and L. Tora. 1998. Function of TAF_{II}-containing complex without TBP in transcription by RNA polymerase II. *Nature* 393:187–91.