Summary

A native gel electrophoresis DNA binding assay was used to resolve complexes formed on the adenovirus Major Late Promoter by general transcription factors and RNA polymerase II. Five sets of complexes containing distinct components were identified. These complexes were generated by sequential binding of TFIID, TFIIA, TFIIB, RNA polymerase II, and TFIIE. The relative positions of each of the factors in the complexes were determined by DNAase I footprint analysis. TFIID, derived from yeast or mammalian cells, formed a complex with yeast TFIID and the TATA element. TFIIA bound to this complex and probably acts as a "bridge" to the polymerase and the initiation site. The addition of ATP or dATP, necessary for "activation" of transcription, resulted in an alteration of the footprint in the +20 to +30 region, the same area protected upon addition of TFIIE to the initiation complex. Addition of ribonucleotide triphosphates generated new complexes that contained accurately initiated transcripts associated with the transcription machinery and the template DNA. A model for the interactions of components in initiation of transcription by RNA polymerase II is proposed.

Introduction

The rate of or initiation of transcription is determined by factors that recognize sequences in enhancer elements, sites upstream of the TATA element, and the TATA element. The upstream binding factors probably influence events at the TATA element; promoting either formation of a complex or initiation by the polymerase. Analysis of the regulation of transcription has been limited by a lack of understanding of the TATA-directed initiation process. Recent results suggest that some aspects of the mechanisms by which these stimulatory factors act are conserved between yeast and higher eukaryotes (see Struhl, 1987; Guarente, 1988, for review). Some aspects of the basic initiation reaction are also conserved, as a TATA binding protein from yeast will direct initiation by mammalian factors and polymerase (Buratowski et al., 1988; Cavallini et al., 1988). This conservation in mechanism is also reflected in the conserved structure of RNA polymerase II (pol II) over evolution (Allison et al., 1985; Corden et al., 1988).

Purified pol II will not accurately initiate transcription in vitro. Accurate initiation can be observed in whole cell (Weil et al., 1979; Manley et al., 1980) or nuclear extracts (Dignam et al., 1983a), and can be reconstituted with partially purified fractions. In addition to pol II, four activities have been shown to be required. These general transcription factors have been partially purified and designated TFIIA, -B, -D, and -E (Matsui et al., 1980; Samuels et al., 1982; Dignam et al., 1983b; Sawadogo and Roeder, 1984). This combination of transcription factors will accurately initiate from a minimal promoter containing only a TATA element and start site.

The factor TFIID (also known as DB or TFIH) contains a protein that specifically binds to the TATA element (Sawadogo and Roeder, 1985b; Nakajima et al., 1986). The first step in the initiation of transcription is correlated with this binding (Davison et al., 1983; Fire et al., 1984). Recently, it has been shown that TFIID from the yeast Saccharomyces cerevisiae is functionally interchangeable in vitro with mammalian TFIID (Buratowski et al., 1988; Cavallini et al., 1988). The two proteins may differ in some physical characteristics, as the purified yeast TFIID behaves as a single protein of approximately 25 kilodaltons (Buratowski et al., 1988), while the mammalian factor behaves as a much larger entity (Samuels et al., 1982; Reinberg et al., 1987).

The factor TFIIA (AB, STF) is also required for the efficient interaction of TFIID and the TATA element (Davison et al., 1983; Fire et al., 1984), although its mechanism of action is unknown. TFIIA activity has been purified from calf thymus and HeLa cells as a set of proteins of 19 and 13 kd (Samuels and Sharp, 1986) and as a single protein of 43 kd (Cgly et al., 1986), respectively. TFIIA may exist as a dimer, or these proteins may represent proteolytic products, as native size estimates of crude preparations suggest a molecular weight of 34 kd for TFIIA from calf thymus (Samuels and Sharp, 1986) and 62 kd for TFIIA from HeLa cells (Reinberg et al., 1987).

The factor TFII B (CBII, BTF3) has been purified from HeLa cells as a protein of 27-30 kd (Reinberg and Roeder, 1987a; Zheng et al., 1987). Although its specific function is unknown, it binds to pol II and possibly TFII E in solution and presumably promotes transcription through this interaction (Zheng et al., 1987; Reinberg and Roeder, 1987a).

The factor TFII E (CBII, BTF2) also binds to pol II in solution (Reinberg and Roeder, 1987a). The only hint to its function in transcription is that fractions containing partially purified TFII E also contain a DNA-dependent ATPase (Sawadogu and Roeder, 1984, Reinberg and Roeder, 1987a). Gel filtration and sedimentation analysis suggest a molecular weight of 76 kd for the TFII E activity (Reinberg and Roeder, 1987a). TFII E is probably related to the...
sis, a method that combines high resolution and sensitivity. Preparative gels have not been resolved by native gel electrophoresis. To date, initiation complexes that are resistant to challenges by inhibitors and template commitment assays, as well as the higher resolution assay of DNAase I footprinting. The complexes react to inhibitor challenges and nucleotide triphosphates in a manner consistent with previous studies of transcription initiation complexes. The hierarchical nature of the complexes suggests a model for transcription initiation and predicts functions for the general factors.

**Results**

TFIIA Affects Binding of TFII to the TATA Element

RNA pol II transcription can be inhibited by poly(dI-dC) or Sarkosyl. Preincubation of the transcription template with TFIIA and TFII before addition of inhibitor renders transcription resistant to such a challenge (Fire et al., 1984; Reinberg et al., 1987). This and other assays have been used to define the first stable complex in initiation: a preinitiation complex "committed" to transcription of a particular template, presumably involving specific binding of TFIIA to the TATA element. Although it clearly acts at this step, there has been a variable requirement for TFIIA. At high levels of yeast TFIIA, TFIIA was not required for poly(dI-dC)-resistant complex formation, but at lower TFIIA levels, TFIIA significantly increased complex formation (Buratowski et al., 1988). This suggests that TFIIA stimulates binding of TFIIA to the TATA element. To examine this issue further, increasing amounts of yeast TFIIA were tested for DNAase I protection of the adenovirus Major Late Promoter (MLP) in the presence or absence of TFIIA. The TFIIA was purified to near homogeneity from yeast whole cell extracts (Buratowski et al., 1988). TFIIA (the generous gift of M. Samuels) was purified to near homogeneity from calf thymus (Samuels and Sharp, 1986). Both factor preparations were purified on the basis of their ability to substitute for the corresponding HeLa factor, and will efficiently function together in an in vitro transcription reaction containing pol II from calf thymus and TFIIA and TFIIIE from HeLa cells.

The presence of TFIIA, which demonstrated no protection by itself, induced two changes in the DNAase I footprint of TFIIA (Figure 1). The first change was quantitative: approximately two-fold less TFIIA was required to saturate binding of the MLP TATA element in the presence of TFIIA. Furthermore, another TATA sequence (TATAAA), at -70 to -75 on the opposite strand, bound TFIID only in the presence of TFIIA, even at the highest levels of TFIIA tested. This upstream footprint is clearly different from footprints of the upstream activators MLTF (Carruthers et al., 1985; Sawadogo and Roeder, 1985b; Moncolin et al., 1988) or CP1 (Chodosh et al., 1988a). The second change seen was qualitative. In the absence of TFIIA, TFIIA protected the coding strand from -17 to -39. In the presence of TFIIA, the protection was extended to approximately -42, and the frequency of cleavage at the upstream boundary increased three-fold (denoted by an arrow in Figure 1). This enhanced cleavage is not due to interactions be-
2) are sequencing ladders of the probe fragment. 

TFIIA Complexes with TFIIID and the TATA Element

Several mechanisms can be postulated for the effect of TFIIA on TFIIID binding. TFIIA could modify TFIIID and/or the promoter DNA to allow a higher affinity interaction. Alternatively, TFIIA could be a component of the preinitiation complex. To explore these possibilities, the gel shift assay was tested for resolution of protein–DNA complexes containing TFIIID (Figure 2A). Incubation of TFIIA alone did not yield a complex with the MLP (lane 1), in agreement with previous studies (Egly et al., 1984; Samuels and Sharp, 1986; Reinberg et al., 1987). Addition of TFIIID alone yielded a fast migrating protein–DNA complex (lane 2). However, this complex was due to non-specific binding of a protein, as it was competed by addition of unrelated fragments and formed on fragments not containing a TATA sequence (data not shown). It is thought that this complex is due to a contaminating DNA binding protein, as does not chromatograph exactly with the TFIIID transcription activity (data not shown). When TFIIID and TFIIA were incubated together with the MLP probe DNA, a new set of complexes was formed (lane 3). These complexes did not form on a probe from a fragment containing a double point-mutation (TAGAGAA) in the TATA element (data not shown). For footprint analysis, the binding mixture was treated with DNAase I before loading on the native gel, and DNA in the complexes was recovered (Figure 2B). The free probe (F) and the non-specific complex (NS) yielded the same DNAase I pattern as DNA in solution (lanes 3 and 4). The specific complexes, dependent on TFIIA and TFIIID, showed protection of the TATA element (lane 5). Furthermore, this footprint had the 5' extension and the enhanced cleavage at the upstream site observed previously only when both TFIIID and TFIIA were added to a footprint assay in solution. Since the specific complexes always behaved identically and probably resulted from heterogeneity in the purified TFIIA (see below), they will be referred to as a single complex.

The gel shift assay suggested that TFIIA was required to form a stable and specific complex between TFIIID and the TATA element, but it did not address whether TFIIA was part of the complex. Knowing that different sources of TFIIA yield factors of different apparent molecular weights (Egly et al., 1984; Samuels and Sharp, 1986; Reinberg et al., 1987) various samples of TFIIA were tested with and without TFIIID in the gel shift assay (Figure 3). HeLa fraction [AB] (a second column preparation of human TFIIA; Samuels et al., 1982) and a TFIIA-like activity partially purified from yeast (unpublished data) both failed to form specific complexes in the absence of TFIIID (lanes 3 and 4, respectively). However, in the presence of TFIIID, a single new complex was formed in each case (lanes 6 and 7). These complexes showed the same DNAase I protection and specificity as the calf thymus TFIIA-dependent complexes (data not shown). The mobilities of the complexes formed with each of the TFIIA sources was different. Since the different TFIIA sources are functionally interchangeable in a transcription reaction in vitro, and produce the same changes in the TFIIID footprint, the different mobilities of their complexes with TFIIID and the MLP suggest that TFIIA is a component of the complex. The different properties of TFIIA from various sources might reflect inherent differences between species and/or partial proteolytic cleavages.
Figure 3. Different Sources of TFIIA Generate Different Mobility Complexes with TFIID on the MLP, Suggesting That TFIIA Is a Component of the Complex

Probes (from plasmid pRW, containing MLP sequences from -53 to +33) were incubated with 2 µl of yeast TFIID (lane 1), 2 µl of TFIIA purified from calf thymus (CT II A, lane 2), 0.5 µl of HeLa fraction [AB], a partially purified preparation of human TFIIA (lane 3), or 1.0 µl of a partially purified yeast activity that substitutes for TFIIA (Yeast II A, lane 4). No specific complexes were formed by any of the factors alone. Each of the TFIIA sources was then tested in the presence of 2 µl of TFIID (lanes 5-7). In each case, a new complex was formed. Although the three complexes were of different mobilities, all three showed the same DNAase I protection pattern over the TATA element as described in Figure 2. The TFIIA + TFIID complex in lane 5 is more clearly visible upon darker exposure, and is identical to that shown in Figure 2A, lane 3. A lighter exposure of the gel is shown here so that the different positions of the complexes in lanes 6 and 7 can be distinguished.

Higher Order Complexes

Identification of a specific complex containing TFIIA and TFIID by the gel shift assay raised the possibility that addition of the other general transcription factors and Pol II might generate other complexes amenable to similar analysis. This was tested by incubating the MLP probe with purified TFIIID, TFIIA, and Pol II, as well as partially purified TFIIIB and -E (Figure 4). Individually (lanes 1-5), none of these fractions yielded complexes except for the nonspecific complex formed in the TFIIID reaction (lane 1). When all the fractions were incubated together (lane 6), in addition to the TFIIIA + TFIID complex (complex 2), five other complexes were formed (complexes 3-7). These complexes were shown to be specific in two ways. First, each was competed 3-to-5-fold more efficiently by a fragment containing MLP sequences spanning the TATA element (-42 to -17) than by an oligonucleotide that was identical, except for the mutations TATAAAA → TAGAGAA (data not shown). These same two fragments compete with similar relative efficiencies for the DNAase I protection footprint of TFIIID on the MLP TATA element (Buratowski et al., 1988). Second, labeled probes of the aforementioned fragments were tested in the gel shift assay. Complexes 2-7 formed on the wild-type probe, but not on the mutant probe (data not shown). This suggests that these complexes are specific for a TATA sequence, and that the only sequences required for their formation reside in or near the TATA element.

Individual fractions were omitted from the binding mixture to test the effect on complex formation. In the absence of TFIIID (Figure 4, lane 7), none of the specific complexes were formed. A background of nonspecific complexes was obtained that was due to contaminating binding proteins in the TFIIB and TFIIE fractions. Omission of TFIIA (lane 8) did not affect complexes 4-7. The doublets that make up complexes 2 and 3 were no longer seen, although a complex migrating close to complexes 2 and 3 was observed. This can be seen more clearly by comparing lanes 11 and 13, where less probe is shifted into higher order complexes, and there is less background due to the absence of nonspecific binding proteins contained in the TFIIE fraction. There are several possible ex-
RNA Polymerase II Initiation Complexes

A. 8. 2xIlD + IIA IID + IIA + IIB + pal It + 3 + DEAE Fractions + DEAE Gradient Fractions DEAE Gradient Fractions + 5 6 7 6 9 10 11 12 13 14 11 6 7 6 9 10 11 12 13 14 3 12 13 14

Figure 5. Separation of TFIIB and TFIIE by Gradient Elution Chromatography over DEAE-Sephacel

(A) Transcription assay of gradient fractions for TFIIB and TFIIE activity. Assays were performed as described in Experimental Procedures. Reactions received the following additions: none (−), 0.5 μl of HeLa fraction [CB] (+Load), or 1.0 μl of each of the DEAE-Sephacel gradient fractions. Fractions were tested for the presence of TFIIB and TFIIE together (−IFIB, −TFIIE), or for TFIIB alone (−IFIB) or TFIIE alone (−TFIIE).

(B) Gel shift analysis of DEAE-Sephacel gradient fractions in the presence of TFIID and TFIIE. All reactions contained 2.0 μl of yeast TFIID and probe containing MLP sequences from −53 to +33. One microliter of each of the gradient fractions was added and incubation and native gel electrophoresis were carried out as described in Experimental Procedures.

(C) Gel Shift analysis of DEAE-Sephacel gradient fractions in the presence of TFIID, TFIIA, TFIIB, and pol II. All reactions contained 2.0 μl of yeast TFIID, 2.0 μl of yeast TFIIE, 0.5 μl of yeast TFIIA, and 1.0 μl of DEAE-Sephacel gradient fraction 7 as a source of TFIIB. Reactions received either no addition (−) or 1.0 μl of each of the DEAE-Sephacel gradient fractions.

plansations for complex formation in the absence of added TFIIA. The simplest is the possibility that the TFIIB fractions were contaminated with TFIIA activity. Alternatively, it is possible that TFIID and TFIIB, and subsequently the other factors, can form stable complexes in the absence of TFIIA. Potentially, formation of these complexes could be stimulated by TFIIA in the transcription reaction. In either case, these results are consistent with previous experiments showing that an in vitro transcription reaction using these fractions is only partially dependent upon addition of TFIIA (Buratowski et al., 1988). Leaving out the crude fraction containing TFIIB allowed formation of complex 2, but not complexes 3-7 (lane 9). Omission of purified pol II resulted in formation of complexes 2 and 3, but not 4-7 (lane 10). All complexes except 6 and 7 were obtained when the fraction containing TFIIE was not added (lane 11). The omission of TFIIID or TFIIA in combination with the absence of TFIIE (lanes 12 and 13, respectively) had the same effects on complex formation as when TFIIE was present.

Since the TFIIB and TFIIE preparations were only partially purified relative to the highly purified TFIID, TFIIA, and pol II used in the previous analysis, further confirmation that formation of the complexes was dependent on the actual transcription factors, and not some other protein in those fractions, was sought. For this, formation of complexes was tested after more extensive fractionation of the transcription activities. Fraction [CR], containing both TFIIB and TFIIE, was chromatographed by salt gradient elution over a DEAE-Sephacel column. The transcription activities of TFIIB and TFIIE were assayed for either together (−TFIIB, −TFIIE) or separately (Figure 5A). TFIIB activity was found predominantly in fractions 7-10, peaking in fraction 8 (−TFIIA). There is also a small amount of TFIIB activity in fraction 6 (−TFIIE). The same fractions were assayed in the gel shift system by mixing with TFIIA and TFIID (Figure 5B). Complex 3 formed upon addition of fractions 7-10, with some in fraction 11, suggesting that complex 3 results from addition of TFIIB to complex 2 (TFIIA + TFIID). Addition of purified RNA pol II to a binding reaction containing TFIIA, -D, and -B (DEAE fraction 7) resulted in further formation of complexes 4 and 5 (Figure 5C, lane 1). Addition of only the DEAE gradient fractions 9-11 to this mixture generated complexes 6 and 7. These fractions contained the TFIIE transcription activity. It was not possible to test strict coelution of complex 3 formation with TFIIB activity, and complexes 6 and 7 formation with TFIIE activity, as the DNA binding assays were not necessarily linearly responsive, and high levels of background proteins may have interfered with the gel shift assay. The fact that complexes 6 and 7 are dependent on all the factors known to be required for in vitro transcription suggests that they represent complete initiation complexes. It is interesting to note that pol II-dependent complexes formed in pairs (complexes 4 and 5 and complexes 6 and 7). Possible reasons for this are discussed below.
The effect of various nucleotide triphosphates on MLP transcription initiation has been studied. The standard binding reactions were performed with the following protein components: 2.0 µl of calf thymus TFIIA, 2.0 µl of yeast TFIID, and 0.5 µl of HeLa fraction [CB] containing TFII B and TFII E (lane 1), or calf thymus pol II added to TFIIA and TFII D (lane 2), or TFIIA, TFIID, fraction [TFI], and pol II (lanes 3-5). In addition, binding reactions in lane 4 contained 100 µM dATP. Besides the indicated protein components, the binding reaction in lane 5 also contained 60 nM ATP, 60 µM UTP, 0.5 µM CTP, and 10 µM 3'-O-methyl GTP.

The first G residue of the MLP transcript is at position +11, addition of ATP, UTP, CTP, and 3'-O-methyl GTP (a chain terminator) yields transcription only to +11. Such a "paused" complex, assembled in crude extract (Cai and Luse, 1987a) or with partially purified factors (Sawadogo and Roeder, 1984), has been shown to be relatively stable. When assayed in the gel shift system, addition of these nucleotides to the reaction resulted in the formation of two new diffuse complexes: 8 and 9 (lane 5). Complex 8 migrated slightly faster than complex 4, and complex 9 migrated slightly slower than complex 5. When unlabeled MLP DNA was added in place of probe DNA to the transcription–binding reaction along with [α32P]CTP and ATP, UTP, and 3'-O-methyl GTP, label was incorporated in two diffuse complexes that migrated identically with complexes 8 and 9 (Figure 6B, lane 1). Such incorporation was not observed when a control DNA fragment of similar size was substituted for the MLP fragment (lane 7). Furthermore, labeled RNA recovered from these complexes migrated in a denaturing gel at the rates expected for accurately initiated MLP transcripts terminated at the first two G residues (data not shown).

The relative positions of the general transcription factors in the various complexes were determined by footprint analysis. For this, the complete binding mix was treated with DNAase I before loading onto the native gel. The DNA in each of the complexes was recovered and electrophoresed on a sequencing gel, and the patterns of protection on the coding and noncoding strands were shown in Figure 7A and 7B, respectively.
RNA Polymerase II Initiation Complexes

The protective patterns observed with complexes 3 and 5 are consistent with the formation of a DNA polymerase II complex, which typically exhibits a footprint covering a region spanning from -47 to +15 on the coding strand. In contrast, complexes 6 and 7, which are formed in the absence of TFIIIE, show a protection pattern similar to that of complex 2, with additional partial protections of some cleavage sites in the -10 to +10 region of the coding strand. This suggests that TFIIIE is likely bound in the complex downstream of the TATA element, perhaps loosely associated with regions of the coding strand. An asymmetric association would be very interesting, as it would leave the initiation site on the template (noncoding) strand accessible to the polymerase.

Complexes 4 and 5, which are generated when purified pol II is added to TFIIA, -IID, and -IIIB, showed identical footprints in all cases. The coding strand had the same upstream boundary (-42) as complexes 2 and 3, but downstream cleavages were strongly suppressed to about +15. On the coding strand, downstream protection was also extended to approximately +15, but surprisingly, the upstream boundary now was extended to about -47. The expanded protection relative to complex 3 presumably reflects the binding of the large RNA pol II molecule to the protein-DNA complex. This binding extends not only over the TATA element and the initiation site, but also two helical turns downstream of the initiation site.

Complexes 6 and 7 were only observed when TFIIIE was
Figure 8. Schematic Model of the Molecular Structure of the Complexes

The factor requirements and DNAase I footprint patterns of each of the complexes are reflected in this diagram. The arrow indicates the transcription initiation site, and the numbers at the top are nucleotide positions relative to that site. The asterisk on TFIIA in complexes 4–7 is to signify that it was not possible to determine whether TFIIA was present in these complexes. Complexes 6 and 7 are believed to represent complete initiation complexes.

The DNAase I protection patterns of complexes 6 and 7 were identical. The patterns were similar to those of complexes 4 and 5, but showed approximately 10 bp (one helical turn) extension of protection downstream. The coding strand was now protected to about +30, and the noncoding strand was protected to about +25 with some enhanced cleavage in sequences further downstream. These additional protections suggest that TFIIE binds downstream of the polymerase.

Addition of nucleotide triphosphates had no effect on the DNAase I footprints of complexes 2–5 (data not shown). The protection patterns of both complexes 6 and 7 were affected, and were still identical. Addition of ATP (data not shown) or dATP (complexes 6* and 7*) to the reaction caused not only a partial loss of the TFIIE-dependent protections between +20 and +30, but also a general suppression of all cleavage sites from +30 to the downstream end of the probe. As previously noted, complexes 6 and 7 apparently dissociate upon addition of ATP or dATP. These are the same nucleotide triphosphates that generate complexes 6* and 7*, which exhibit the downstream alterations in the footprint. Since the sets of complexes differ in the DNAase I pattern in an area of protection likely to be due to the TFIIE binding, it is theorized that an ATP-dependent release of TFIIE occurs during conversion of the complexes. This transition is probably distinct from the apparent mobility decrease observed upon conversion of complex 7 to 7*, because complex 6 undergoes the same footprint changes without an apparent mobility shift. In addition, some purine nucleotide triphosphates (such as GTP or ADP) will convert complex 7 to the slower migrating form without producing the changes in protection patterns (data not shown).

Addition of NTPs that allow transcription to +11 generated complexes 8 and 9. The protection patterns of complexes 8 and 9 were identical to each other and to those of complexes 4 and 5 (data not shown). However, since complexes 8 and 9 migrated with mobilities similar to complexes 4 and 5 respectively, it was difficult to prepare DNA confidently from the individual complexes. Therefore, the protection patterns of complexes 8 and 9 remain uncertain.

Discussion

We have resolved a series of complexes by native gel electrophoresis that suggest an ordered assembly of the general transcription factors and pol II on the Ad2 MLP template. The mobility of each complex, the factor requirements for its formation, and its DNAase I footprint have been combined to predict a molecular structure for each complex. A schematic representation of the composition and relative positions of the general transcription factors in each of these ordered complexes is shown in Figure 8. We propose that these complexes represent intermediates in the initiation reaction.

Complex 1

The first step in initiation is probably recognition of the TATA element by TFIID, perhaps in combination with TFIIA (see below). TFIID was required for formation of all of the specific complexes (2–7). TFIID will specifically protect the TATA element sequences −37 to −17 from DNAase I in the absence of any other transcription factors. A specific complex can be inferred from this result, and has been termed complex 1. For unknown reasons, no specific TFIID–TATA element complex was detectable in the gel shift assay. Complex 1 may be unstable in the gel shift assay, even though solution studies suggest a relatively long half-life (>20 min) for the TFIID–TATA element complex (unpublished data).

Published reports of DNAase I footprints on the MLP by partially purified mammalian TFIID are somewhat different from those observed with the yeast TFIID. The HeLa TFIID fractions were found to protect sequences from −38 to −4, followed by alternating hypersensitive sites (spaced about 10 bp apart) and protected areas to +35 (Sawadogo and Roeder, 1985b; Nakajima et al., 1988). While the protection in the region of the TATA element is similar to that seen with the purified yeast TFIID, the
downstream protections are strikingly different. Curiously, the downstream protections were observed on the MLP and the histone H4 promoter, but not several other promoters; these others showed only the TATA element protection (Nakajima et al., 1988). While the different TFIID footprints may reflect differences caused by proteolysis or conformation changes, another possible explanation is that the fractions containing mammalian TFIID (purified about 300-fold; Nakajima et al., 1988) also contain one or more other factors responsible for the downstream protections. Alternatively, it is possible that the yeast and mammalian factors are fundamentally different. These two factors have significantly different sizes when analyzed by sedimentation and chromatography (see Buratowski et al., 1988, for discussion).

Complex 2

Complex 2 contains both TFIID and TFIIA bound specifically to the -42 to -17 region of the MLP. The strongest evidence for the presence of TFIIA in the complex is that different sources of TFIID transcription activity (calf thymus, HeLa cells, or yeast) generated complexes of different mobilities, but identical DNAase I footprints, in conjunction with TFIID. The heteromeric complex 2 is reminiscent of some upstream transcription factors. The yeast transcription activator HAP 2/3 (Hahn and Guarente, 1988) and a family of human CCAAT binding proteins (Chodosh et al., 1988b) require at least two different polypeptides to reconstitute sequence-specific binding in a gel shift assay. In addition, the heteromeric contacts between the subunits of the yeast and mammalian activators are highly conserved, as subunits from the different species can be combined to regenerate specific binding (Chodosh et al., 1988b). Similarly, the TFIID-TFIIA interactions must be conserved. Both mammalian and yeast TFIID can combine with yeast TFIIA to generate TATA element-specific complexes, and any combination of yeast and/or mammalian TFIID and TFIIA can function together in a reconstituted in vitro transcription reaction (data not shown).

Addition of TFIIA slightly enhanced the binding of TFIID and modified the DNAase I protection pattern of TFIID on the TATA element. Interestingly, TFIIA more effectively enhanced the binding of TFIIA to a cryptic TATA sequence at -70 than to the MLP TATA element. Thus, TATA elements from different promoters may differ in their dependence upon TFIIA. Comparison of the footprints of complexes 1 and 2 demonstrates a TFIIA-dependent extension and a 3-fold enhancement in cleavage of a site at the upstream boundary of protection (about -43). Solely for this reason, TFIIA is positioned on the upstream side of TFIID in Figure 8.

Inhibition studies have previously defined a preinitiation complex whose formation required both TFIID and TFIIA, and which was resistant to challenges with either another TATA element, poly (dl-dC), or Sarkosyl (Davison et al., 1983; Fire et al., 1984; Reinberg et al., 1987). The complex formed from these solution studies probably corresponds to complex 2, as it has the same factor requirement and response to poly(dl-dC) challenge (data not shown). Studies of the role of TFIIA in initiation have been confounded by a variable requirement for the factor in in vitro transcription reactions; reports have ranged from TFIIA being totally dispensable (Sawadogo and Roeder, 1985a) to being strongly stimulatory (Egly et al., 1984; Samuels and Sharp, 1986; Buratowski et al., 1988) to being absolutely required (Reinberg et al., 1987).

Order of addition experiments have led to the suggestion that TFIIA acts before binding of TFIID, perhaps through a nonspecific interaction with the DNA (Reinberg et al., 1987). This would be surprising, as TFIIA does not bind to negatively charged columns. A potential complication in such experiments is the DNA-dependent inactivation of mammalian TFIID during preincubation in the absence of TFIIA (Fire et al., 1984). While the results presented here do not argue for or against the hypothesis that TFIIA acts before TFIID, they do demonstrate that TFIIA remains stably associated after binding of TFIID to the TATA element, at least during the early steps of initiation complex formation.

Complex 3

After binding to TFIID and TFIIA to the TATA element, TFIIB probably binds to the initiation complex. Fractions containing TFIIB are required for formation of complexes 3-9. The evidence that TFIIB is actually present in complex 3 is the difference in mobilities of complexes 2 and 3, which is dependent on fractions known to contain TFIIB activity, and the differences in DNAase I protection in complex 3, compared with complex 2. TFIIA remains in complex 3, as the mobility of this complex varies with different sources of TFIIA (data not shown).

The DNAase I protection pattern of complex 3 is quite interesting. In addition to protection of the TATA element as observed in complex 2, partial protection of some cutting sites from -10 to +10 on the coding strand was observed. No corresponding protections were detected on the noncoding strand. This pattern suggests that TFIIA may be associated specifically with one strand and extend from the TFIID-TFIIA complex to beyond the transcription initiation site. Purified RNA pol II added to the binding mixtures did not generate any new complexes unless TFIIA as well as TFIID was present, suggesting that in the absence of TFIIB, polymerase does not bind stably to TFIID. It is interesting that TFIIB has been shown to associate with RNA pol II in solution (Zheng et al., 1987). Therefore, it seems likely that TFIIB acts as a "bridging" molecule between the TFIID-TFIIA-TATA element complex and pol II.

If TFIIA does act as a bridge, the TFIIA-pol II interaction may also be involved in "measuring" the distance from the TATA element to the initiation site. In this regard, it will be informative to isolate the yeast homolog of TFIID, because in S. cerevisiae the initiation site is generally a greater distance from the TATA element (60-120 bp) than in higher eukaryotes (30 bp; see Guarente, 1987, for review). At least one domain of the TFIID protein, that involved in contacting the TFIID-TFIIA complex, must be conserved over evolution, since the yeast TFIID and TFIIA function with the other mammalian transcription factors. It remains to be seen whether the domain of TFIID that interacts with polymerase is also conserved.
tion may explain why we have so far been unable to substitute yeast pol II for the mammalian enzyme.

Complexes 4 and 5
The addition of purified RNA pol II is necessary for generation of complexes 4–7. As discussed above, polymerase does not stably bind to the TFIIID–TATA element complex unless TFIIB is also present in the complex. Complexes 4 and 5 apparently represent binding of pol II to complex 3. Protection from DNAase I cleavage is quite extensive in these complexes, extending as far downstream as +20. In addition, the upstream boundary of protection on the template strand, which is about −43 in complex 3, extends to −47 upon pol II binding (complexes 4–7). If this protection is due to the polymerase molecule, and not a conformational change in TFIIA or TFIID, then a portion of pol II would be in close proximity to upstream sites and any stimulatory transcription factors bound there. The MLP transcription factor (MLTF or USF) binding site, which was not contained on the probes used in the gel shift assays presented here, is located from −50 to −66 (Carlton et al., 1986; Sawadogo and Roeder, 1985b; Moncollin et al., 1986). This juxtaposition would allow for direct interactions between polymerase and at least the most proximal upstream factor.

Curiously, the pol II-dependent complexes occur in pairs: complexes 4 and 5 (pol II binding without TFIIE) and complexes 6 and 7 (binding with TFIIE). The members of each pair always show identical DNAase I footprints. While the basis of the mobility difference between the members of each pair is not known, at least three plausible explanations exist. First, the mobility difference may be due to binding by different forms of pol II. It is well documented that purification of pol II results in the isolation of intact and proteolyzed forms (Hodo and Blatti, 1977). A second possibility is that the mobility difference reflects the binding of another protein to the RNA polymerase complex. A potential candidate is the protein S-II (also called TFIIIS and RAP 38), a 37–40 kd molecule that has been shown to bind RNA pol II and to stimulate transcription elongation (Tirronen et al., 1983; Reinberg and Roeder, 1987b; Sopta et al., 1985). A third possibility is that the mobility difference reflects a conformation change in the polymerase–DNA interaction. Such an explanation has been proposed for a doublet observed with purified E. coli polymerase. In similar gel assays, purified bacterial holoenzyme generated two promoter-specific complexes, the relative amounts of which could be affected by varying binding and gel temperatures (Straney and Crothers, 1986). Further experiments are required to decide which, if any, of the above explanations are correct.

We have not observed a strict requirement for added TFIIA in order to form the higher order complexes 4–7. This might be interpreted as indicating that TFIIA is not required for later steps in initiation complex formation. Interactions between TFIIA and TFIID may occur in the absence of TFIIA, since incubation of the purified TFIIA and the TFIIA fraction generated a complex with a mobility intermediate to those of complexes 2 and 3 (data not shown and Figure 4). If TFIIA facilitates, but is not absolutely required for complex formation, varying conditions (such as the concentration of the other factors) might make the rate of the TFIIA-promoted step limiting, or not limiting, and therefore lead to conflicting observations. An alternative explanation is that other fractions are contaminated with TFIIA activity. Either possibility is consistent with the partial requirement for added TFIIA in the reconstituted in vitro transcription reaction.

Kinetic experiments, involving preincubation of transcription factors with a DNA template, followed by challenge with inhibitors or a second template, have suggested an association of pol II with complexes dependent upon TFIIA and TFIID, in the absence of TFIIE and TFIIE (Fire et al., 1984, Reinberg et al., 1987). A stable association of pol II with the TFIIA–TFIID–TATA element complex was not observed in this study, nor in other studies that assayed stable complex formation by DNAase I footprint analysis (Van Dyke et al., 1988) or exchange of drug-resistant and -sensitive pol II (Carthew et al., 1988). Thus, it is likely that the earlier kinetic studies either were flawed by the use of contaminated fractions, or were complicated by a non-specific association of pol II with the template DNA during preincubation, or that the hypothetical association between pol II and the TFIIA–TFIID–TATA element complex is not sufficiently stable to be detected in the other assays.

Complexes 6 and 7
Addition of TFIIE to complexes 4 and 5 generates complexes 6 and 7. TFIIE apparently binds downstream of the polymerase, protecting sequences in the +20 to +30 regions from DNAase I cleavage. TFIIE has been shown to bind pol II in solution (Reinberg and Roeder, 1987a). Indeed, TFIIE is related or identical to the RAP 30/74 complex (Flores et al., 1988), two complexed proteins that were isolated on the basis of their affinity for pol II immobilized on a column matrix (Sopta et al., 1985). The RAP 30/74 complex has been shown to be necessary for accurate transcription initiation (Burton et al., 1986), but can be substituted for by a partially purified TFIIE fraction (Flores et al., 1988). Although it is possible to separate TFIIB, TFIIIE, and pol II and to construct complexes representing the binding of each factor, it is possible that in vivo pol II and TFIIIE, and perhaps TFIIB, are associated before binding to the TFIIA–TFIID–TATA element complex.

Fractions containing TFIIIE (Sawadogo and Roeder, 1984; Reinberg and Roeder, 1987a) and RAP 30/74 (Burton et al., 1986) have been reported to contain a DNA-dependent ATPase activity. Transcription has an energy requirement that can be satisfied by hydrolysis of either ATP or dATP, generating what has been termed an "activated" transcription complex (Runick et al., 1982; Sawadogo and Roeder, 1984). These same two nucleotides cause an apparent dissociation of complexes 6 and 7 and generation of complexes 6* and 7*. At least with dATP, the loss of complexes 6 and 7 also correlated with an apparent increase in complexes with mobility similar to complexes 4 and 5. This dissociation is accompanied by a loss (in complexes 6* and 7*) of the TFIIA-dependent DNAase I protection between +20 and +30. A similar loss of
DNAase I protection was observed following nucleotide triphosphate addition to transcription complexes assembled in solution on the MLP in nuclear extracts (Cai and Luse, 1987b) or with partially purified factors (Van Dyke et al., 1988). Both studies showed that complete complexes had an upstream DNAase I protection boundary of about −42, and a downstream boundary of +30 that retreated to +25 upon addition of either ATP or dATP. Although these studies did not have the sensitivity of the native gel shift assay to resolve the intermediate complexes 3–5, the protections described were very similar to the protection seen with complexes 6 and 7, and probably represent the same entity: a complete initiation complex. We propose that the loss of the downstream protection and the instability of complexes 6 and 7 result from an ATP-dependent dissociation of TFIIE from the initiation complex. This may be related to the DNA-dependent ATPase detected in TFIIE fractions. One likely function for a DNA-dependent ATPase that binds downstream of the polymerase molecule might be as a helicase, which would unwind the DNA, making the template strand accessible to the polymerase. On linear molecules such as those used in this study, the putative helicase would run off the template, which would explain the ATP-dependent dissociation of TFIIE.

Complexes 8 and 9
Addition of NTPs that allow transcript elongation to +11 results in formation of complexes 8 and 9. These complexes have been shown to contain accurately initiated transcripts. Complexes 8 and 9 migrate near complexes 4 and 5 respectively, consistent with the model that TFIIE dissociates from the activated transcription complex. The toxin a-amanitin blocks formation of complexes 8 and 9, but not other complexes, and not the ATP/dATP effects discussed above. This is consistent with its role as an inhibitor of elongation, but not binding, by pol II (Cochet-Meilhac and Chambon, 1974).

The identification of an ordered set of complexes, each containing some subset of the previously isolated general transcription factors, defines a reaction pathway for transcription initiation by pol II. While the binding of each factor to the complex may not be strictly sequential in vivo, with some binding together as pre-existing complexes, it is likely that the contacts between factors defined by the ordered pathway shown in Figure 8 occur in vivo. As has been long appreciated, initiation of transcription in eukaryotic systems is a complicated reaction, involving many proteins. As a conceptual analogy, prokaryotic DNA replication seems to be more applicable than prokaryotic transcription. It should be noted that binding by any or all of the defined transcription factors is potentially rate-limiting for initiation and therefore a potential step for regulation of transcription. It is clear that analysis of complexes by native gel electrophoresis will not only be useful for further elucidating the mechanisms of action of the general factors in transcription, but also for discovering which step(s) and general factors are affected by the regulatory transcription factors that bind upstream of the TATA element.

Experimental Procedures

Protein Purification
RNA polymerase II was purified from calf thymus according to Hodo and Blatt (1977) to a final protein concentration of 150 µg/ml. By silver-stained gel analysis, the polymerase is ~90% pure and has the expected subunit composition. TFIID was purified approximately 80,000-fold from S. cerevisiae, to the Superose-12 FPLC step, as previously described (Buratowski et al., 1988). The final protein concentration was 2–5 µg/ml, and the protein was judged to be 30–50% pure by silver-stained gel analysis (Hahn et al., unpublished data). Calf thymus TFIIA was purified to step V (20,000-fold; 2.0 µg/ml of final protein concentration) as previously described (Samuels and Sharp, 1986). HeLa fraction [AB] (Samuels et al., 1982), which had a protein concentration of 4 mg/ml, was used as a source of partially purified human TFIIA. Yeast TFIIA was partially purified by chromatography of a yeast extract over Heparin-SEPSEPHOR and DEAE-SEPSEPHOR to a final protein concentration of 0.5 mg/ml; the detailed purification will be published elsewhere (Hahn et al., unpublished data). HeLa fraction [CBI] (Samuels et al., 1982) was used as the source of partially purified human TFIIIB and TFIIE. Separation of TFIIIB and TFIIE was carried out in two ways. One ml of fraction [CBI] (7 mg of protein/ml) in Buffer A + 100 mM KCl (Samuels et al., 1982) was loaded onto a 1 ml DEAE-Sephaloc column (Pharmacia) at a flow rate of 3 column volumeshr. The flowthrough fraction (termed fraction [CBA], containing 1 mg of protein/ml) was collected and contained the majority of TFIIE activity (Dignam et al., 1983b). After washing with three column volumes of Buffer A + 100 mM KCl, the column was then eluted with three column volumes of Buffer A + 300 mM KCl. Fractions of 300 µl were collected and assayed as described below. TFIIB activity was found to peak at 75 mM KCl, while TFIIE was found to elute at 200 mM KCl.

DNA Probes
DNA for probes was from the plasmids pFW (MLP sequences −53 to +33 cloned into the Sma site of pUC13) or pLP (−174 to +33; Chodosh et al., 1989). Fragments were prepared by excising the MLP inserts with either EcoRI and HindIII or EcoRI and HindIII (NEB). Probes were made by end labeling the EcoRI-HindIII fragment with Klenow enzyme (Boehringer Mannheim) and [α32P]-dATP (NEB) or, to label the other strand, by labeling the EcoRI-HindIII fragment with [α32P]-dCTP (NEB). The pUC fragment lanes in the RNA labeling of initiated transcription complexes (Figure 6B) was a 140 bp PvuII-HindIII fragment excised from pUC13 that contained the polyprotinker. Probes and competitors were gel-purified.

Gel Electrophoresis DNA Binding Assay
The indicated protein components were incubated with 0.5–2 ng of probe (roughly 20,000 cpm) for 20–30 min at 30°C. All binding reactions were done at transcription buffer conditions (12 mM HEPES-NaOH [pH 7.9], 12% glycerol, 1 mM EDTA, 0.6 mM DTT, 50–100 mM KCl, 5 mM MgCl2 and also contained 0.2 mg/ml of BSA and 5–20 µg/ml of poly[d(C3G5)CCG(C3G5)] in a total volume of 10 µl. Reactions were loaded onto 4–5% acrylamide native gels (0.15 × 16 cm, 40:1 mono: bis ratio, 2.5% glycerol) and run in buffer consisting of 25 mM Tris base, 190 mM glycine, and 1 mM EDTA (final pH 8.5). Electrophoresis was for 2–3 hr at 20–25 mA at room temperature, until bromophenol blue had run to the bottom of the gel. They were then transferred to Whatman 3MM paper, dried, and autoradiographed.

DNAase I Footprinting
Binding reactions were performed as described above for the gel electrophoresis DNA binding assay. DNAase I (Wortthington) was then added at approximately 10 µg/ml for 1 min at room temperature. For solution footprinting, the reactions were stopped with transcription stop mix (Samuels et al., 1982), phenol-chloroform extracted, chloroform extracted, EtOH precipitated, washed with 70% EtOH, dissolved in loading buffer (95% formamide, 1× TBE), heat denatured, and run
on a standard 8% acrylamide, 83 M urea sequencing gel. Alternatively, binding reactions (scaled up 5-fold) were loaded onto the native gels immediately after DNAase I treatment. After native gel electrophoresis, the complexes were electrophoretically blotted onto nitrocellulose filters (Schleicher & Schull) in native gel running buffer (about 2 hr at 80 mA in a Bio-Hard transblot electrophoretic apparatus). The nitrocellulose membrane was then autoradiographed, and the areas corresponding to the labeled complexes were cut out. Labeled DNA was eluted from the membrane into buffer containing 10 mM Tris-HCl (pH 7.9), 1 mM EDTA, and 1 M NaCl at 68°C for 20 min. The supernatant was then extracted, processed, and electrophoresed as described for solution footprinting.

In Vitro Transcription

Transcriptions were performed essentially as previously described (Buratowski et al., 1988). Assays for TFIIIB and TFIIIE together contained 0.5 μl (0.34 ng) of yeast TFIIID (Mono S fraction), 0.5 μl of HeLa fraction [AB], which contains TFIIA, 0.5 μl of HeLa fraction [CD], and 0.2 μl of calf thymus RNA pol II in addition to the standard buffer, template, and nucleotide concentrations. Assays for TFIIIE in addition contained 0.5 μl of the DEAE-Sephacel flowthrough fraction [CBA] described above, as a source of TFIIIE. Assays for TFIIIB contained an additional 1 μl of the 300 mM KCl fraction [CBB] of the DEAE-Sephacel column, as a source of TFIIIE.

Acknowledgments

We wish to thank Mark Samuels for purified TFIIA and HeLa fractions.

This work has been supported by U.S. Public Health Service Grant number PO1-CA42063 from the National Institutes of Health, Cooperative Agreement number CDR-8803014 from the National Science Foundation, and partially by National Cancer Institute core grant number P30-CA41051 to P. A. S.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734.

Received September 26, 1988; revised November 29, 1988.

References


