Identification of a minimal set of proteins that is sufficient for accurate initiation of transcription by RNA polymerase II

Curtis M. Tyree, Catherine P. George, Lucy M. Lira-DeVito, Sharon L. Wampler, Michael E. Dahmus, Leigh Zawel, and James T. Kadonaga

Department of Biology and Center for Molecular Genetics, University of California, San Diego La Jolla, California 92093 USA; Department of Biochemistry and Biophysics, University of California, Davis, California 95616 USA; University of Medicine and Dentistry of New Jersey, Robert Wood Johnson Medical School Department of Biochemistry, Piscataway, New Jersey 08854 USA

In eukaryotes, initiation of mRNA synthesis is a multistep process that is carried out by RNA polymerase II and auxiliary factors that are commonly referred to as basal or general factors. In this study accurate initiation of transcription was reconstituted with purified, Escherichia coli-synthesized TFIIB, TBP (the TATA box-binding polypeptide of the TFIID complex), and the 30-kD subunit of TFIIF (also known as RAP30), along with purified, native RNA polymerase II from Drosophila embryos, calf thymus, or HeLa cells. This minimal set of factors was able to transcribe a subset of the promoters tested. The addition of both subunits of TFIIE and the 74-kD subunit of TFIIF increased the efficiency of transcription by a factor of 2 to 4. In contrast, the inclusion of a crude TFIID fraction from Drosophila embryos in place of recombinant TBP resulted in a strong dependence on TFIIE. By gel mobility-shift analysis, TFIIB, TBP, RAP30, and polymerase were able to assemble into DB and DBPolF30 complexes with transcriptionally competent (wild type or initiator mutant), but not with transcriptionally inactive (TATA and TATA/initiator mutant), versions of the Drosophila Adh promoter. Thus, it appears that RNA polymerase II is able to initiate transcription subsequent to assembly of the DBPolF30 complex, which is a minitranscription complex that represents the central core of the RNA polymerase II transcriptional machinery.

[Key Words: RNA polymerase II; basal transcription factors; TFIIB; TATA box-binding protein; TFIIF (RAP30/74); in vitro transcription]

Received March 8, 1993; revised version accepted April 14, 1993.

The control of gene expression is a problem of fundamental importance and significant complexity. At the level of transcription, genes appear to be rendered either competent or inert at different stages or tiers, which range from global alterations in chromatin structure (for review, see Grunstein 1990; Kornberg and Lorch 1991; Felsenfeld 1992; Roeder 1991; Adams and Workman 1993; Croston and Kadonaga 1993) to the action of regulatory factors, such as sequence-specific transcriptional activators that interact with promoter and enhancer regions (for review, see Ptashne 1988; Johnson and McKnight 1989; Mitchell and Tjian 1989). The central component in transcription is, however, the minimal set of factors and RNA polymerase II, which catalyze a basal level of RNA synthesis (for review, see Saltzman and Weinmann 1989; Sawadogo and Sentenac 1990; Conaway and Conaway 1991; Roeder 1991; Pugh and Tjian 1992; Weis and Reinberg 1992; Zawel and Reinberg 1992, 1993; Drapkin et al. 1993).

In eukaryotes, synthesis of mRNA is carried out by the RNA polymerase II transcriptional machinery. RNA polymerase II is a multisubunit enzyme that can catalyze nonspecifically the synthesis of RNA from a DNA template, but requires additional factors, commonly referred to as the general or basal factors, to initiate transcription accurately at promoters and to carry out the elongation reaction efficiently. The basal factors include transcription factor IIA (TFIIA), TFIIB, TFIID, TFIIE, TFIIF, TFIH, TFIJ, and TFIIS (for nomenclature and abbreviations, see Materials and methods). Biochemical studies with reconstituted in vitro transcription systems have revealed that TFIIB, TFIID, TFIIE, TFIIF, and TFIJ facilitate initiation of transcription, TFIIS in-
creases the efficiency of elongation, and TFIIH has a dual function in both initiation and elongation.

It is important to elucidate the identity and function of the basal transcription factors. Although the basal factors have been highly purified from native sources, the exact set of proteins that comprises the basal transcription machinery has not been elucidated. This issue is further complicated by recent data suggesting that basal transcription at different promoters occurs by distinct mechanisms [Parvin et al. 1992, Wampler and Kadonaga 1992]. Thus, different genes may not have the same requirements for basal transcription factors. It is also important to consider whether or not each of the basal factors functions directly to synthesize RNA. Some basal factors may participate in the transcription process, whereas others may counteract the action of transcriptional inhibitors. For example, the role of TFIIA in basal transcription has been a subject of controversy. Depending on the reaction conditions, the requirement for TFIIA in basal transcription has been observed to be strong, weak, or nonexistent [Sawadogo and Roeder 1985; Wampler et al. 1990; Cortes et al. 1992; Ranish et al. 1992; Sayre et al. 1992]. Notwithstanding, recent work has led to the purification and cloning of TFIIA, TFIIIB, TBP (the TATA box-binding polypeptide of the native TFIID complex; Pugh and Tjian 1992), TFIIIE, TFIIIF, two subunits of TFIIH, and TFIIIS [see references cited in Zawel and Reinberg (1992, 1993) and Drapkin et al. (1993) in addition to Aso et al. 1992; Finkelstein et al. 1992; Fischer et al. 1992; Gildea et al. 1992].

The analysis of basal transcription has been founded on the development of fractionated in vitro transcription systems, and we have chosen Drosophila embryos as the source of factors because of the high transcriptional activity of Drosophila extracts [Kadonaga 1990, Kamakaka et al. 1991], the low cost of embryos, and the functional conservation of the basal transcription factors between Drosophila and humans [Wampler et al. 1990]. This work has led to the biochemical separation of Drosophila embryo extracts into purified RNA polymerase II and fractions containing TFIIIB, TFII/E/F, and TFIID, where the identity of the Drosophila homologs of the HeLa transcription factors was deduced from cross-substitution experiments in which the Drosophila fractions were substituted into a fractionated HeLa transcription system and vice versa [Wampler et al. 1990]. There was no apparent TFIIA activity in this fractionated system.

The development of the Drosophila transcription system has provided an independent means of examination of basal transcription that was distinct from the well-characterized mammalian and yeast systems. Moreover, to analyze basal transcription, we have transcribed a set of different promoters, including those of cellular Drosophila genes. We have also monitored transcriptional activity by primer extension analysis, which requires synthesis of transcripts <100 nucleotides, rather than by runoff analysis, which requires synthesis of transcripts of several hundred nucleotides. In this manner, it has been possible to focus primarily on initiation of transcription as well as to investigate potential variations in the transcriptional properties of different promoters [Parvin et al. 1992, Wampler and Kadonaga 1992]. In this study we describe the use of the fractionated Drosophila transcription system to elucidate the minimal components that are sufficient for accurate initiation of transcription by RNA polymerase II.

Results and Discussion

Transcription with purified RNA polymerase II and recombinant basal factors

By using purified Escherichia coli-synthesized basal factors, we have been able to compare the biochemical activity of each of the recombinant transcription factors with that of the native protein fractions in the Drosophila transcription system. For example, purified, E. coli-synthesized dTFIIB was observed to possess activity for both basal and sequence-specific factor-activated transcription that was indistinguishable from that of the native protein [Wampler and Kadonaga 1992, see also Ha et al. 1991]. The comparison of purified, recombinant dTBP with native TFIIID however, was not as straightforward. For instance, it was found that TBP was the TATA box-binding component of a larger, multisubunit TFIIID complex and that TBP was active for basal, but not factor-activated transcription [for review, see Pugh and Tjian 1992]. In addition, we observed that recombinant dTBP could substitute for the Drosophila TFIIID fraction with some, but not all, promoters for basal transcription [data not shown]. These initial experiments provided a reconstituted transcription system that used purified, recombinant dTFIIB along with a purified, albeit abbreviated, version of TFIIID.

We then proceeded to carry out transcription in reactions containing purified, recombinant dTFIIB, dTBP, hTFIIIE [Ohkuma et al. 1991; Peterson et al. 1991; Sumimoto et al. 1991], and hTFIIF [Sopta et al. 1989, Aso et al. 1992; Finkelstein et al. 1992], along with purified RNA polymerase II from Drosophila embryos and the native (impure) fractions in the Drosophila transcription system [Wampler et al. 1990]. The purity of each of the recombinant proteins and the RNA polymerase II is shown in Figure 1A. In an extension of our comparative study of dTBP versus native TFIIID fraction, we carried out a survey of different promoters, including the Drosophila alcohol dehydrogenase [Adh] [proximal], Drosophila Krüppel, Drosophila hunchback [P1 and P2], Drosophila Fockey, Drosophila hsp70, adenovirus major late [AdML], adenovirus E1B, and adenovirus E4 promoters and found that accurate transcription could be reconstituted with a combination of purified, recombinant dTFIIB, hTFIIIE, and hTFIIIF, along with the native TFIIID fraction [Wampler et al. 1990] and purified RNA polymerase II from Drosophila embryos. In most instances, dTBP was a poor substitute for the native TFIIID fraction. With the Drosophila Adh proximal promoter, however, dTBP could substitute effectively for the TFIIID fraction. It was thus possible to transcribe the Adh promoter with purified, recombinant dTFIIB, dTBP, hTFIIIE, hTFIIIF, and
Tyree et al.

**Figure 1.** The *Drosophila* alcohol dehydrogenase (Adh) proximal promoter can be transcribed accurately in vitro with purified, recombinant basal transcription factors and RNA polymerase II purified from *Drosophila* embryos. The nomenclature and abbreviations of the basal transcription factors are described in Materials and methods. (A) Analysis of the purified transcription factors by SDS-PAGE. The proteins were visualized by silver staining. (Left) 10% SDS-PAGE of *E. coli*-synthesized, basal transcription factors. (Lane 1) dTBP (60 ng); (lane 2) dTFIIB (40 ng); (lane 3) hTFIIE34 (80 ng); (lane 4) hTFIIE56 (90 ng); (lane 5) hTFIIF30 (60 ng); (lane 6) hTFIIF74 (500 ng—the protein stained better with Coomassie blue than silver). (Right) 4-25% polyacrylamide–SDS gradient gel electrophoresis of RNA polymerase II from *Drosophila* embryos (500 ng protein). The sizes of molecular mass markers [in kD] are indicated. The IIA and IIB forms of the largest subunit of the polymerase are noted. (B) Transcription from the Adh proximal promoter. Reactions were performed with pPA386-56 (100 ng) template DNA at 21°C. The complete, reconstituted transcription reaction is shown in lanes 8 and 10. The complete reaction [lanes 8–11] contained dTFIIB (5 ng), dTBP (100 ng), hTFIIE34 (2 ng), hTFIIE56 (10 ng), hTFIIF30 (6 ng), hTFIIF74 (10 ng), and *Drosophila* RNA polymerase II (20 ng). In lanes 1–7, the indicated transcription factors were omitted from the otherwise complete reaction. The minimal set [lanes 12,13] contained dTFIIB (5 ng), dTBP (100 ng), hTFIIF30 (6 ng), and *Drosophila* RNA polymerase II (20 ng). As a control, α-amanitin was included in the reactions at a concentration of 4 μg/ml in lanes 9, 11, and 13. The reverse transcription products are indicated by brackets.

purified RNA polymerase II [Fig. 1B]. Transcription initiated accurately at the correct in vivo start sites [data not shown] and was completely inhibited by 4 μg/ml of α-amanitin, which indicates that RNA synthesis was carried out by RNA polymerase II. Transcription was strongly dependent on the presence of dTFIIB, dTBP, hTFIIF30, and RNA polymerase II, whereas omission of hTFIIE34, hTFIIE56, or hTFIIF74 from the reaction did not result in a significant decrease in RNA synthesis. These findings suggested that dTFIIB, dTBP, hTFIIF30, and RNA polymerase II may comprise a core, or minimal set, of factors, required for basal transcription. Therefore, we performed transcription reactions with either this minimal set of factors or the complete set [dTFIIB, dTBP, hTFIIE, hTFIIF, and polymerase] and found that both yielded accurate transcription from the Adh promoter, although the efficiency of transcription with the minimal set was ~33% of that with the complete set. These data suggested that TFIIB, TBP, TFIIF30, and polymerase may constitute a central, catalytic core of the RNA polymerase II transcriptional machinery.

The properties of the purified transcription system

We then examined transcription with a homologous, mammalian-derived system with the human versions of TFIIB, TBP, TFIIE, and TFIIF along with RNA polymerases IIA or IIB from calf thymus. The IIA and IIB forms of the polymerase, which either possess [IIA form] or lack [by proteolysis during purification of the enzyme; IIB form] the carboxy-terminal heptapeptide repeat in the largest subunit [carboxy-terminal domain, or CTD; for review, see Peterson and Tjian 1992; Zawel and Reinberg 1992, 1993], were each purified to near homogeneity from calf thymus [Fig. 2A]. The calf thymus polymerases were therefore used to test the transcriptional requirement for the CTD, as well as to examine the ability of different preparations of RNA polymerase II to carry out accurate transcription with the purified basal factors [Fig. 2B]. The Adh promoter was transcribed accurately with the purified, mammalian-derived system with either the IIA or the IIB form of the polymerase. Hence, with the purified transcription factors, the presence or the absence of the CTD did not affect basal transcription from the Adh promoter, which was consistent with results obtained with impure transcription systems with the *Drosophila* actin 5C and AdML promoters [Zehring et al. 1988; Kim and Dahmus 1989; Buratowski and Sharp 1990]. The strong dependence on TBP, TFIIB, TFIIF30, and polymerase for transcription, as well as the weak requirement for TFIIE34, TFIIE56, and TFIIF74, was identical to that observed with the mixed *Drosophila* embryos.
with either the complete or minimal set of factors. Thus, did not observe transcription from the other promoters Drosophila hsp70, Drosophila Jockey, purified transcription system. different genes, we carried out a series of reactions with a variety of promoters, including the Adh, proximal promoter. These results suggest further that the CTD is not essential for transcription with the purified components. Finally, because similar results were obtained with the mammalian transcription system [Fig. 2B] and the heterologous Drosophila and human system [Fig. 1B], the use of the mixed system does not appear to be a critical issue in the evaluation of the purified transcription system.

To test the ability of the purified factors to transcribe different genes, we carried out a series of reactions with a variety of promoters, including the Drosophila Adh, Drosophila Krüppel, Drosophila hunchback (P1 and P2), Drosophila hsp70, Drosophila Jockey, AdML, adenovirus E1B, adenovirus E4, human immunodeficiency virus I[HIV-1], HIV-2, and human apolipoprotein A1 promoters. In these experiments we observed reproducible, accurate transcription from only the Drosophila Adh proximal, Drosophila Krüppel, and AdML promoters with the complete and minimal sets of factors (Fig. 3A). We did not observe transcription from the other promoters with either the complete or minimal set of factors. Thus, the purified factors do not possess the ability to transcribe all genes but, rather, a subset of genes. These findings indicate that the Adh promoter is not unique in its ability to be transcribed with the purified factors but, rather, that this property is shared by other promoters, including the commonly studied AdML promoter. Both the AdML (Fig. 3B) and the Krüppel [data not shown] promoters displayed the same basal factor dependence as the Adh promoter. Moreover, the AdML promoter could be transcribed with either RNA polymerase IIA or IIB. Hence, under the conditions used in this study, the basal transcription factor requirements of the Drosophila Adh proximal, Drosophila Krüppel, and AdML promoters were indistinguishable. Examination of the DNA sequences of the promoters used in these experiments did not, however, reveal any similarity between the Adh, Krüppel, and major late promoters; therefore, promoter elements that might mediate transcription by the purified factors are not yet apparent.

An important point that remains to be addressed is the possibility that the preparations of RNA polymerase II used in this study contained small, but transcriptionally active amounts of basal factors, such as TFIIE, TFIIF, TFIHH, or TFIIF. Therefore, we tested the ability of purified RNA polymerase II from HeLa cells [Lu et al. 1991] to transcribe the AdML promoter along with the complete (TFIIB, TBP, TFIIE, and TFIIF) and minimal (TFIIB, TBP, hTFIIF30) sets of basal factors. This preparation of HeLa RNA polymerase II has been used for the fractionation and purification of basal transcription factors from

---

**Figure 2.** Transcription from the Adh proximal promoter with the human versions of the purified, basal transcription factors and RNA polymerases IIA and IIB purified from calf thymus. (A) Analysis of the calf thymus RNA polymerase IIA and IIB purified by 4–25% polyacrylamide–SDS gradient gel electrophoresis. (Lane 1) RNA polymerase IIA [1.1 μg]; [lane 2] RNA polymerase IIB [3 μg]. The polypeptides were visualized by silver staining. The sizes of molecular mass markers (in kD) are indicated. The IIA and IIB forms of the largest subunit of the polymerase are also noted. (B) Transcription from the Adh proximal promoter. Reactions were performed with pPD386-56 [100 ng] template DNA at 30°C. The complete, reconstituted transcription reaction with RNA polymerase IIA is shown in lane 8. The complete reaction [lanes 8–11] contained hTFIIB [30 ng], hTBP [50 ng], hTFIE34 [2 ng], hTFIE56 [10 ng], hTFIIF30 [6 ng], hTFIIF74 [10 ng], and calf thymus RNA polymerases [37 ng].
Figure 3. The *Drosophila* Krüppel, AdML, and *Drosophila* Adh promoters can be transcribed with the purified basal transcription factors and RNA polymerase II purified from calf thymus. Complete transcription reactions contained dTFIIB (5 ng), dTBP (100 ng), hTFIIE34 (2 ng), hTFIIE56 (10 ng), hTFIIF30 (6 ng), hTFIIF74 (10 ng), and calf thymus RNA polymerases IIA or IIB, as noted (37 ng). Minimal reactions contained dTFIIB (5 ng), dTBP (100 ng), hTFIIF30 (6 ng), and calf thymus RNA polymerases IIA or IIB, as indicated (37 ng). The reverse transcription products are designated by brackets. (A) Transcription from the Adh, Krüppel, and AdML promoters. Reactions were performed with pPA386-56, pKr-31/+13, or pLAX template DNAs (100 ng) at 21°C. (B) The AdML promoter displays the same transcription factor requirement as the *Drosophila* Adh promoter. Reactions were carried out with pLAX (100 ng) template DNA at 25°C. The complete, reconstituted transcription reaction with calf thymus RNA polymerase IIA is shown in lane 8. In lanes 1–7, the indicated factors were omitted from the otherwise complete reactions containing calf thymus RNA polymerase IIA. In lanes 8–15, the complete or minimal sets of transcription factors were transcribed with RNA polymerase IIA or IIB. In lanes 16–18, HeLa RNA polymerase II (90 ng) was used instead of the calf thymus enzyme. As a control, α-amanitin was included in the reactions at a concentration of 4 μg/ml in lanes 9, 11, 13, 15, and 18. The reverse transcription products are indicated by bracket.

We were also able to examine whether the RNA polymerase II from *Drosophila* was contaminated with biochemically active amounts of TFIIE by using a partially purified fraction containing native, but impure, TFIID instead of purified, recombinant TBP. As shown previously in Figure 1B, when transcription reactions were performed with dTBP, dTFIIB, hTFIIE, hTFIIF, and *Drosophila* RNA polymerase II, omission of either subunit of TFIIE resulted in only a slight decrease in transcription. This absence of a requirement for TFIIE could...
have been attributable to the presence of contaminating TFIIE in the purified RNA polymerase II. Yet, when transcription reactions were carried out under identical conditions, except for the use of a native, impure TFIID fraction instead of recombinant dTBP, there was a strong requirement for both subunits of recombinant hTFIIE (Fig. 4). Therefore, it appears unlikely that the TFIIE insensitivity of the purified transcription system was attributable to TFIIE contamination of the polymerase because identical amounts of the same preparations of hTFIIE and polymerase were used in both experiments (Figs. 1B and 4). Rather, the ability of TFIIE to stimulate basal transcription was dependent on the use of the impure TFIID fraction instead of recombinant dTBP. This effect could have been the result of either a defect in the recombinant TBP that circumvented a normal requirement for TFIIE, or alternatively, the presence of a transcriptional inhibitor in the TFIID fraction whose repression was counteracted by TFIIE.

It is important to consider the apparent absence of a requirement for the other basal factors, such as TFIIF74, TFIIF, and TFIIFH, because it has been well established that transcription of the AdML promoter requires TFIIB, TBP, TFIIE, TFIIF, TFIIF, and TFIIFH in addition to RNA polymerase II. A few possible explanations are as follows. First, as mentioned in the introductory section, we used a primer extension assay for the analysis of the in vitro-synthesized transcripts instead of the more commonly used runoff assay. For the detection of RNA, the primer extension assay enables the use of either supercoiled or linear DNA templates (in this study we transcribed supercoiled DNA), yields high resolution (single nucleotide) of the transcription start sites, and focuses primarily on initiation of transcription because it requires only ~70 nucleotides of synthesized RNA. In contrast, the runoff assay is usually performed with linear DNA templates and requires the synthesis of a transcript of several hundred nucleotides. The differences in the basal factor requirements may be attributable to the use of the different transcription assays, and additional basal factors beyond the minimal set may be required for initiation from linear versus supercoiled templates, as well as for efficient elongation of the transcripts. For instance, the requirement for TFIIFH, which has been shown recently to possess DNA helicase activity (Schaefter et al. 1993), may be circumvented by the use of negatively supercoiled, rather than linear, DNA templates. Second, the remaining basal factors might function to increase the efficiency of transcription by the minimal set of proteins. With the purified factors, the efficiency of transcription from the Adh and AdML promoters was typically in the range of 0.002–0.005 transcripts/template per 30 min, which is comparable to that observed with crude HeLa extracts (Weil et al. 1979; Manley et al. 1980; Dignam et al. 1983), but considerably lower than the high level of transcription that has been obtained with a soluble nuclear fraction from Drosophila embryos (0.45 transcripts/template per 30 min; Kamakaka et al. 1991). Third, the use of the recombinant factors may uncouple or even negate the requirement for basal factors by a variety of mechanisms. Finally, it is also important to note that the purified factors were able to transcribe only a subset of promoters. These differences in the basal factor requirements with different promoters suggest that some basal transcription factors may be involved in the transcription of distinct subsets or classes of genes.

### Figure 4.
TFIIE is essential for transcription when the reactions are reconstituted with a native, but impure TFIID fraction from Drosophila embryos. The Adh proximal promoter (pPA386-56; 100 ng) was transcribed with native TFIID fraction (6 μg) [Wampler et al. 1990], dTFIIB (5 ng), hTFIIF30 (6 ng), hTFIIF74 (10 ng), and Drosophila RNA polymerase II (20 ng). hTFIIE34 (2 ng) and/or hTFIIE56 (10 ng) were also included in the reactions as indicated. As a control, α-amanitin was included in the reactions at a concentration of 4 μg/ml in lane 5. The reverse transcription products are designated by bracket.

### Gel mobility-shift analysis of transcription factor–template DNA complexes
To monitor the assembly of putative transcription complexes, we performed a gel mobility-shift analysis (Bartowski et al. 1989; Maldonado et al. 1990) of the basal factor–template DNA complexes. In studies of the AdML promoter, gel mobility-shift experiments with fractionated transcription factors have led to the hypothesis that the order of assembly of basal factors into a transcription initiation complex is as follows (Bartowski et al. 1989; Maldonado et al. 1990; Flores et al. 1992, Killeen et al. 1992): TFIID and TFIIB bind to the template DNA to form a TFIID–TFIIB (DB) complex; RNA polymerase II and TFIIF assemble onto the DB complex to give a DPaII complex; TFIIE can then be incorporated to give a DPaIIFE complex, which is then followed by the sequential binding of TFIIFH and TFIIF to
Tyree et al. give the final DBPolFEHJ complex. TFIIA can assist the binding of TFIID to the template DNA and can associate with each of the intermediate complexes [Cortes et al. 1992] but is not essential for the assembly of the transcription complex or for transcription in vitro. In addition, Killean et al. (1992) have shown that purified, recombinant yeast TBP, hTFIIB, and hTFIIF30, along with purified calf thymus RNA polymerase II, were able to form a DBPolF30 complex in a gel-shift assay with the AdML promoter.

In our experiments with the purified basal factors, we used a template DNA probe containing a truncated version of the Drosophila Adh proximal promoter, termed pAdh-40/+25, which possessed sequences of the Adh promoter from -40 to +25 relative to the RNA start site. The purified, recombinant basal factors along with Drosophila RNA polymerase II were able to assemble into DB and DBPolF complexes on the Adh promoter (Fig. 5A). In accord with the previous findings of Peterson et al. (1991), incorporation of TFIIIE into the DBPolF complex required both subunits of TFIIE [Fig. 5B]. Moreover, TFIIF30 and TFIIF74 were each able to form their respective DBPolF30 and DBPolF74 complexes, whereas the presence of both the 30- and 74-kD subunits of TFIIF appeared to increase the efficiency of the assembly of the DBPolF complex. We also examined inhibition of complex assembly with competitor DNAs that contained the wild-type or mutant versions of the Adh promoter. The wild-type and *Inr DNAs, which contained transcriptionally active promoters, were more effective competitors than the *TATA and *TATA*Inr DNAs, which contained mutant versions of the Adh promoter that were transcriptionally inert. Thus, these results indicate that the purified, recombinant basal factors could form DB, DBPolF, and DBPolFE complexes on the Drosophila Adh promoter in a manner that was consistent with the properties of native, mammalian factors with the AdML promoter.

We then carried out both gel-mobility shift and in vitro transcription analyses with minimal Adh promoter variants containing the wild-type sequence from -40 to +25 relative to the start site [Adh], mutations in the TATA box, mutations in the initiator element, or mutations in both the TATA and initiator elements. The primary objective of these experiments was to compare the ability to form DB and DBPolF30 complexes with the transcriptional competence of the promoters. The gel-shift analysis of these promoters with purified dTFIIB, dTBP, hTFIIF30, and Drosophila RNA polymerase II showed that DB and DBPolF30 complexes assembled...
Transcription by RNA polymerase II

with comparable efficiency on the wild-type and *Inr promoters, whereas there was no detectable formation of DB and DBPolF30 complexes with the *TATA and *TATA*Inr promoters [Fig. 6A]. The complementary in vitro transcription analysis of these promoters revealed that the wild-type and *Inr promoters were transcriptionally competent, whereas the *TATA and *TATA*Inr promoters were transcriptionally inert [Fig. 6B]. Furthermore, the discrimination between active and inactive promoter variants by the mixture of transcription factors present in a crude nuclear extract was identical to that displayed by the purified factors [Fig. 6B].

Thus, the efficiency of assembly of the DB and DBPolF30 complexes with wild-type and mutant Adh promoters correlated well with the ability of the promoters to be transcribed. Moreover, the transcriptional and gel-shift activities of the promoter variants were strongly dependent on the presence of the TATA box and, therefore, with the Adh promoter, the initiator element did not appear to be required for assembly of the minimal transcription complexes. We do not suggest, however, that the Inr element is generally dispensable for transcription by the purified transcription factors. In the specific case of the Adh promoter, it appears that the TATA box is the dominant promoter element, whereas, in contrast, with the Drosophila Krippel promoter, it appears that both the TATA box and the initiator elements are required for transcription by the purified factors [C.P. George and J.T. Kadonaga, unpubl.]. These data, along with the previous observations, suggest that TFIIIB, TBP, TFIIIF30, and RNA polymerase II can assemble into a competent DBPolF30 complex with some, but not all, promoters and that the polymerase can then initiate transcription accurately from this minitranscription complex upon the addition of ribonucleoside 5′-phosphates.

Summary and perspectives

We have been able to reconstitute accurate transcription, as well as the assembly of DBPolF30 complexes, with purified, recombinant TFIIIB, TBP, and TFIIIF30 and purified RNA polymerase II. This minimal set of factors was able to transcribe a subset of class II promoters. Studies of the mechanism of basal transcription have suggested that TFIIIB, TFIIID, TFIIF, and polymerase are the first four factors to assemble into the transcription complex, but accurate transcription by these factors had not been demonstrated previously. In this work, we used four different preparations of RNA polymerase II that were purified by different methods from Drosophila embryos, calf thymus, or HeLa cells. Straightforward interpretation of these data suggests that the minimal set of factors, TFIIIB, TBP (or TFIIID), TFIIF, and polymerase, con-

![Figure 6. Analysis of transcription factor–template DNA complex formation with wild-type or mutant Adh minimal promoter fragments. Four versions of the minimal Drosophila Adh proximal promoter were used. pAdh -40/+ 25 contains the wild-type Adh sequence from -40 to +25 relative to the RNA start site; pAdh *TATA contains a mutant TATA box; pAdh *Inr contains a mutant initiator element; and pAdh *TATA*Inr contains mutant TATA and initiator elements. [A] Gel mobility-shift analysis of transcription factor–template DNA complex formation. The assembly of TBP–TFIIIB (DB) and TBP–TFIIIB–RNA polymerase II–TFIIIF30 (DBPolF30) complexes was examined with the different promoter fragments under conditions identical to those described in Fig. 5. The positions of the free DNA probe, DB complex, and DBPolF30 complex are indicated. [B] Transcription of the wild-type and mutant Adh minimal promoters. In vitro transcription reactions were carried out with either a crude, soluble nuclear fraction from Drosophila embryos (SNF; 60 μg of total protein) [Kamakaka et al. 1991] or purified transcription factors under conditions identical to those described in Fig. 3A. The reverse transcription products are indicated by brackets.]

Cold Spring Harbor Laboratory Press
Tyree et al.

constitutes a minitranscription complex that represents the central core of the basal transcriptional machinery.

We do not yet understand the molecular basis for the ability of the purified proteins to transcribe only a fraction of the genes that we have tested. All of the promoters used in this study could be transcribed in vitro with either crude or partially purified preparations of transcription factors; thus, there may be one or more components, such as TFIIH, TFIIJ, or the TATA box-binding protein-associated factors (TAFs) [for review, see Pugh and Tjian 1992], that are required for transcription of the remaining genes. We also examined the DNA sequences of each of the promoters but did not find any common sequence motifs that correlated with either the ability or the inability of the genes to be transcribed by the purified factors. In the future, it will be important to elucidate both the protein factors and the DNA sequence elements that are responsible for the variations in the transcriptional properties of the different genes.

In vivo, the natural mechanism of transcription is likely to involve many other factors. Some of those factors may participate directly in the transcription process, whereas others may function to counteract repression by transcriptional inhibitors, which might include components of chromatin. We do not presume that transcription in vivo requires only the minimal set of factors described in this work but suggest that this set of factors might function as the central, catalytic core of the transcriptional machinery. In this regard, the minimal transcriptional complex will provide a starting point for further studies that should lead to a more complete understanding of the mechanisms of basal and regulated transcription in eukaryotes.

Materials and methods

Nomenclature of basal transcription factors

Transcription factors IIA through IJJ are abbreviated as TFIIA, TFIIIB, and so on. The TATA box-binding polypeptide of the TFIIID complex is referred to as TBP. All of the purified basal factors used in this study were synthesized in E. coli. A lower case d or h designates the original source of the factor as either Drosophila or human. Hence, dTFII is recombinant, Drosophila transcription factor IIB. The 34- and 56-kD subunits of human TFIIIE are referred to as hTFIIIE34 (or hE34) and hTFIIIE56 (or hE56). The 30- and 46-kD subunits of human TFIIIE are designated as hTFIIIE30 (or hE30) and hTFIIIE34 (or hE34). Alternate nomenclature for the basal transcription factors is as follows [for review, see Zawel and Reinberg 1992]: TFIIA (α, FA, factor α); TFIID (τ, factor d, BTF-1); TFIIIE (ε, factor α); TFIIIF (RAP30/74, βγ, factor 5, FC, BTF4, factor g); TFIIH [BTF-2, factor b, 8]; TFII IS (S-II).

Purification of the basal transcription factors

All of the recombinant basal transcription factors were synthesized in E. coli with T7 RNA polymerase expression vectors [pAR and PET vectors; Studier et al. 1990]. The cDNAs encoding the transcription factors were inserted into the expression vectors in a manner that allowed the initiating methionine codons of the cDNAs to be used by the E. coli translational machinery. Hence, the resulting proteins did not contain extraneous amino acid residues derived from the expression vector. dTFIIIB was purified as described by Wampler and Kadonaga (1992); hTFIIIB [plasmid pBlI (Ha et al. 1991), generously provided by I. Ha and D. Reinberg], was purified by the method of Ha et al. (1991). dTFIIA [plasmid pTFIIA35 (Hoey et al. 1990)], generously provided by T. Hoey, B. Dynlacht, and R. Tjian, and hTFIIA [plasmid pGGP63 (Peterson et al. 1990), generously provided by G. Peterson, T. Hoey, and R. Tjian], were purified by successive chromatography on DEAE–Sepharose, heparin–Sepharose CL-2B, and Mono Q resins. TBP was assayed by its activity in a reconstituted in vitro transcription reaction [Wampler et al. 1990], and only a small fraction (<5%) of the TBP, which appeared to be >95% homogeneous by SDS–polyacrylamide gel analysis, was active for basal transcription. hTFIIIE34 [plasmid pAR34 (Peterson et al. 1991), generously provided by M. Maxon, G. Peterson, R. Tjian, J. Inostroza, and D. Reinberg], was purified by successive chromatography on S-Sepharose and Mono Q resins. hTFIIIE56 [plasmid pAR56 (Peterson et al. 1991), generously provided by M. Maxon, G. Peterson, R. Tjian, J. Inostroza, and D. Reinberg], was purified by successive chromatography on DEAE–Sepharose, Q-Sepharose, Mono Q, Superdex 200, and Mono Q resins followed by precipitation with 1.4 M ammonium sulfate. hTFIIE30 [plasmid pETRAP30-11d (Sopha et al. 1989; Finkstein et al. 1992), generously provided by Z. Burton and J. Greenblatt], was purified by the method of Finkelstein et al. (1992). hTFIIE34 [plasmid pETRAP74-3d (Aso et al. 1992), generously provided by T. Aso and S. Weissman], was purified to ~30% homogeneity by successive chromatography on P11 phosphocellulose [Whatman], Superdex 200, and Mono Q resins. The preparations of hTFIIIE34 contained two contaminating polypeptides that appeared to be fragments of hTFIIE34. All of the chromatography resins, with the exception of P11 phosphocellulose [Whatman] and heparin–Sepharose CL-2B, can be obtained from Pharmacia-LKB. Detailed procedures for the purification of the E. coli-synthesized basal transcription factors are available upon request. RNA polymerase II from Drosophila embryos was purified as described by Wampler et al. (1990), except that an additional 15–50% (vol/vol) glycerol gradient step was included after Mono Q chromatography. The glycerol gradient purification yielded a preparation of RNA polymerase II from Drosophila embryos that was lacking the 46-kD polypeptide that had been observed previously in preparations of polymerase obtained from embryos [Greenleaf et al. 1979; Wampler et al. 1990, but not larvae [Greenleaf and Bautz 1975]. Calf thymus RNA polymerase IIA was purified by the method of Kim and Dahmus [1988] through the heparin–Sepharose CL-4B, DE-52 [Whatman], and DEAE–SPW [TosoHaas] columns [Kim and Dahmus 1988] and then purified further by alkyl–Superose (Pharmacia-LKB) chromatography. Calf thymus RNA polymerase IIB was purified according to the method of Hodo and Blatti [1977] through the DE-52 [Whatman] and P11 phosphocellulose [Whatman] columns followed by chromatography on heparin–Sepharose CL-4B and Mono Q (Pharmacia-LKB) resins as described by Kim and Dahmus [1988]. RNA polymerase II from HeLa cells [generous gift of D. Reinberg, Robert Wood Johnson Medical School, Piscataway, NJ] was purified as described by Lu et al. (1991).

In vitro transcription analyses

Transcription reactions were carried out with the indicated amounts of each protein as described by Wampler et al. (1990), except that the transcription factors and the supercoiled DNA template [100 ng] were preincubated at 21°C for 30 min [to allow assembly of the preinitiation complexes] before initiation of transcription by the addition of ribonucleoside 5’ triphosphates.
Transcription by RNA polymerase II

The transcription reactions were then incubated at 21°C for 30 min. Reactions with mammalian RNA polymerases were performed at slightly higher temperatures (either 25°C or 30°C, as noted in the figure legends). Where indicated, α-amanitin was included in the reactions at a concentration of 4 μg/ml, which is sufficient to inhibit transcription by RNA polymerase II but not that of RNA polymerases I and III. The resulting transcripts were subjected to primer extension analysis as described previously (Kadonaga 1990; Wampler et al. 1990). Quantitation of the in vitro-synthesized RNA was carried out with a PhosphorImager (Molecular Dynamics). All experiments were performed a minimum of two times (but typically, at least four times) to ensure reproducibility of the data. The plasmid pPA386-56 contains sequences from -426 to +383 (relative to the transcription start site) of the Drosophila proximal Adh promoter with a deletion between positions -868 and -56 (Heberlein et al. 1985). Transcripts synthesized from pPA386-56 were analyzed with the Adh proximal primer (5'-TGACAATCTTAGCGCAGAAGTCATGCG). The plasmid pAdh-40/+25 was constructed by insertion of complementary, synthetic oligonucleotides containing sequences from -40 to +25 (relative to the RNA start site) of the Drosophila Adh proximal promoter region into the XbaI and PstI restriction sites in the polylinker of pUC119. The plasmid pAdh*TATA is identical to pAdh-40/+25, except that the TATA box sequence from -32 to -25 was changed from TATAAAA to ACCTCGCCG. The plasmid pAdh*TATA*Inr is identical to pAdh-40/+25, except that the initiator element from +1 to +6 was changed from AACAC to GTTGGT. The plasmid pAdh*TATA*Inr1 is identical to pAdh-40/+25, except that it contains both sets of TATA box and initiator mutations present in pAdh*TATA and pAdh*Inr. The plasmid pKr-31/+13(119) contains sequences from -31 to +13 (relative to the major upstream RNA start site) of the Drosophila Krüppel promoter in pUC119 (Kerrigan et al. 1991). Transcripts synthesized from pAdh-40/+25, pAdh*TATA, pAdh*Inr, pAdh*TATA*Inr, and pKr-31/+13(119) were analyzed with the M13 reverse sequencing primer (5'-AGCGGATAACAATTTCACACAGGA), which hybridizes to pUC sequences. The plasmid pLAX contains sequences from -259 to +535 (relative to the transcription start site) of the AdML promoter from -40 to +25 relative to the RNA start site. This fragment was gel purified and 3′ end-labeled by filling in the restriction enzyme overhangs with the large fragment of E. coli DNA polymerase I, dGTP, dCTP, dTTP, and [α-32P]dATP to give a 114-bp fragment. The labeled probe (−0.1 to −0.2 ng; 10,000 cpm) was incubated with the transcription factors (as indicated in the figure legends) in a medium containing 25 mM HEPES (K+); pH 7.6, 7.5 mM MgCl2, 60 mM KCl, 2.5 mM (NH4)2SO4, 0.06 mM EDTA, 8% [vol/vol] glycerol, 2% [wt/vol] polyethylene glycol 8000, and 50 ng of poly[d(G-C)] in a total volume of 10 μl. The binding/assembly reactions were carried out at 25°C for 30 min, and the resulting samples were subjected to electro- phoresis on a 4% (wt/vol) polyacrylamide gel. All experiments were performed a minimum of two times (but typically, at least four times) to ensure reproducibility of the data.

Acknowledgments

We thank D. Reinberg, R. Tjian, Z. Burton, J. Greenblatt, T. Aso, and S. Weissman for the generous sharing of reagents and information that were essential for this study. We are grateful to E.P. Geiduschek, D. Reinberg, E. Blackwood, M. Levine, R. Kamakaka, M. Pazin, S. Paranjape, and G. Croston for critical reading of the manuscript. J.T.K. is a Lucille P. Markey Scholar in the Biomedical Sciences and a Presidential Faculty Fellow. This work was supported by a grant from the National Institutes of Health [NIH (GM 33300) to M.E.D. and grants from the NIH, the National Science Foundation, and the Lucille P. Markey Charitable Trust to J.T.K.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 USC section 1734 solely to indicate this fact.

References

Adams, C.C. and J.L. Workman. 1993. Nucleosome displacement in transcription. Cell 72: 1–20.

Aso, T., H.A. Vasavada, T. Kawaguchi, F.I. Germino, S. Ganguy, S. Kitajima, S.M. Weissman, and Y. Yasukochi. 1992. Characterization of the cDNA for the large subunit of the transcription initiation factor TFIIF. Nature 355: 461–464.

Buratowski, S. and P.A. Sharp. 1990. Transcription initiation complexes and upstream activation with RNA polymerase II lacking the C-terminal domain of the largest subunit. Mol. Cell. Biol. 10: 5562–5564.

Buratowski, S., S. Hahn, L. Guarente, and P.A. Sharp. 1989. Five intermediate complexes in transcription initiation by RNA polymerase II. Cell 56: 549–561.

Conaway, J.W. and R.C. Conaway. 1991. Initiation of eukaryotic messenger RNA synthesis. J. Biol. Chem. 266: 17721–17724.

Cortes, P., O. Flores, and D. Reinberg. 1992. Factors involved in specific transcription by mammalian RNA polymerase II: Purification and analysis of transcription factor II A and identification of transcription factor III. Mol. Cell. Biol. 12: 413–421.

Croston, G.E. and J.T. Kadonaga. 1993. Role of chromatin structure in the regulation of transcription by RNA polymerase II. Curr. Opin. Cell Biol. 5 [in press].

Dignam, J.D., R.M. Lebovitz, and R.G. Roeder. 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res. 11: 1475–1489.

Drapkin, R., A. Merino, and D. Reinberg. 1993. Advances in regulation of transcription by RNA polymerase II. Curr. Opin. Cell Biol. 5 [in press].

Dynan, W.S. and R. Tjian. 1983. Isolation of transcription factors that discriminate between different promoters recognized by RNA polymerase II. Cell 32: 669–680.

Felsenfeld, G. 1992. Chromatin as an essential part of the transcriptional mechanism. Nature 355: 219–224.

Finkelstein, A., C.F. Kostrub, J. Li, D.P. Chavez, B.Q. Wang, S.M. Fang, J. Greenblatt, and Z.F. Burton. 1992. A cDNA encoding RAP74, a general initiation factor for RNA polymerase II. Nature 355: 464–467.

Fischer, L., M. Gerard, C. Chalut, Y. Lutz, S. Humbert, M.
Tyree et al.

Kanno, P. Chamblon, and J.-M. Egly. 1992. Cloning of the 62-kilodalton component of basic transcription factor BTF2. Science 257: 1392–1395.

Flores, O., H. Lu, and D. Reinberg. 1992. Factors involved in specific transcription by mammalian RNA polymerase II: Identification and characterization of factor IIH. J. Biol. Chem. 267: 2786–2793.

Gileadi, O., W.J. Feaver, and R.D. Kornberg. 1992. Cloning of a subunit of yeast RNA polymerase II transcription factor b and CTD kinase. Science 257: 1389–1392.

Greenblatt, J. 1991. RNA polymerase-associated transcription factors. Trends Biochem. Sci. 16: 408–411.

Greenleaf, A.L., L.M. Borsett, P.F. Jiamachello, and D.E. Coulter. 1979. α-Amanitin-resistant D. melanogaster with an altered RNA polymerase II. Cell 18: 613–622.

Grunstein, M. 1990. Histone function in transcription. Annu. Rev. Cell Biol. 6: 643–678.

Hoey, T., B.D. Dynlacht, M.G. Peterson, B.F. Pugh, and R. Tjian. 1992. Diverse transcriptional functions of human general transcription factor IIE. Nature 354: 369–373.

Kamakaka, R.T., C.M. Tyree, and J.T. Kadonaga. 1991. Accurate transcription by mammalian RNA polymerase II: Role of transcription factors IIA, IIB, and IIB during formation of a transcription-competent complex. Mol. Cell. Biol. 10: 6335–6347.

Manley, J.L., A. Fire, A. Cano, P.A. Sharp, and M.L. Gefer. 1980. DNA-dependent transcription of adenovirus genes in a soluble whole-cell extract. Proc. Natl. Acad. Sci. 77: 3855–3859.

Mitchell, P.J. and R. Tjian. 1989. Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. Science 245: 371–378.

Ohkuma, Y., H. Sumimoto, A. Hoffman, S. Shimasaki, M. Hori-koshi, and R.G. Roeder. 1991. Structural motifs and potential α homologies in the large subunit of human general transcription factor TFIIIE. Nature 354: 398–401.

Parvin, J.D., H.Th.M. Timmers, and P.A. Sharp. 1992. Promoter specificity of basal transcription factors. Cell 68: 1135–1144.

Peterson, M.G. and R. Tjian. 1992. The tell-tale trigger. Nature 358: 620–621.

Peterson, M.G., N. Tanase, B.F. Pugh, and R. Tjian. 1990. Functional domains and upstream activation properties of cloned human TATA binding protein. Science 248: 1625–1630.

Peterson, M.G., J. Inostroza, M.E. Maxon, O. Flores, A. Admon, D. Reinberg, and R. Tjian. 1991. Structure and functional properties of human general transcription factor IIE. Nature 354: 369–373.

Ptashne, M. 1988. How eukaryotic transcriptional activators work. Nature 335: 683–689.

Pugh, B.F. and R. Tjian. 1992. Diverse transcriptional functions of the multisubunit eukaryotic TFIIID complex. J. Biol. Chem. 267: 679–682.

Ranish, J.A., W.S. Lane, and S. Hahn. 1992. Isolation of two genes that encode subunits of the yeast transcription factor IIA. Science 255: 1127–1129.

Roeder, R.G. 1991. The complexities of eukaryotic transcription initiation: Regulation of preinitiation complex assembly. Trends Biochem. Sci. 16: 402–408.

Saltzman, A.G. and R. Weinmann. 1989. Promoter specificity and modulation of RNA polymerase II transcription. FASEB J. 3: 1723–1733.

Sawadogo, M. and R.G. Roeder. 1985. Factors involved in specific transcription by human RNA polymerase II: Analysis by a rapid and quantitative in vitro assay. Proc. Natl. Acad. Sci. 82: 4394–4398.

Sawadogo, M. and A. Sentenac. 1990. RNA polymerase B [II] and general transcription factors. Annu. Rev. Biochem. 59: 711–754.

Sayre, M.H., H. Tschochner, and R.D. Kornberg. 1992. Reconstitution of transcription with five purified initiation factors and RNA polymerase II from Saccharomyces cerevisiae. J. Biol. Chem. 267: 23376–23382.

Schaeffer, L., R. Roy, S. Humbert, V. Moncollin, W. Vermeulen, J.H.J. Hoeijmakers, P. Chambon, and J.-M. Egly. 1993. DNA repair helicase: A component of BTF2 (TFIIH) basic transcription factor. Nature 360: 398–401.

Wampler, S.L. and J.T. Kadonaga. 1992. Functional analysis of Drosophila transcription factor IIB. Genes & Dev. 6: 1542–
Wampler, S.L., C.M. Tyree, and J.T. Kadonaga. 1990. Fractionation of the general RNA polymerase II transcription factors from Drosophila embryos. J. Biol. Chem. 265: 21223–21231.

Weil, P.A., D.S. Luse, J. Segall, and R.G. Roeder. 1979. Selective and accurate initiation of transcription at the Ad2 major late promoter in a soluble system dependent on purified RNA polymerase II and DNA. Cell 18: 469–484.

Weis, L. and D. Reinberg. 1992. Transcription by RNA polymerase II: Initiator-directed formation of transcription-competent complexes. FASEB J. 6: 3300–3309.

Wolffe, A.P. 1992. Chromatin: Structure and function. Academic Press, San Diego, CA.

Zawel, L. and D. Reinberg. 1992. Advances in RNA polymerase II transcription. Curr. Opin. Cell Biol. 4: 488–495.

——., 1993. Initiation of transcription by RNA polymerase II: A multi-step process. Prog. Nucleic Acids Res. Mol. Biol. 44: 67–108.

Zehring, W.A., J.M. Lee, J.R. Weeks, R.S. Jokerst, and A.L. Greenleaf. 1988. The C-terminal repeat domain of RNA polymerase II largest subunit is essential in vivo but is not required for accurate transcription initiation in vitro. Proc. Natl. Acad. Sci. 85: 3698–3702.
Identification of a minimal set of proteins that is sufficient for accurate initiation of transcription by RNA polymerase II.

C M Tyree, C P George, L M Lira-DeVito, et al.

Genes Dev. 1993, 7:
Access the most recent version at doi:10.1101/gad.7.7a.1254