PROTEIN-PROTEIN INTERACTIONS INVOLVED IN TRANSCRIPTIONAL ACTIVATION BY ACTIVATORS

by

Hua Xiao

A thesis submitted in conformity with the requirements for the Degree of Doctor of Philosophy
Graduate Department of Molecular and Medical Genetics
University of Toronto

©Copyright by Hua Xiao 1996
The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author’s permission.

L’auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L’auteur conserve la propriété du droit d’auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.
Abstract

The mechanism by which transcriptional activators increase the levels of messenger RNA is poorly understood. It is generally believed that activators directly or indirectly contact the basal transcription apparatus through their activation domains to stimulate transcription. In this thesis, I first showed that two acidic transactivators, herpes simplex virus VP16 and human p53, directly interact with the multi-subunit human general transcription factor TFIIH and its Saccharomyces cerevisiae counterpart, factor b. The VP16- and p53-binding domains in these factors lie in the p62 subunit of TFIIH and in the homologous subunit, TFB1, of factor b. Point mutations in VP16 that reduce its transactivation activity in both yeast and mammalian cells weaken its binding to both yeast and human TFIIH. This suggests that binding of activation domains to TFIIH is an important aspect of transcriptional activation. As well, I also demonstrated that the human immunodeficiency virus activator Tat, but not the cellular proteins Sp1 and CTF, bound TFIIH. The correlation between the ability of activators to stimulate elongation and their ability to bind TFIIH suggests that TFIIH-activator interaction may be involved in stimulation of transcriptional elongation. Finally, I have identified two human proteins (57 and 30 kDa) that bind to the activation domain of Tat (amino acids 1-48). One (57 kDa) is the human cellular protein CD46 and the other (30 kDa) is a novel human protein for which there is an existing cDNA. Point mutations in the amino-terminal portion of Tat that reduce its transactivation activity also weaken its binding of CD46 and the 30 kDa protein. Moreover, CD46 can bind to both of the glutamine-rich activation domains of Sp1. Coexpression of CD46 in human HeLa cells, which already contain CD46, further increased transcriptional activation by Tat of the HIV1-LTR. Therefore, Tat and Sp1 may form a complex mediated by CD46 to increase the efficiency of transcription.
TABLE OF CONTENTS

Abstract ii
Table of contents iii
List of figures v
Chapter I. Introduction 1
   A. General transcription factors and RNA polymerase II. 4
   B. Elongation factors. 22
   C. Transcriptional activators. 27
   D. Mechanisms of transcriptional activation. 32
   E. Regulation of transcription by HIV I Tat. 42
Chapter II. Binding of basal transcription factor TFIH to the acidic activation domains of VP16 and p53. 48
   Abstract 50
   Introduction 51
   Materials and Methods 54
   Results 57
   Discussion 64
   References 79
Chapter III. The activation domain of Tat binds to the general transcription factor TFIH. 92
   A. Summary 93
   B. Introduction 94
   C. Materials and Methods 96
   D. Results 97
   E. Discussion 104
Chapter IV. Identification of other human cellular proteins that interact with the activation domain of Tat. 106
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td></td>
<td>107</td>
</tr>
<tr>
<td>A.</td>
<td>Introduction</td>
<td>108</td>
</tr>
<tr>
<td>B.</td>
<td>Materials and Methods</td>
<td>110</td>
</tr>
<tr>
<td>C.</td>
<td>Results</td>
<td>113</td>
</tr>
<tr>
<td>D.</td>
<td>Discussion</td>
<td>124</td>
</tr>
</tbody>
</table>

Chapter V. Conclusions, perspectives and directions for future research. 127

Chapter VI. References. 132
<table>
<thead>
<tr>
<th>LIST OF FIGURES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chapter II. Figure 1</td>
</tr>
<tr>
<td>Chapter II. Figure 2</td>
</tr>
<tr>
<td>Chapter II. Figure 3</td>
</tr>
<tr>
<td>Chapter II. Figure 4</td>
</tr>
<tr>
<td>Chapter II. Figure 5</td>
</tr>
<tr>
<td>Chapter II. Figure 6</td>
</tr>
<tr>
<td>Chapter III. Figure 1</td>
</tr>
<tr>
<td>Chapter III. Figure 2</td>
</tr>
<tr>
<td>Chapter III. Figure 3</td>
</tr>
<tr>
<td>Chapter III. Figure 4</td>
</tr>
<tr>
<td>Chapter IV. Figure 1</td>
</tr>
<tr>
<td>Chapter IV. Figure 2</td>
</tr>
<tr>
<td>Chapter IV. Figure 3</td>
</tr>
<tr>
<td>Chapter IV. Figure 4</td>
</tr>
<tr>
<td>Chapter IV. Figure 5</td>
</tr>
<tr>
<td>Chapter IV. Figure 6</td>
</tr>
</tbody>
</table>
Chapter I
Introduction
Gene expression is largely regulated at the level of transcription. Understanding how transcription is regulated in a cell is very important for the development of new therapies for treatment of human diseases. Over the past few decades, we have learned that DNA carries genetic information not only for synthesizing protein and RNA, but also for regulating gene expression. The DNA elements that regulate transcription of genes consist of the core promoter, upstream regulatory sequences and enhancers. The core promoter, which contains the TATA-box element or/and the initiator (Inr), controls the basal level of transcription and the start point of transcriptional initiation. The upstream regulatory sequences and enhancers, which are recognized by transcriptional regulatory proteins, mediate the regulation of transcription. It is estimated that up to 10% of human genes may encode transcriptional regulatory proteins (Kingston and Green, 1994). A great deal of effort has been devoted to the analysis of how these factors activate transcription. It is generally believed that transcription factors must ultimately modulate RNA polymerase activity to regulate transcription. In 1959, Weiss and Gladstone first identified a mammalian RNA polymerase activity (Weiss and Gladstone, 1959). Three types of RNA polymerases, RNA polymerase I, RNA polymerase II and RNA polymerase III, were subsequently identified in mammalian cells (Roeder and Rutter, 1969). RNA polymerase I synthesizes the 18 S and 28 S ribosomal RNAs. RNA polymerase II synthesizes messenger RNA precursors and some snRNAs. RNA polymerase III synthesizes other snRNAs, all tRNAs and the 5 S ribosomal RNA. The introductory chapter of this thesis reviews current advances in the study of regulation of transcription by RNA polymerase II.

The structures of the core promoters controlled by RNA polymerase II are highly variable. They often contain both a TATA box, typically located at position -25 to -30, and an initiator element (weak consensus PyPyA+1NT/APyPy) overlapping the transcription start site (Grosschedl and Birnstiel, 1980; Wasylyk et al., 1980; Corden et al., 1980). However, many cellular genes contain only one of these elements and some genes do not contain any of them.
The core promoter mediates not only the assembly of an initiation complex containing RNA polymerase II and general transcription factors, but also gene regulation.
A. General transcription factors and RNA polymerase II.

Unlike prokaryotic RNA polymerase, which can directly recognize certain strong promoters, eukaryotic RNA polymerase II does not accurately initiate transcription on its own at any known promoter. Instead, RNA polymerase II was shown to produce specific transcripts from adenovirus DNA templates only when S-100 extracts derived from tissue culture cells were added into the reactions (Weil et al., 1979). Subsequently, whole cell extracts and nuclear extracts have been developed in order to reconstitute accurate transcription in vitro (Manley et al., 1980; Dignam et al., 1983). It was found that a DNA element containing only a TATA-box sequence and the cap site was able to direct initiation of a specific basal level of transcription (Hu and Manley, 1981; Reinberg and Roeder, 1987). Fractionation of HeLa whole cell extracts and nuclear extracts resulted in the identification of the factors required for reconstitution of transcription by RNA polymerase II in vitro (Zawel and Reinberg, 1992). The factors that are required to produce a basal level of transcription along with RNA polymerase II are called basal transcription factors. Because these basal factors may be required for transcription of most or all class II genes, they are also called general transcription factors. The human general transcription factors that have been identified so far are TFIIA, TFIIB, TFIID, TFIIE, TFIIF (RAP30 and RAP74), TFIIH and TFIIJ (Zawel and Reinberg, 1992). Based on experiments involving sequential gel mobility shift assays, they assemble together with RNA polymerase II in an ordered manner onto promoter-containing DNA (Van Dyke et al., 1988; Buratowski et al., 1989; Maldonado et al., 1990).

TFIID

TFIID is a multisubunit factor that is the first general transcription factor to bind to the TATA-box DNA element (Davison et al., 1983; Fire et al., 1984; Reinberg et al., 1987) and is the only general transcription factor known to specifically interact with DNA. It has been extensively studied in the yeast, *Drosophila* and human systems. The fact that a yeast fraction can substitute functionally for human TFIID in a reconstituted *in vitro* transcription system
enabled several research groups to purify yeast TFIID to homogeneity and to clone the cDNA encoding it (Horikoshi et al., 1989; Hahn et al., 1989; Schmidt et al., 1989; Cavallini et al., 1989; Eisenmann et al., 1989). Subsequently, the cDNAs encoding human, Drosophila and Arabidopsis TFIIDs were also cloned (Peterson et al., 1990; Kao et al., 1990; Hoffman et al., 1990; Hoey et al., 1990; Muhich et al., 1990; Gasch et al., 1990). This recombinant TFIID, now referred to as the TATA box-binding protein (TBP), binds specifically to the TATA-box DNA element and can replace the TFIID fraction for basal transcription in vitro. Comparison of the amino acid sequences of TBP from a large variety of species revealed that the carboxy-terminal portion of TBP is highly conserved and contains two imperfect repeat motifs separated by a basic region (Greenblatt, 1991b). In contrast, the amino-terminal portion of TBP is very divergent in amino acid sequence and size among diverse species. Any small deletion in the conserved region abolishes both DNA-binding activity and transcription activity (Horikoshi et al., 1990). Unlike most DNA-binding proteins that interact with DNA in its major groove, TBP contacts DNA in its minor groove (Starr and Hawley, 1991; Lee et al., 1991a). The crystal structures of the Arabidopsis thaliana and Saccharomyces cerevisiae TBPs were solved by using X-ray diffraction analysis (Nikolov et al., 1992; Chasman et al., 1993). The structure of TBP was described as a saddle with the underside of the molecule containing the DNA interaction surface. The DNA binding surface is a curved eight-stranded antiparallel β-sheet. The co-crystal structure of TBP with a TATA-box element subsequently revealed that the TATA box is centered on the pseudodyad of the protein and the beginning of TATA box contacts the second direct repeat (Kim et al., 1993a; Kim et al., 1993b). The binding of the saddle-shaped TBP to DNA induces a conformational change in DNA, in which the TATA box is bent towards a narrowed major groove and away from a widened minor groove with sharp kinks at both its 5' end and its 3' end. Between the kinks, the right-handed, double-stranded DNA is smoothly curved and partially unwound (Kim et al., 1993a; Kim et al., 1993b). The seat of the saddle presents a large convex surface to which other proteins bind during transcription initiation (Nikolov et al., 1992). Upon DNA binding, the structure of TBP also undergoes a conformational change, in
which a $10^\circ$ twist of one domain with respect to the other shifts the relative positions of the stirrup-loops by 5 Å.

Unlike yeast TBP, human and *Drosophila* TBPs were found to be tightly associated with other polypeptides, now known as TBP-associated factors (TAFs) (Dynlacht et al., 1991; Tanese et al., 1991; Zhou et al., 1992). These TAFs are not required for basal transcription. Four distinct sets of TBP-TAF complexes, including SL1 (Comai et al., 1992), TFIID (Dynlacht et al., 1991; Tanese et al., 1991), TFIIB (Lobo et al., 1992) and SNAPc (Sadowski et al., 1993), have been found to be specifically required for transcription from different promoters in eukaryotic cells. SL1 is an essential factor for transcription of the large rRNA genes by RNA polymerase I. It consists of TBP and three TAFs of sizes 110, 63 and 48 kDa (Comai et al., 1992). TFIIB and SNAPc are distinct multisubunit factors containing TBP and different TAFs that are involved in transcription by RNA polymerase III. TATA-less RNA polymerase III promoters are recognized by TFIIC which, sometimes in conjunction with TFIIB, binds to internal promoter elements. TFIIC then recruits TFIIB which, in turn, recruits RNA polymerase III (White et al., 1992b). SNAPc is required for transcription of small-nuclear RNA genes, including the U6 snRNA gene, by RNA polymerases II and III. It consists of at least three TAFs, including SNAP43, SNAP45 and SNAP50, in addition to TBP (Henry et al., 1995). SNAPc appears to bind specifically to a non-TATA-box DNA-binding site, the proximal sequence element PSE (Henry et al., 1995).

Although much of the TBP in yeast extracts exists in monomeric form, a complex containing TBP and TAFs can be immunopurified from yeast whole cell extracts with antibodies against TBP (Poon and Weil, 1993). Thus, yeast TBP may also associate with TAFs *in vivo*.

To date, eight cDNAs encoding *Drosophila* TAFs, nine cDNAs encoding human TAFs and some of the yeast TAF cDNAs that are involved in RNA polymerase II transcription have been cloned. TAF11250 was previously cloned and identified as CCG1, a gene rescuing the G1 arrest of a temperature-sensitive hamster cell line (ts13) (Sekiguchi et al., 1991; Hisatake et al., 1993a). TAF11250 binds directly to TBP (Hisatake et al., 1993a) and other TAFs. Other
*Drosophila* TAFs, such as TAF\(_{130}\alpha\) and TAF\(_{150}\), have also been shown to directly interact with TBP (Yokomori et al., 1993b; Verrijzer et al., 1994). In contrast, *Drosophila* TAF\(_{110}\) and TAF\(_{60}\) appear to interact with TBP through other TAFs. As will be discussed in more detail later, many TAFs function as co-activators to participate in events of transcriptional activation by activators. *Drosophila* TAF\(_{110}\) and TAF\(_{250}\) can partially support transcriptional activation by Sp1. The level of activated transcription can be further increased by TAF\(_{150}\) (Chen et al., 1994).

TBP is also required for transcription of genes that do not contain a consensus TATA box. These genes include class I genes transcribed by RNA polymerase I and TATA-containing and TATA-less class III genes transcribed by RNA polymerase III, as well as TATA-less class II genes. Pugh and Tjian (1991) demonstrated that transcription of TATA-less class II genes also requires a multisubunit TFIID complex. Indeed, TAF\(_{150}\) was found to bind directly to sequences overlapping the Inr and extending to position +35 of some promoters (Verrijzer et al., 1995). The utilization of the Inr and downstream promoter elements minimally requires a trimeric TBP-TAF\(_{250}\)-TAF\(_{150}\) complex. Depending on the structures of core promoters, the TAFs can stabilize or destabilize the DNA-binding of TFIID and modulate promoter utilization and selectivity (Verrijzer et al., 1995).

**TFIIB**

TFIIB is a general transcription factor with a molecular mass of 33 kDa. It binds to a TBP-DNA complex or TFIIA-TFIID-DNA complex in gel mobility-shift assays (Buratowski et al., 1989; Maldonado et al., 1990). Genes encoding TFIIB have been cloned from human, rat, *Xenopus*, *Drosophila*, and yeast. Comparison of these amino acid sequences has revealed some similarities. The amino acid sequence of TFIIB deduced from a human TFIIB cDNA clone was found to have short stretches of weak similarity with the bacterial initiation factor σ70 (Ha et al., 1991; Malik et al., 1991). Two independent functional domains, the amino-terminal domain and carboxy-terminal domain of TFIIB, have been identified by several groups. The amino-terminus of TFIIB, containing a putative zinc-binding domain, is essential for TFIIB function *in vitro*.
(Buratowski and Zhou, 1993). This portion of TFIIB is able to interact with the small subunit of TFIIF (RAP30) (Ha et al., 1993). Such a protein-protein interaction may enable TFIIB to recruit TFIIF and RNA polymerase to the promoter (Ha et al., 1993; Hisatake et al., 1993b; Barberis et al., 1993; Buratowski and Zhou, 1993). Substitutions of cysteine codons within the putative zinc finger of TFIIB eliminate the ability of TFIIB to recruit the TFIIF-RNA polymerase II complex to the promoter (Buratowski and Zhou, 1993), suggesting that the zinc finger plays a role in the incorporation of RNA polymerase into the initiation complex. The carboxy-terminal portion of TFIIB consists of two imperfect direct repeats and a putative basic amphipathic α-helix located between the two repeats. This domain is necessary and sufficient for interaction with TBP. The crystal structure of a TFIIB-TBP-TATA element ternary complex was recently described at 2.7 Å resolution. The carboxy-terminal portion of TFIIB is a two-domain structure resembling cyclin A. The two repeat domains are rotated by 90° with respect to one another and connected by a short linker (Nikolov et al., 1995). Each of these domains consists of five α-helices. The structural similarity between TFIIB and cyclin A implies that TFIIB may regulate the activity of Cdk7/cyclinH, a component of TFIIB in the initiation complex. The co-crystal structure also revealed that TFIIB binds the C-terminal stirrup of TBP and contacts the phosphoribose backbone of the DNA both upstream and downstream of the center of the TATA element (Nikolov et al., 1995). The three-dimensional structure in solution of the C-terminal core domain of human TFIIB was also determined by using nuclear magnetic resonance (NMR) spectroscopy (Bagby et al., 1995). The solution structure of TFIIB is similar to the crystal structure of TFIIB described by Nikolov et al. (1995).

TFIIB may control the start site of transcription since a mutation in the yeast TFIIB gene (SUA7) can change transcription start-site selection (Pinto et al., 1992). This idea was further supported by a study in vitro in which replacement of TFIIB from S. cerevisiae by its counterpart from Schizosaccharomyces pombe could switch start-site selection to the pattern characteristic of S. pombe (Li et al., 1994). In this study it was also found that use of TFIIB from a particular
species necessitates the use of RNA polymerase II derived from the same species, indicating that there is a functional interaction between TFIIB and RNA polymerase II.

**TFIIF**

The further purification of a crude TFIIE fraction resulted in the identification of a general transcription factor that is now known as TFIIF (Flores et al., 1988). The subunits of TFIIF are identical to RAP30 and RAP74 (Sopta et al., 1985), which were initially purified by using RNA polymerase II affinity chromatography (Burton et al., 1988). RAP30 and RAP74 were shown to be essential factors for initiation of transcription by RNA polymerase II on several promoters when linear templates were used (Burton et al., 1988). Native RAP30 and RAP74 may form a tetrameric complex containing two subunits of RAP30 and two subunits of RAP74 (Flores et al., 1990; Kitajima et al., 1990). The human cDNAs encoding RAP30 and RAP74 have been cloned and sequenced (Sopta et al., 1989; Finkelstein et al., 1992; Aso et al., 1992). Two regions of RAP30 share sequence similarity with the bacterial sigma factors (Sopta et al., 1989). One of these regions in $\sigma^{70}$ is needed for interaction with bacterial RNA polymerase (McCracken and Greenblatt, 1991). Interestingly, like bacterial sigma factors, TFIIF was shown to be able to bind to *E. coli* core RNA polymerase. This interaction between TFIIF and bacterial core RNA polymerase could be specifically disrupted by *E. coli* $\sigma^{70}$ (McCracken and Greenblatt, 1991). These findings led to the proposal that TFIIF and bacterial sigma factors may have some similar functional role in transcription. Consistent with this idea, the RAP30/74 purified from HeLa cell extracts and its rat counterpart, $\beta\gamma$, purified from rat liver, as well as recombinant RAP30, can prevent RNA polymerase II from binding non-specifically to DNA (Killeen and Greenblatt, 1992; Conaway and Conaway, 1990). RAP30 can also recruit RNA polymerase into a preinitiation complex containing TBP and TFIIB (Flores et al., 1991; Killeen et al., 1992). This property of RAP30 probably partly occurs because of its ability to interact with RNA polymerase II (McCracken and Greenblatt, 1991) and TFIIB (Ha et al., 1993) and perhaps at least partly occurs because its C-terminal region resembles region 4 of $\sigma$ factors and binds non-specifically to DNA (Tan et al., 1994). The amino-terminal 110 amino acids of RAP30 and the amino-
terminal residues 62-171 of RAP74 are required for interacting with each other to form a heteromer (Yonaha et al., 1993).

The sequence of RAP74 revealed three structurally distinct regions, a globular amino-terminal domain (residues 1-179), a charged domain (residues 180-356), and a globular carboxy-terminal domain (residues 357-517) (Aso et al., 1992; Finkelstein et al., 1992). Yonaha et al. (1993) demonstrated that both N-terminal residues 73-205 and carboxy-terminal residues 356-517 are essential for full transcription activity in vitro. Similarly, the Drosophila factor F5a, which is equivalent to mammalian RAP74, was found to contain a hydrophobic amino-terminal domain (residues 1-205), a central charged domain (residues 206-436), and a hydrophobic carboxy-terminal domain (residues 437-577). TFIIF (RAP30/RAP74) was shown to stimulate transcription at the levels of both initiation and elongation (Bengal et al., 1991). By deletion analysis of Drosophila F5a protein, the amino-terminal region was shown to be essential and sufficient for stimulating elongation in the presence of RAP30 (Kephart et al., 1994). It is noteworthy that a mutant F5a protein lacking residues 417-577 stimulates elongation more efficiently than wild type F5a (Kephart et al., 1994). The carboxy-terminal region of human RAP74 was also found to interfere with its ability to interact with RAP30. Thus, the carboxy-terminal region (417-577) may be involved in regulation of its elongation activity.

Both RAP30 and RAP74 are phosphorylated in vivo (Sopta et al., 1985). Kitajima et al. (1994) reported that alkaline phosphatase treatment of native TFIIF reduced binding of TFIIF to RNA polymerase II, impaired assembly of TFIIF into a preinitiation complex and decreased its initiation and elongation stimulating activities by 80-85% (Kitajima et al., 1994). This result supports a hypothesis in which phosphorylation of TFIIF stimulates its transcription activity at the levels of both initiation and elongation (Kitajima et al., 1994).

**TFIIE**

The transcription factor TFIIE was originally defined as an activity present in a chromatographic fraction derived from a HeLa cell extract and required for transcription in vitro. This fraction was later found to contain at least three basal transcription factor activities: TFIIE,
TFIIF, and TFIIH (reviewed by Zawel and Reinberg, 1992). Two groups simultaneously reported (Ohkuma et al., 1990; Inostroza et al., 1991) that TFIIE is a heterotetramer consisting of two polypeptides, 54 kDa for TFIIEα and 34 kDa for TFIIEβ. Both subunits of human TFIIE have been cloned and sequenced (Peterson et al., 1991; Ohkuma et al., 1991; Sumimoto et al., 1991). The amino acid sequences of TFIIE subunits revealed some interesting features. TFIIEα has 439 amino acids with a pI of 4.5 and several putative motifs: an amphipathic α-helix is present from residues 12 to 41, a potential zinc finger was found in residues 129-157, and a region with similarity to a protein kinase consensus sequence is present in amino acids 242-259. To date, however, no kinase activity has been found associated with highly purified TFIIE.

TFIIEβ is a basic protein having 291 amino acids. Sumimoto et al. (1991) found that one region of TFIIEβ was somewhat homologous to RAP30. TFIIE can be assembled into the preinitiation complex after RNA polymerase II is recruited to a promoter (Buratowski et al., 1989; Inostroza et al., 1991), and, consistent with this, it was shown that TFIIE interacts weakly with RNA polymerase II by a co-sedimentation assay (Reinberg and Roeder, 1987). Recently, Maxon et al. (1994) reported that TFIIEα binds selectively to the nonphosphorylated form of RNA polymerase II. TFIIEα was also shown to interact with TFIIH as well as with TBP and TFIID (Maxon et al., 1994). The TFIIE-TFIIH interaction is mediated by ERCC3, a subunit of TFIIH (Maxon et al., 1994). As well, TFIIEβ interacts with RAP30 and RNA polymerase II (Maxon et al., 1994). The role of TFIIE in transcription is not clear, although it is required for transcription from most, if not all, promoters when linear templates are used in *in vitro* transcription assays. The possible existence of multiple interactions of TFIIE with other factors, assuming they are all correct, within the preinitiation complex suggests that TFIIE may be involved in stabilizing the preinitiation complex. Based on their suggestion that TFIIE and TFIIH are needed for promoter clearance, Goodrich and Tjian (1994) suggested that TFIIE recruits TFIIH into initiation complexes by creating a region of single-stranded DNA in the promoter.

A variable requirement for TFIIE in the absence of TFIIH was found for transcription *in vitro* from various supercoiled core promoters (Tyree et al., 1993; Timmers, 1994), suggesting
an additional role of TFIIE in transcription. This idea was further supported by the observation that transcription from the adenovirus major late core promoter is more dependent on TFIIE than transcription from the adenovirus E4 or mouse mammary tumor virus promoters. In addition, the dependence of transcription on TFIIE for all three promoters is influenced by the ionic strength in the reaction (Timmers, 1994). Pan and Greenblatt demonstrated that efficient and accurate initiation in vitro from DNA templates with a premelted region upstream of the initiation site only required TBP, TFIIB and RNA polymerase II (Pan and Greenblatt, 1994). Similarly, Holstege et al. (1995) showed that transcription from a relaxed E4 promoter template with a premelted -8 to +2 region was completely independent of both TFIIE and TFIIH. It was therefore proposed that TFIIE may facilitate initiation by stimulating promoter melting on linear and some negatively supercoiled templates (Holstege et al., 1995).

**TFIIH**

TFIIH is a basal transcription factor which was initially co-purified with TFIIF and TFIIID (Flores et al., 1992). When highly purified TFIIF along with recombinant factors were used in a reconstituted transcription system, a factor, termed TFIIH, was required for transcription. Its rat counterpart is known as δ and its yeast equivalent is called factor b. Egly and his colleagues purified a basal transcription factor named BTF2 which was required for formation of a functional initiation complex (Gerard et al., 1991) and was later revealed to be identical to TFIIH. TFIIH (BTF2) is a multisubunit factor consisting of at least nine subunits with molecular masses of 34, 36, 40, 41, 44, 50, 62, 80 and 89 kDa (Gerard et al., 1991; Roy et al., 1994). TFIIH can bind to the preinitiation complex after the formation of the DABPolFE complex (Flores et al., 1992). Human TFIIH and rat δ factor, as well as yeast factor b, are associated with a protein kinase activity that specifically phosphorylates the carboxy-terminal domain (CTD) of the large subunit of RNA polymerase II (Lu et al., 1992; Serizawa et al., 1993a; Feaver et al., 1991). TFIIH is also associated with a DNA-dependent ATPase activity and an ATP-dependent DNA helicase activity (Schaeffer et al., 1993; Drapkin and Reinberg, 1994).
The p62 subunit of human TFIIH and the 72 kDa TFB1 subunit of yeast factor b were the first subunits of TFIIH to be cloned and sequenced (Fischer et al., 1992; Gileadi et al., 1992). It was found that the amino acid sequences of p62 and TFB1 share considerable similarity. Amino acid sequence analysis of the tryptic digest of purified polypeptides led to the demonstration that the 89 kDa subunit of BTF2 is the ERCC3 gene product, a presumed helicase implicated in the human DNA excision repair disorders Xeroderma Pigmentosum (XP) and Cockayne's Syndrome (Weeda et al., 1990; Schaeffer et al., 1993). ERCC3 contains seven consensus helicase motifs, a putative Mg²⁺-binding site, and a consensus nucleotide-binding site. The yeast counterpart of p89 has been identified as RAD25 or SSL2, which was previously shown to have a role in nucleotide excision repair (Gulyas and Donahue, 1992; Park et al., 1992). Indeed, both purified ERCC3 and RAD25 proteins have single-stranded DNA-dependent ATPase and ATP-dependent DNA helicase activities. A conditional lethal mutation (rad25-ts24) in the nucleotide-binding site of RAD25 demonstrated a general requirement for RAD25 in RNA polymerase II transcription in vivo (Qiu et al., 1993; Guzder et al., 1994b). This was further confirmed by the observation that the mutant protein rad25-ts24 is defective for transcription in vitro (Guzder et al., 1994b).

The genes encoding the 44 and 34 kDa subunits of BTF2 have also been cloned and sequenced (Humbert et al., 1994). The amino acid sequence of p44 revealed a significant overall homology with SSL1, a yeast protein that is also present in yeast TFIIH and is involved in UV resistance (Yoon et al., 1992). The p34 subunit has a zinc finger domain that is also found in p44 and SSL1, suggesting that p44 and p34 may share some common function.

The 80 kDa subunit of TFIIH was identified as ERCC2 (Schaeffer et al., 1994), which is the product of a gene responsible for the D complementation group of the DNA repair syndrome XP (Fletjer et al., 1992). Yeast RAD3, the homologue of human ERCC2, was also found to associate with factor b (Higgins et al., 1983). RAD3 binds directly to both the SSL2 and SSL1 proteins (Bardwell et al., 1994). As revealed by the two hybrid assay in yeast, SSL1 interacts with itself and with RAD3 and TFB1 (Bardwell et al., 1994). Unlike RAD25, a mutation in RAD3 which abolishes its ATPase/helicase activity and reduces its DNA repair function does not
affect transcription. Nevertheless, the RAD3 gene is essential for transcription by RNA polymerase II since a rad-3-ts mutant is defective in synthesizing mRNA (Guzder et al., 1994a). The direct role of RAD3 in RNA polymerase II transcription was confirmed by the finding that RAD3 protein can complement the transcription defect in an extract made from the rad3-ts14 strain (Guzder et al., 1994a). Both ERCC3 and ERCC2 possess helicase activity, but ERCC3 is a 3' to 5' helicase while ERCC2 is a 5' to 3' helicase (Schaeffer et al., 1994). The ATPase and helicase activities in purified BTF2 reflect the functioning only of its ERCC3 p89 polypeptide (Roy et al., 1994b), which is consistent with observations that only the helicase activity of the yeast ERCC3 homologue, RAD25, and not the helicase activity of the yeast ERCC2 homologue, RAD3, is needed for transcription by RNA polymerase II in vivo.

The presence of a CTD kinase activity in TFIIH (Feaver et al., 1991; Lu et al., 1992) focused attention on identifying this kinase. Amino acid sequences of tryptic peptides derived from 40 kDa and 36 kDa polypeptides that copurified with human TFIIH led to the identification of the TFIIH-associated kinase as the cyclin-dependent protein kinase CDK7/Cyclin H, also known as MO15 (Roy et al., 1994a; Shiekhattar et al., 1995; Serizawa et al., 1995). Similarly, a 33 kDa polypeptide that co-purified with yeast holo-TFIIH was identified as KIN28, which is a putative serine/threonine kinase from yeast and a member of the p34cdc2/CDC28 family of CDKs (Feaver et al., 1994). KIN28 is one of three subunits of TFIIK, a component of holo-TFIIH that can be separated from core TFIIH during purification (Feaver et al., 1994). MO15 and KIN28 are able to phosphorylate CTD polypeptides but are not sufficient to phosphorylate the CTD on RNA polymerase II (Roy et al., 1994a; Feaver et al., 1994). Mutations in KIN28 cause a rapid decrease in the steady-state mRNA levels for many genes, including CLB2, which encodes the major mitotic cyclin; CDC2, which encodes a cyclin-dependent kinase; ACT1, which encodes actin; and URA3, which encodes orotidine-5'-phosphate decarboxylase. These mutations do not, however, affect the steady-state levels of rRNA synthesized by RNA polymerase I (Cismowski et al., 1995). This identification of a cell cycle kinase as a component of the basal transcription machinery provides a potential link between control of the cell cycle and
transcription. Although human MO15 was identified as the catalytic subunit of a cyclin-dependent kinase-activating kinase (CAK) (Roy et al., 1994; Shiekhattar et al., 1995; Serizawa et al., 1995), KIN28 was found neither to regulate the phosphorylation state of the yeast cell cycle CDK nor to possess CAK activity in vitro (Cismowski et al., 1995).

**TFIIA**

TFIIA is a general transcription factor that can interact with TBP and stabilize the binding of TBP to the TATA-box DNA element. Early studies produced contradictory results regarding the requirement of TFIIA for transcription in a reconstituted transcription assay in vitro. Cortes et al. (1992) showed that the TFIIA fraction which was used in previous studies consists of at least two separate components, TFIIA and TFIIJ, that can be separated by TBP affinity chromatography. TFIIA binds to a TBP affinity column (Cortes et al., 1992; Coulombe et al., 1992) while TFIIJ activity flows through the column. By using a combination of conventional chromatography and TBP affinity chromatography methods, TFIIA was purified to homogeneity from various organisms. TFIIA purified from human cells consists of three polypeptides of 37, 19 and 13 kDa (Coulombe et al., 1992; Cortes et al., 1992). *Drosophila* TFIIA also consists of three subunits of 30, 20 and 14 kDa polypeptides. All three subunits of human TFIIA were required for formation of DAB complexes on promoters (Cortes et al., 1992). In contrast, yeast TFIIA is composed of two polypeptides, TOA1 (32 kDa) and TOA2 (13.5 kDa) (Ranish et al., 1992).

The cDNAs encoding TFIIA have been cloned from yeast, *Drosophila* and human cells. The cDNA encoding TOA1 has an open reading frame of 286 amino acids, whereas the cDNA encoding TOA2 has an open reading frame of 122 amino acids (Ranish et al., 1992). In humans and *Drosophila*, a single gene encodes a larger polypeptide, termed TFIIAα. The predicted amino acid sequence of the TFIIAα is 376 amino acids with an isoelectric point (pI) of 4.2, but recombinant TFIIAα migrates on SDS-polyacrylamide gels with an apparent molecular mass of 55 kDa. Similarly, *Drosophila* TFIIAα has an apparent molecular mass of 48 kDa. Recombinant TFIIAα can functionally replace the two largest subunits of natural TFIIA in the
formation of DA complexes. Biochemical studies have indicated that the two largest subunits of natural TFIIA are generated by processing precursors encoded by the human and *Drosophila* TFIIAα genes (DeJong and Roeder, 1993; Yokomori et al., 1993a; Ma et al., 1993). The smallest subunits of the *Drosophila* and human TFIIA (γ) have also been cloned (Yokomori et al., 1994; Sun et al., 1993; Ozer et al., 1994). The deduced amino acid sequence of human TFIIAγ has 85% similarity with *Drosophila* TFIIAγ and 64% similarity with yeast TFIIAγ.

TFIIA was found to have no effect on basal transcription when recombinant TBP was used in a transcription reaction and was found to have a stimulatory effect when native TFIIID was used in a reconstituted transcription system (Cortes et al., 1992). This implies that TFIIID fractions can contain transcriptional inhibitors and, indeed, TFIIA may be involved in competing with inhibitors that interact with TBP (Meisterernst and Roeder, 1991; Meisterernst et al., 1991). This was supported by the observation that repression of basal transcription by a specific TFIIID repressor, Dr2, is counteracted by TFIIA (Merino et al., 1993). However, TFIIA can stimulate transcription in a reaction containing highly purified TFIIID, which does not contain repressors. Therefore, TFIIA may directly stimulate transcription by interacting with a TAF. Indeed, Tjian and his colleagues demonstrated that TFIIA is associated with the *Drosophila* TAFIII110 and that TAFs and TFIIA can mediate differential utilization of the tandem Adh promoters (Yokomori et al., 1993a; Hansen and Tjian, 1995). Perhaps partly for these reasons, there is considerable evidence that TFIIA is also involved in cooperating with transcriptional activators to stimulate initiation of transcription. TFIIA participates in the rapid formation of an open promoter complex promoted by the activator protein GAL4-AH (Wang et al., 1992). Moreover, antibodies against TFIIAα can block activation by GAL4-VP16 and GAL4-Pro *in vitro* (Ma et al., 1993). Consistent with this, human TFIIA was shown to stimulate activation of transcription *in vitro* as well as *in vivo* in transfected HeLa cells (Sun et al., 1993). Furthermore, *Drosophila* TFIIA can stimulate transcriptional activation by VP16, Sp1 and NTF1 (Yokomori et al., 1994). Since human TFIIA was found to interact directly with the activator Zta to stimulate TFIIID binding to
the TATA region and downstream promoter sequences (Ozer et al., 1994), TFIIA may often help activators to stimulate formation of a preinitiation complex containing TFIID.

**TFIIJ**

TFIIJ has not yet been purified to homogeneity. It is an essential factor for transcription when native TFIID is replaced by TBP in vitro in a reconstituted transcription system (Cortes et al., 1992) and may therefore be one of the TAFs associated with TBP. TFIIJ can be assembled into the preinitiation complex after TFIIH. TFIIJ has not been completely purified and its role in transcription remains undetermined.

**RNA polymerase II and its holoenzyme**

RNA polymerase II is responsible for the transcription of eukaryotic protein-coding genes and some genes encoding small nuclear RNAs (snRNAs). The nuclear RNA polymerases in eukaryotes have three common features: first, they are multisubunit enzymes consisting of 10-12 subunits; second, some subunits of RNA polymerase II are also components of RNA polymerase I and RNA polymerase III; finally, the two large subunits of 220 and 140 kDa are different in amino acid sequence, but conserved in structure among all three nuclear RNA polymerases (Allison et al., 1985; Sawadogo and Sentenac, 1990; Kolodziej et al., 1990; Young, 1991; Woychik et al., 1990).

The genes encoding most yeast RNA polymerase II subunits have been cloned, but the activity of the enzyme has not yet been reconstituted with cloned subunits. The genes encoding the largest subunits of RNA polymerase II from yeast (Allison et al., 1985), *Drosophila* (Biggs et al., 1985; Allison et al., 1988; Zehring et al., 1988), mouse (Ahearn et al., 1987; Corden et al., 1985), and hamster (Allison et al., 1988) have been cloned and sequenced. Comparison of the sequences has revealed a unique feature present at the carboxy-terminus. The carboxy-terminal domain (CTD) consists of multiple, sometimes degenerate repeats of the consensus sequence Tyr-Ser-Pro-Thr-Ser-Pro-Ser. The number of repeats varies among species and is 26 in *S. cerevisiae* and 52 in humans.
The CTD has been proposed to be involved in transcriptional activation, elongation and displacement of histones or other DNA binding proteins. However, all these models have been challenged by conflicting experimental results. Zehring et al. (1988) demonstrated that the Drosophila CTD is essential in vivo but is not always required for accurate transcription initiation in vitro. Although the CTD is essential for the transcription of the mouse dihydrofolate reductase promoter in vitro (Thompson et al., 1989), it is not required for transcription of other promoters, such as the adenovirus major late and Drosophila actin 5C, histone H3, and histone H4 promoters (Zehring et al., 1988; Kim and Dahmus, 1989). Whereas complete deletion of the CTD results in inviability in yeast (Allison et al., 1988; Nonet et al., 1987), partial deletions of the CTD lead to conditional growth phenotypes and defects in gene expression (Nonet et al., 1987; Bartolomei et al., 1988). Some contradictions in the requirement for the CTD for transcription in vitro may be due to the transcription system used, since dependence on the yeast CTD can be reconstituted by adding two fractions, one negative-acting and the other positive-acting, in the basal factor system (Li and Kornberg, 1994). Recent studies on the mouse RNA polymerase II CTD showed that the CTD is required for transcriptional activation by some activators like VP16, but is not required for transcriptional activation by Sp1 (Liao et al., 1991; Gerber et al., 1995). Peterson et al. (1991) reported that the CTD binds nonspecifically to DNA with an affinity of $10^6$ M$^{-1}$ and might be involved in removing SIN1, a yeast homologue of the chromatin-associated HMG1 from higher eukaryotes, as suggested by genetic evidence that deletion of the SIN1 gene suppresses the transcriptional and growth defects caused by a CTD truncation in yeast. However, consistent with the idea that the CTD participates directly in activation, RNA polymerase II CTD truncation mutations were found to affect the transcription activated by GAL4-VP16 in vitro with templates that are not assembled into nucleosomes as well as those that are assembled into nucleosomes (Liao et al., 1991).

Two forms of RNA polymerase II, designated IIA and IIB, can be distinguished based on the level of phosphorylation of the largest subunit (Kim and Dahmus, 1989; Cadena and Dahmus, 1987). The phosphorylation sites reside in the CTD in RNA polymerase IIB, which is
the form of the enzyme that participates in chain elongation (Laybourn and Dahmus, 1990). The phosphorylation of the CTD has been proposed to serve as a trigger for switching transcription from initiation to elongation (Laybourn and Dahmus, 1990). Lu et al. (1991) showed that the nonphosphorylated form of RNA polymerase II was bound preferentially to the preinitiation complex. Usheva et al. (1992) demonstrated that a non-phosphorylated CTD on RNA polymerase II could directly interact with TBP, while a phosphorylated CTD could not. They proposed a model in which phosphorylation of the CTD results in dissociation of the interaction between TBP and the CTD so that RNA polymerase II and associated factors can start elongation (Usheva et al., 1992). Using an in vivo ultraviolet crosslinking technique and antibodies directed specifically against the CTD of IIA or the CTD of IIO, O'Brien et al. (1994) demonstrated that the paused polymerases at the 5' ends of many Drosophila genes have an unphosphorylated CTD, whereas most of the elongating polymerases on the fully heat-induced hsp70 and hsp83 genes contain either a partially phosphorylated CTD or hyper-phosphorylated CTD. This suggests that the CTD is normally phosphorylated during or after initiation in response to particular activators like the heat shock factor, HSF.

Because of these observations, many researchers have tried to identify biologically important CTD kinases. Firstly, the CTD was demonstrated to be phosphorylated by the MO15 (CDK7) in the general factor TFIIH after entering the preinitiation complex (Lu et al., 1992). Other cellular kinases from yeast, Aspergillus, mouse and human cells have also been reported to be capable of phosphorylating the CTD. These kinases include cdc2 kinase (Corden and Ingles, 1992), mammalian c-abl kinase (Baskaran et al., 1993), yeast SRB10/11 kinase (Liao et al., 1995) and the DNA-dependent kinase in mammalian cells (Dvir et al., 1992). Greenleaf and his colleagues identified a yeast CTD kinase composed of 58, 38 and 32 kDa subunits (Lee and Greenleaf, 1991). The 58 kDa subunit has a high degree similarity to the cdc2/CDC28 family of protein kinases. The CTD may indeed be a substrate for several protein kinases but not enough has been done to test which kinases actually phosphorylate the CTD in vivo.
RNA polymerase II molecules can be separated from the general transcription factors during chromatographic purification. On the other hand, RNA polymerase II can interact with several general initiation factors, as well as with the elongation factor TFIIS, and growing evidence supports the idea that cells contain one or more forms of RNA polymerase II holoenzyme in which the enzyme is tightly associated with many other proteins, including several general initiation factors. Initial clues were based on genetic evidence produced by members of the Young laboratory, who selected for suppressor mutations in yeast strains bearing a partially deleted CTD and having conditional growth phenotypes. Mutations in a total of nine genes encoding suppressor proteins SRB2, SRB4, SRB5, SRB6, SRB7, SRB8, SRB9, SRB10 and SRB11 were able to suppress the defects of yeast cells containing a CTD truncation. Moreover, strains deleted for the SRB2 gene exhibit the same set of conditional growth phenotypes and very similar defects in gene expression as the mutant with a CTD truncation. This suggested that SRB proteins might interact with the CTD (Nonet and Young, 1989, Koleske et al., 1992; Thompson et al., 1993; Hengartner et al., 1995). Indeed, SRB2 was found to directly interact with TBP and to be essential for efficient assembly of the preinitiation complex in vitro (Koleske et al., 1992). Moreover, all the SRB proteins were found to be tightly associated in a complex containing a portion of RNA polymerase II as well as the general transcription factors TFIIF, TFIH and TFIIB and other unidentified polypeptides (Thompson et al., 1993; Koleske and Young, 1994; Hengartner et al., 1995). Although this form of the RNA polymerase II holoenzyme seems to include only 2% of the total RNA polymerase II in a yeast cell (Koleske and Young, 1994), experiments with yeast SRB mutants have indicated that virtually all of the transcription in vivo involves the SRB proteins (Thompson and Young, 1995).

Kim et al. (1994) purified an RNA polymerase II holoenzyme that seemed to comprise at least half of the RNA polymerase II in a yeast cell. In this case, purification of the holoenzyme was based on its co-purification with a mediator required for transcriptional activation. Their form of the holoenzyme contains this multiprotein mediator and core RNA polymerase II (Kim et
al., 1994). The mediator can be dissociated from the RNA polymerase II holoenzyme by using a column containing an immobilized monoclonal antibody directed against the CTD. The mediator consists of some 20 polypeptides, including TFIIF and the GAL11, SUG1, SRB2, SRB3, SRB4, SRB5, SRB6, SRB7, SRB8, SRB9, SRB10 and SRB11 proteins (Kim et al., 1994; Hengartner et al., 1995). Genetic experiments have previously implicated GAL11 and SUG1 in transcriptional activation by GAL4 and other yeast activators (Swaffield et al., 1995). However, cells lacking SRB10 and SRB11 exhibit a set of phenotypes common to cells defective in SSN6/TUP1-mediated transcriptional repression (Wahi and Johnson, 1995). Therefore, the mediator in RNA polymerase II holoenzyme may act as a signal transducer for both activation and repression.
B. Elongation factors

Transcription of some proto-oncogenes and viral genes, including HIV-1, have been found to be regulated at the level of transcriptional elongation (Miller et al., 1989; Wright et al., 1989; Kao et al., 1987). While much has been learned about regulation of the initiation of transcription by RNA polymerase II, much less is known about regulation of transcriptional elongation by RNA polymerase II. After RNA polymerase II escapes from the promoter, it continues to catalyze the processive addition of ribonucleotides to the 3' end of the elongating RNA chain until specific attenuation or termination signals are encountered. However, growing evidence indicates that RNA polymerase II molecules are often arrested shortly after initiating transcription of many genes and produce abortive transcripts. For example, RNA polymerase II pauses just downstream of a heat-shock promoter after synthesizing a short nascent RNA (Rougvie and Lis, 1990). By using various assay systems in vitro, several factors from mammalian, Drosophila and yeast cells have been identified as transcriptional elongation factors, including TFIIF (described above), TFIIS, TFII X, SII and YES.

TFIIS

TFIIS was first discovered by Natori and co-workers as a factor from Ehrlich ascites tumor cells that stimulated nonspecific transcription by RNA polymerase II (Sekimizu et al., 1976). It was later shown that human TFIIS acts subsequent to the initiation step to stimulate the rate of elongation (Reinberg and Roeder, 1987). TFIIS proteins have been purified from a variety of organisms including yeast (Sawadogo et al., 1981), Drosophila (Sluder et al., 1988), mouse (Sekimizu et al., 1976), calf thymus (Rappaport et al., 1987) and humans (Reinberg and Roeder, 1987), and their genes have been cloned. The human TFIIS protein is composed of 280 amino acids. Human TFIIS was found to share 92% identity in the DNA sequence of its open reading frame with mouse TFIIS and 40% identity in its deduced amino acid sequence with yeast TFIIS, also called PPR2 (Clark et al., 1991). Sequence similarity was also observed between the mouse and yeast TFIIS proteins as well as Drosophila and mouse TFIIS proteins. These sequence similarities of TFIIS proteins among various species indicate that TFIIS plays an
important role in the cell. Carboxy-terminal regions of TFIIS are particularly well conserved, indicating the functional significance of this region. It was shown that the carboxy-terminal region of TFIIS contains at least two domains: one for binding to RNA polymerase II and the other for binding to nucleic acid and stimulating transcriptional elongation (Agarwal et al., 1991).

TFIIS is not essential for yeast growth since deletion of the PPR2 gene is not lethal (Clark et al., 1991). This may imply that yeast contain a related gene that substitutes for PPR2. However, deletion of the PPR2 gene from the yeast genome causes cells to be sensitive to the uracil analog 6-azauricil (6AU) (Hubert et al., 1983). This sensitivity to 6AU can also be conferred by mutations in the gene encoding the largest subunit of RNA polymerase II (Archambault et al., 1992). Moreover, these mutations can be suppressed by overexpression of TFIIS (Archambault et al., 1992). These results suggest that TFIIS is functionally associated with RNA polymerase II and that the largest subunit of RNA polymerase II is involved in the process of transcriptional elongation. Under nearly physiological conditions, TFIIS enables polymerase to read efficiently through intrinsic arrest sites by interacting with RNA polymerase II (Horikoshi et al., 1984; Rappaport et al., 1987; Reinberg et al., 1987). In the presence of TFIIS, arrested RNA polymerase II complexes cleave their nascent RNA transcripts in a 3' to 5' direction (Izban and Luse, 1992; Reines, 1992). Since highly purified mammalian RNA polymerase II ternary complexes without TFIIS still exhibit a low level of cleavage activity, it was proposed that TFIIS may stimulate a nuclease activity in RNA polymerase II. Studies on the structure of the nucleic acid binding domain of TFIIS by NMR revealed that it contains a three-stranded \( \beta \)-sheet and a disordered loop to form a zinc ribbon (Qian et al., 1993). Amino acid substitution of either residue Asp-261 or Glu-262 within the zinc ribbon eliminates the TFIIS activity, suggesting these residues play a critical role in stimulation of elongation and RNA cleavage and possibly participate in a two-metal-ion phosphoryltransfer reaction catalyzed by RNA polymerase II (Jeon et al., 1994).

SIII
Transcription factor SIII, also termed Elongin, was identified as a novel transcription factor in rat liver extracts stimulating synthesis of accurately initiated transcripts by RNA polymerase II. This factor is composed of three subunits of 110, 18, and 15 kDa (Elongin A, B and C) with a native molecular mass of approximately 140 kDa (Bradsher et al., 1993a). SIII was initially discovered as an activity that was required for synthesis of full length transcripts in a highly purified reconstituted in vitro transcription system in which yeast TBP was substituted for rat TFIID and dATP or AMP-PNP was used to replace ATP in the reactions. The stimulation of transcription by SIII was observed when any of the ribonucleoside triphosphates were present at limiting concentrations, indicating that it acts at the stage of elongation of transcription (Bradsher et al., 1993b). SIII does not affect initiation or the formation of the first few phosphodiester bonds. In a pulse-chase analysis, general transcription factors and RNA polymerase II are sufficient to synthesize short transcripts, but not full-length transcripts, in the absence of SIII (Bradsher et al., 1993b). Moreover, SIII is able to stimulate transcription after initiation occurs. Therefore, SIII may directly act on the elongation complex (Bradsher et al., 1993b).

The cDNA encoding the 15 kDa subunit of SIII (Elongin C) has been cloned and sequenced (Garrett et al., 1994). The p15 gene encodes a polypeptide of 112 amino acids with a predicted isoelectric point of 4.6. Analysis of its deduced amino acid sequence revealed that p15 shares significant sequence similarity with a 60 amino acid region of the E. coli transcription termination factor ρ and a 19 amino acid stretch that is similar to a sequence in the E. coli antitermination factor NusB. Bacterially expressed recombinant p15 was demonstrated to be able to stimulate elongation along with the native p110 and p18 subunits of SIII (Garrett et al., 1994). Elongin A is the transcriptionally active component of Elongin. The cDNA encoding Elongin A was also cloned recently and its deduced amino acid sequence revealed a 53% similarity with TFIIS at its amino-terminal region (Aso et al., 1995). The cDNA for Elongin B has also been cloned. The amino-terminal portion (84 amino acids) of Elongin B contains striking sequence similarity to ubiquitin, suggesting that Elongin B may be a member of the ubiquitin homology
gene family (Garrett et al., 1995). Unlike Elongin C, it acts like a chaperone to promote the assembly and thermostability of the Elongin complex.

Elongin was recently found to be the target of the von Hippel-Lindau (VHL) tumor suppressor protein (Duan et al., 1995; Kibel et al., 1995). The VHL protein binds to Elongin B and C and inhibits the transcriptional activity of Elongin, indicating that VHL protein is a regulatory factor of transcriptional elongation.

**TFIIX**

TFIIX was first described as an activity required for transcription in an *in vitro* reconstituted system, in which a DNA template containing Ad2-MLP downstream sequences between nucleotides +33 to +536 and a heparin-Ultrogel TFIID fraction were used (Reinberg et al., 1987). It was further shown that TFIIX enables RNA polymerase II complexes to elongate after they are blocked in transcription elongation at the Minute virus of mice attenuator (Bengal et al., 1991) or at the adenovirus major late transcriptional unit (Krauskopf et al., 1991). The polypeptide composition of TFIIX has not been identified yet.

**YES**

RNA polymerase II can initiate and transcribe without general initiation factors on double-stranded DNA templates with single-stranded poly-deoxycytosine(polydC) tails (Chafin et al., 1991). A yeast factor, termed YES, can specifically stimulate the rate of transcriptional elongation by yeast and *Drosophila* RNA polymerase II but has no effect on transcription by RNA polymerase I and RNA polymerase III. The effect of YES on elongation was shown to be different from RNase H (Chafin et al., 1991). The activity of YES was further purified and corresponded to a 116-kDa protein. The gene encoding it has not yet been identified and cloned.

**P-TEF**

Using *Drosophila* KcN extract, Price and his colleagues have also studied the activity of elongation complexes in vitro. They observed that only a fraction of RNA polymerase II molecules can generate transcripts longer than several hundred nucleotides (Kephart et al., 1992). The ability of this fraction of RNA polymerase II to synthesize long transcripts was
C. Transcriptional activators

Regulation of transcription is largely mediated by various transcriptional activators that bind to specific DNA sequence elements that are often located hundreds or even thousands of base pairs upstream of the TATA-box. Such a gene-specific DNA sequence element is called an upstream activation sequence (UAS) or enhancer (Banerji et al., 1981; Guarente et al., 1982). Transcriptional activators generally contain two separate domains: one is involved in site-specific DNA binding or interacts with another DNA-binding protein, and the other is needed for transcriptional activation. Based on the amino acid compositions of their transcriptional activation domains, activators have been classified into at least three categories: acidic residue-rich activators, glutamine-rich activators, and proline-rich activators (Mitchell and Tjian, 1989). There are other activation domains which do not fall into any of these categories and have not yet been classified.

Acidic activators

Acidic activation domains were the first ones to be identified. Studies of the activation domains of GAL4 and GCN4, two yeast transcriptional activators, revealed two common features: richness in acidic residues and putative amphipathic \( \alpha \)-helical structures (Ma and Ptashne, 1987; Hope and Struhl, 1986). However, the importance of acidic residues and \( \alpha \)-helical structures for transcriptional activation is challenged by both biochemical and genetic studies (Hahn, 1993). For example, mutagenic analysis of the C-terminal activation domain of GAL4 suggested that its acidic amino acids are not required for activation. Furthermore, this activation domain is not unstructured or \( \alpha \) helical, but may form a \( \beta \) sheet (Leuther et al., 1993). This is further supported by circular dichroism spectroscopy experiments which showed that the acidic activation domains of the GCN4 and GAL4 proteins are not \( \alpha \) helical, but form \( \beta \) sheets in slightly acidic solution (Van Hoy et al., 1993). It is not clear whether the \( \beta \) sheet structure represents the active or inactive state for transcriptional activation (Leuther et al., 1993; Van Hoy et al., 1993).
Acidic activation domains have also been found in a number of other transcriptional activators, including the herpes simplex virus transactivator protein VP16 (Triezenberg et al., 1988), the tumor suppressor protein p53 (O'Rourke et al., 1991; Fields and Jang, 1990; Raycroft et al., 1990) and the human estrogen receptor (Tora et al., 1989).

The herpes simplex virus transactivator protein VP16 is responsible for activating the expression of viral immediate early genes (Campbell et al., 1984). VP16 itself does not bind to DNA, but forms a heteromeric complex with the cellular factor HCF and subsequently associates with a cellular protein OCT1 that recognizes an octamer sequence in DNA (Stern et al., 1989; Kristie and Sharp, 1990; Xiao and Capone, 1990; Stern and Herr, 1991). Dissection of VP16 revealed that its amino-terminal region appears to interact with cellular proteins and its carboxy-terminal 78 amino acids is a transcriptional activation domain. The 78 amino acid activation domain functions in vivo and in vitro when fused to the DNA-binding domain of the yeast activator GAL4 (Sadowski et al., 1988; Chasman et al., 1989; Carey et al., 1990). Substitution of amino acids within the activation domain of VP16 has provided evidence that negatively charged residues and certain hydrophobic residues are important for activation, but the putative amphipathic α-helix appears to be unimportant for the structure of the activation domain (Cress and Triezenberg, 1991). The 78 amino acid activation domain was further demonstrated to contain two independent activation domains: the region from residues 412 to 456 and the region from residues 452 to 490 (Regier et al., 1993).

Cullen and his colleagues reported that the activation domain of VP16 is able to stimulate transcription when fused to the RNA-binding domain of the HIV-1 Rev protein and targeted to the promoter by a rev response element (RRE) in the transcribed region (Tiley et al., 1992). This finding indicated that VP16 and the human immunodeficiency virus transactivator protein Tat, which naturally binds to HIV TAR RNA, might share similar mechanisms to activate transcription. Yankulov et al. (1994) measured the ability of a set of activation domains to stimulate RNA polymerase II processivity in vivo. They expressed the processivity as the percentage of total transcripts that reads through to the termination site in each gene. Like Tat
protein, the activation domain of VP16 was shown to significantly stimulate the processivity of RNA polymerase II (Yankulov et al., 1994). Similar effects on processivity by GAL4-AH and GAL4-E1a were also observed (Yankulov et al., 1994). Therefore, acidic activators may stimulate transcription at the levels of both initiation and elongation.

The human tumor suppressor p53 gene encodes a protein which functions to suppress oncogenic transformation (Eliyahu et al., 1989; Finlay et al., 1989), negatively regulate cell cycle progression (Baker et al., 1990; Diller et al., 1990; Martinez et al., 1991), and induce apoptosis in certain physiological situations (Yonish-Rouach et al., 1991; Shaw et al., 1992). It was originally identified as a polypeptide that associates with DNA tumor virus transforming proteins, such as SV40 large T-antigen (Lane and Crawford, 1979), the adenovirus E1B protein (Sarnow et al., 1982), and the human Papillomavirus E6 protein (Wemess et al., 1990). The p53 gene is not essential for growth since mice lacking p53 appear to have no developmental defects. However, these mice are very susceptible to the development of malignant tumors. Two groups (Fields and Jang, 1990; Raycroft et al., 1990) demonstrated that wild-type p53 protein acts as a transcriptional activator of transcription of class II genes. Like other transcriptional activators, p53 protein contains two separate functional domains: a site-specific DNA binding domain near its carboxy-terminus (Kern et al., 1991) and a transcriptional activation domain within its amino-terminal 42 amino acids (Unger et al., 1992). When the p53 protein or its activation domain is fused to the DNA binding domain of the yeast activator protein GAL4, the chimeric protein stimulates the transcription of genes bearing GAL4-binding sites adjacent to their promoters in yeast and human cells (Fields and Jang, 1990; O'Rourke et al., 1990; Raycroft et al., 1990). Two oncogene products, MDM2 (Oliner et al., 1993; Chen et al., 1994) and adenovirus 2 early 1B 55 kDa protein (Kao et al., 1990), bind to the amino-terminal portion of p53 and inhibit its transcription activity. The gene encoding MDM2 was originally cloned as a cellular oncogene amplified on a mouse double minute chromosome (Cahilly-Snyder et al., 1987).
Mutations in the gene encoding the tumor suppressor p53 have been found in many human tumors (Baker et al., 1989; Nigro et al., 1989; Kern et al., 1992). Most mutations result in loss of p53 ability to activate the transcription of genes that contain p53 binding sites in their promoters. It has been demonstrated that this effect is most often due to the elimination of the DNA-binding activity of p53 (Kern et al., 1992; Unger et al., 1992). Probably because p53 can form multimers (Stenger et al., 1992; McCormick et al., 1981; Kraises et al., 1988; O'Reilly and Miller, 1988), many of the mutant p53 proteins act biologically as dominant negatives to inhibit the activity of wild type p53 (Raycroft et al., 1990; Kern et al., 1992; Unger et al., 1992).

Although most p53 mutations occur outside of the activation domain, some point mutations can nevertheless alter the folding of p53 to prevent the activation domain from functioning (Raycroft et al., 1990). Mutations in the activation domain can also affect activation by p53. Lin et al. (1994) carried out an extensive genetic analysis of the activation domain of p53 by using site-directed mutagenesis. The results indicated that hydrophobic residues Leu22 and Trp23 are critical for the transactivation activity of p53, but acidic residues are not essential for its transcription activity. Crook et al. (1994) examined a number of naturally occurring point mutations detected in tumors as well as a series of deletion and truncation mutations for their activities in transactivation, growth suppression, and transformation. They concluded that transcriptional activation by p53 correlates well with suppression of growth but not with transformation.

**Glutamine-rich activators**

Analysis of the transcriptional activator Sp1 resulted in the identification of glutamine-rich activation domains. The two most potent activation domains in Sp1 appear to have no obvious similarity to any other known protein and contain about 25% glutamine residues (Courey and Tjian, 1988). This type of glutamine-rich motif was also observed in other transcriptional activators, such as the Drosophila Antennapedia, Ultrabithorax, and Zeste proteins (Courey and Tjian, 1988), yeast HAP1, HAP2, and GAL11 (Pfeifer et al., 1988, Pinkham et al., 1987, Suzuki et al., 1988), and mammalian OCT-1 (Stern et al., 1989) and OCT-2 (Ko et al., 1988).
Proline-rich activators

Transcription factor CTF/NF1 contains a proline-rich (20 to 30%) activation domain in its carboxy-terminal region (Mermod et al., 1989). Proline-rich motifs have also been identified in some other mammalian transcription activators, such as SRF (Norman et al., 1987) and AP-2 (Williams et al., 1988). The activation domain of CTF/NF1 functions in yeast when fused to the yeast GAL4 DNA binding domain (Kim and Roeder, 1993). Xiao et al. (1994) demonstrated that the activation domain of CTF/NF1 contains a CTD-like motif which is critical for its transcription activity. When this motif is replaced with CTD repeats from RNA polymerase II, CTF increases its activation activity when there are more CTD repeats (Xiao et al., 1994).
D. Mechanisms of transcriptional activation

How transcriptional activators stimulate transcription has been a long standing question. A variety of models have been proposed to explain how transcription is activated. In bacteria, it has been demonstrated that transcriptional activators can directly interact with the $\alpha$ subunit of RNA polymerase to stimulate formation of a closed or open complex (Hochschild et al., 1983; Igarashi and Ishihama, 1991). The fact that acidic activation domains from yeast, Drosophila and mammalian transcriptional activators can function in all eukaryotic organisms strongly implies that a common mechanism for transcriptional activation by acidic activation domains must exist in all eukaryotes. The observations that activation domains can function when attached to the DNA-binding domain of a heterologous protein (Brent and Ptashne, 1985) and that activator proteins can control transcription from a promoter when they are bound to DNA hundreds or even thousands of base pairs away led to a hypothesis in which activators bound to a UAS or enhancer may interact with or modulate the preinitiation complex by looping out of the DNA between the enhancer and the proximal promoter (Ptashne, 1988). This interaction would then result in transcriptional activation. One important question then arises: which factors within the basal transcription machinery are direct targets for transcriptional activators?

**Direct targets in the basal transcription machinery**

The large subunit of RNA polymerase II has been proposed to be the target of transcriptional activators. Allison and Ingles (1989) reported that an increase in the length of the RNA polymerase II CTD suppressed the transcriptional defects of mutations in the yeast activator GAL4. This result indicated that the RNA polymerase II large subunit CTD may directly or indirectly contact activators. Brandl and Struhl (1989) showed that GCN4 can directly bind to RNA polymerase II by performing affinity chromatography experiments with immobilized GCN4 protein. The portion of GCN4 that contacted RNA polymerase II was not located in its acidic activation domain, but instead resided in its DNA-binding domain, and thus the interaction...
may not be meaningful. To date, no transcriptional activator has been shown to directly contact the CTD in RNA polymerase II.

Because TFIID is the first basal transcription factor recruited into the initiation complex and plays an important role in initiation of transcription, it has also been proposed to be the target of activators. DNase I footprint analysis suggested that TFIID might be a direct target for the activators USF, ATF and GAL4, since these activators can enhance protection by partially purified native TFIID of a region downstream from the TATA-box (Sawadogo and Roeder, 1985; Horikoshi et al., 1988a, 1988b). Using affinity chromatography with columns containing an immobilized protein A-VP16 chimeric protein, Stringer et al. (1990) examined whether the activation domain of VP16 could bind directly and selectively to RNA polymerase II or one of its general initiation factors. They found that an activity eluted from the VP16 affinity column could restore transcription activity to a heated HeLa nuclear extract that was lacking TFIID activity. The activation domain of VP16 also selectively and directly bound yeast TBP that had been expressed in bacteria. The binding of TBP to various mutant VP16 activation domains in vitro correlated well with their abilities to activate transcription in vivo (Ingles et al., 1991). Kim et al. (1994) demonstrated that a single amino acid substitution in TBP, L114K, reduces its binding to VP16 and also abolishes its ability to support activated transcription, but not basal transcription. These results suggested that TBP is a functional target of the acidic activator VP16.

TBP has been demonstrated to directly interact with a number of transactivators, including the acidic activators p53 (Seto et al., 1992; Truant et al., 1993; Ragimov et al., 1993; Martin et al., 1993; Liu et al., 1993; Chen et al., 1993), c-Myc (Hateboer et al., 1993), v-Rel and c-Rel (Kerr et al., 1993; Xu et al., 1993), and E2F-1 (Hagemeier et al., 1993b). An oncogenic mutation which is not in the activation domain of p53 was reported to reduce binding of p53 to TBP (Truant et al., 1993). Using a gel mobility shift assay, Chen et al. (1993) demonstrated that p53 and TBP cooperate with each other to bind a DNA element containing specific p53-binding sites and a TATA-box. This result suggested that binding of p53 to TBP may lead to a more stable p53 -TFIID-promoter complex and subsequently activate transcription.
Two groups have independently determined that the adenovirus large E1A protein, which contains a potent acidic activation domain, can also interact with human TBP (Horikoshi et al., 1991; Lee et al., 1991b). Lee et al. (1991b) performed far-western protein blotting and co-immunoprecipitation experiments. They demonstrated that E1A directly contacts a 51 residue region in the conserved carboxy-terminal domain of human TBP. Mutational analysis suggested that this interaction is necessary, but not sufficient, for the transactivation activity of E1A (Lee et al., 1991b).

Other types of activation domains, including glutamine-rich domains, proline-rich domains and other unclassified activation domains, have also been demonstrated to directly interact with TBP. These include the Epstein-Barr virus proteins Zta (Lieberman and Berk, 1991) and R (Manet et al., 1993), the human T-cell leukemia virus type 1 activator Tax1 (Caron et al., 1993), the human immunodeficiency virus type 1 transactivator Tat (Kashanchi et al., 1994), and human c-fos, c-jun (Ransone et al., 1993), PU-1 (Hagemeier et al., 1993a) and Sp1 (Emili et al., 1994). A particularly interesting case is the proline-rich activator CTF, which binds to TBP via the CTD-like repeat in its proline-rich domain (Xiao et al., 1994).

How does the activator-TBP interaction lead to transcriptional activation? Binding of TBP to the TATA-box element is a rate limiting step in initiation and is strongly inhibited by repressors such as nucleosomes (Workman and Roeder, 1987; Klein and Struhl, 1994). Several experiments suggest that the interaction between activator and TBP may increase recruitment of TBP or TFIIB to the promoter. The Epstein-Barr virus protein Zta binds directly to TBP and significantly reduces the rate of dissociation of TBP from the TATA-box element in vitro (Lieberman and Berk, 1991). Kim et al. (1994) demonstrated that a TBP mutation that prevented binding to the activation domain of VP16 prevented VP16 from recruiting TFIIB into the initiation complex in vitro. This result implied that the VP16-TBP interaction may actually help the entry of TFIIB into the initiation complex. To assess the biological relevance of activator-TBP interactions that are observed in vitro for transcriptional activation in vivo, Klages and Strubin (1995) used two hybrid proteins to show that a direct interaction of TBP with an
activator is sufficient to increase initiation of transcription by RNA polymerase II in yeast cells. In similar experiments, Chatterjee and Struhl (1995) demonstrated that *E. coli* LexA-TBP, a chimeric protein containing the DNA binding domain of LexA, efficiently activates transcription from a promoter bearing a LexA operator upstream of a TATA element. These experiments imply that recruitment of TBP to a promoter is sufficient for activation.

TFIIB is another basal factor that has been demonstrated to interact directly with transcriptional activators. Lin and Green (1991) showed that the activation domain of VP16 could also bind TFIIB activity in addition to TFIID. An amino acid substitution (FP442) in the activation domain of VP16, which reduces the activity of VP16 *in vivo* (Cress and Triezenberg, 1991), may also reduce the binding of TFIIB (Lin et al., 1991) to VP16 *in vitro*, although this observation is controversial (Goodrich and Tijian, 1994; Walker et al., 1993). Mutant TFIIB proteins that are unable to bind to VP16 can support basal transcription but not VP16-activated transcription (Roberts et al., 1993). The interaction of TFIIB with the activation domain of VP16 is direct since VP16 can bind recombinant TFIIB present in bacterial extracts (Lin et al., 1991). Using both an immobilized DNA template assay and a gel-filtration assay, it was shown that certain activators, including GAL4-AH, GAL4-VP16, GAL4-Pro, and GAL4-Gln, could enhance the entry of TFIIB to the preinitiation complex when TBP was bound to the promoter (Lin and Green, 1991; Choy and Green, 1993). GAL4-AH also increased the entry of other basal factors and RNA polymerase II when TFIID was present (Choy and Green, 1993). The enhancement of the binding of other basal factors and RNA polymerase II requires the binding of activators to multiple binding sites (Choy and Green, 1993) and may result from the disruption of an intramolecular interaction between the amino- and carboxy-terminal domains of TFIIB by the acidic activator (Roberts and Green, 1994). The recruitment of TFIIB by activators may be required also for each new round of transcription since some biochemical evidence indicates that all basal transcription factors and RNA polymerase II, except TFIID, were dissociated from the promoter when transcription was initiated (Zawel et al., 1995).
Other transcriptional activators have also been found to interact directly with TFIIB.
These include the glutamine-rich Drosophila activator Ftz and certain members of the superfamily
of nuclear receptors, such as the estrogen receptor, thyroid hormone receptor, chicken ovalbumin
upstream promoter transcription factor and vitamin D receptor. The Drosophila transcription
factor Krüppel (Kr), an essential factor in organogenesis during later embryonic development,
was also found to interact with TFIIB (Sauer et al., 1995). It is of interest that monomeric Kr
activates transcription and specifically binds to TFIIB, whereas Kr dimers repress transcription
and bind to TFIIEβ (Sauer et al., 1995).

Based on studies on SRF-activated transcription, Zhu et al (1994) proposed that
transcriptional activation by SRF, and perhaps some other activators, involves their interaction
with TFIIF. SRF is an essential factor for serum and growth factor induction of the c-fos gene.
It binds to the serum response element in the c-fos gene and contains an activation domain that
has no homology with other known activation domains (Treisman and Ammerer, 1992; Prywes
and Zhu, 1992). It was shown that transcriptional activation by SRF and GAL4-VP16 required
a high concentration of TFIIF, whereas transcriptional activation by Sp1 did not. Using a gel
mobility assay and the yeast two hybrid system, SRF was demonstrated to specifically interact
with RAP74, the large subunit of TFIIF (Zhu et al., 1994; Joliot et al., 1995). Moreover, a
portion of RAP74 that interacts with SRF is not required for basal transcription, but is required
for activation by SRF.

Coactivators, adaptors and mediators

Schmidt et al. (1989b) investigated the interaction between the glutamine-rich activator
Sp1 and TFIID. They found that Sp1 increased the number of productive transcription
complexes, even though the binding of TFIID to TATA-box DNA was not stimulated. Sp1 did
not change the extent of the footprint of yeast TBP, and they suggested that Sp1 does not directly
contact TBP. This speculation was further investigated by Pugh and Tjian (1990). They
showed that Sp1 only activates transcription in a reconstituted reaction in which native TFIID is
used instead of TBP. The native TFIID contained co-activators, now known as TAFs, that are
species-specific and are associated with TBP. Pugh and Tjian (1990) proposed a model in which Sp1 directly contacts co-activators to stimulate transcription. Direct interaction of Sp1 and TAF110 has indeed been observed (Hoey et al., 1993). The biological relevance of this interaction was further supported by the finding that mutations of Sp1 that impair the transcription activity of Sp1 weaken the binding of Sp1 to TAF110. TAFs were also required for transcriptional activation by the Drosophila activator NTF-1 (Tanese et al., 1991). Berk and colleagues (Zhou et al., 1992) showed that a "holo-TFIID" supports transcriptional activation by activation domains of several activators, including GAL4-AH, GAL4-E1A, Zta, and Sp1, and even supports transcription from a TATA-less promoter. This "holo-TFIID" was said to consist of TBP and five TAFs of 250, 125, 95, 78 and 50 kDa. These results also suggested that TAFs act as coactivators for transcriptional activators. Drosophila TFIID has now been reconstituted from recombinant TBP and recombinant TAFs produced in a baculovirus expression system (Chen et al., 1994). Since several activators have been found to function in a coactivator-specific manner with recombinant TFIID, TAFs may act as gene-specific coactivators (Chen et al., 1994; Thut et al., 1995).

Other studies on a phenomenon called "squelching" in yeast and HeLa cells indicated that a mediator or adaptor might act as a bridge between acidic activation domains and basal transcription factors (Berger et al., 1990; Kelleher et al., 1990; White et al., 1991). A fraction containing such a mediator was partially purified from yeast cell extract. It did not replace any factor fraction required to reconstitute basal transcription, but it was required for activated transcription by acidic activators (Flanagan et al., 1991). As mentioned above, Kim et al. (1994) further purified the mediator and found that it consists of 20 polypeptides, including the TFIIF, GAL11, SUG1, SRB2, SRB4, SRB5, and SRB6 proteins and other as yet unidentified polypeptides. More recently, the mediator was shown to contain all the genetically identified SRB proteins (Hengartner et al., 1995). The mediator not only supported activated transcription but also stimulated basal transcription. It was also shown that the mediator increased the efficiency of phosphorylation of RNA polymerase II by the TFIIH-associated CTD kinase,
suggesting that there is an interaction between the mediator and the CTD (Kim et al., 1994). The mediator may directly interact with the activation domain of VP16 (Hengartner et al., 1995). Genetic evidence has suggested that GAL11 may interact with certain activators (Nishizawa et al., 1990). Barberis et al. (1995) demonstrated that a point mutation in GAL11 creates a new interacting site for the DNA-binding domain of GAL4, which is devoid of any activating function, and enables it to strongly activate transcription. Thus, it was proposed that a single activator-holoenzyme contact can trigger gene activation simply by recruiting the RNA polymerase II holoenzyme to DNA (Barberis et al., 1995; Hengartner et al., 1995).

In order to identify an adaptor for VP16, Guarente and his colleagues developed a genetic approach to select yeast mutants that can tolerate the toxicity resulting from overexpression of GAL4-VP16. These mutant strains contain mutations in the genes ADA1, ADA2 and ADA3 as well as two other genes, one of which is identical to GCN5 (Marcus et al., 1994). Both the ada2 and ada3 mutants are temperature sensitive. The ada1 mutation reduced the synthesis of GAL4-VP16 and inhibited transcription from several yeast promoters. The ada2 mutation reduced the transcriptional activity of a subset of activators. The ADA2 gene was isolated and found to encode a novel protein which appears to mediate transcriptional activation by GAL-VP16 and GCN4, but not by HAP4 and GAL4 in vivo in yeast (Berger et al., 1992). Thus, the adaptors may act as specific factors for various kinds of activators. However, ADA2 has also been shown to interact with the activation domain of GAL4 (Melcher and Johnston, 1995).

A protein fraction called USA (Meisterernst et al., 1991), that is apparently distinct from the TAFs, not only increases the induced level of transcription in the presence of activators, but also influences the basal level of transcription in the absence of activators. USA appears to be a mixture of two kinds of factors, one of which is a collection of positively acting transcription cofactors (PC1, PC2, PC3 and PC4), and the other of which, known as NC1, interacts with TBP and inhibits basal transcription. The studies on USA have suggested that it is a general coactivator for activated transcription. PC1 has not been further purified and characterized yet. PC2 and PC3 act as positive factors in the presence of activator and as negative factors in the
absence of activator (Kretzschmar et al., 1993; Merino et al., 1993). Purification of the dominant activity in the cofactor fraction has resulted in the identification of a 14 kDa DNA-binding protein, termed PC4, which increases transcriptional activation by various activation domains fused to the DNA-binding domain of GAL4 (Ge and Roeder, 1994; Kretzschmar et al., 1994). The cDNA encoding PC4 was isolated and its deduced amino acid sequences contain two serine-rich regions in the amino-terminal half portion which shows a low degree of homology to viral immediate-early regulatory proteins (Ge and Roeder, 1994; Kretzschmar et al., 1994). PC4 appears to be a general cofactor as it enhances activated transcription by all types of activators. PC4 also interacts specifically with different types of activation domains and TFIIB (Ge and Roeder, 1994). This suggests that PC4 may act as an adaptor to bridge between activators and the basal transcription factors.

Repressors and chromatin

It is becoming increasingly clear that repressors play an important role in the functioning of activators. These repressors include histones associated with chromatin (Workman et al., 1991; Croston et al., 1991) and non-histone proteins which inhibit binding of TBP to the promoter either by interacting with TBP or by competing with TFIIB for binding to TBP (Meisterernst and Roeder, 1991; Meisterernst et al., 1991; Inostroza et al., 1992). Several studies suggest that activators can relieve repression of transcription caused by histone H1 or nucleosomes. Workman et al. (1991) reported that many activators, such as the pseudorabies virus IE protein, USF, Sp1, the GAGA factor, and GAL4-VP16 help TFIID to prevent repression by chromatin. The activator may first mediate displacement of nucleosomes and/or of histone H1 and then interact with PCs and general initiation factors to enhance the formation of the preinitiation complex (Workman et al., 1991; Kretzschmar et al., 1994; Kretzschmar et al., 1993). PC2 was identified as a 500-kDa protein complex and a novel cofactor of activated transcription (Kretzschmar et al., 1994). PC2 is unable to replace TFIIB for basal transcription and is not sufficient to support activated transcription by GAL4-AH without the presence of TAFs in the reaction. Histone H1 protein and PC2 can synergistically mediate transcriptional
activation by GAL4-AH. PC3 was purified to a nearly homogenous 100 kDa polypeptide (Kretzschmar et al., 1994). Two internal peptide sequences were found to be identical to amino acid sequences in DNA topoisomerase I, an enzyme known to relax either negatively or positively supercoiled DNA and presumably to be involved in transcriptional elongation. DNA topoisomerase I is the same as Dr2, a protein that was previously identified and characterized by Reinberg and his colleagues (Merino et al., 1993). In their in vitro reconstituted transcription system, Dr2 and ACF, in addition to the general transcription factors and RNA polymerase II, were necessary for high levels of transcription by GAL4-AH. Dr2 is associated with the TFIID complex and inhibits basal transcription through a direct interaction with TBP. The repression of transcription by Dr2 is specific for TATA-containing promoters and can be relieved by TFIIA or GAL4-AH (Merino et al., 1993).

A SWI-SNF complex including the SNF2/SWI2, SNF5, SNF6, SWI1, and SWI3 proteins was found to be required for the expression of a large number of genes in the yeast Saccharomyces cerevisiae (Winston and Carlson, 1992). The transactivation by several activators in vivo and in vitro was also found to be dependent on the SWI-SNF complex, suggesting that the SWI-SNF complex is a cofactor for those activators (Peterson and Herskowitz, 1992; Peterson and Tamkun, 1995). Genetic evidence indicated that the SWI-SNF complex functions by altering chromatin structure to relieve repression (Hirschhorn et al., 1992; Carlson and Laurent, 1994). The SNF2 protein contains a DNA-dependent ATPase activity required for transcriptional activation (Laurent et al., 1993). The SWI-SNF complex may interact with nucleosomes and facilitate the binding of activators using hydrolysis of ATP, since the purified SWI-SNF complex enhances the binding of GAL-AH to nucleosomal DNA (Cote et al., 1994).

The purified yeast SWI-SNF complex consists of 10 subunits (Cote et al., 1994), whereas the human SWI-SNF complex consists of seven polypeptides with a molecular mass greater than 700 kDa (Kwon et al., 1994). The partially purified human SWI-SNF complex was also shown to be able to mediate the ATP-dependent disruption of a nucleosome and assist GAL4-VP16 and
GAL4-AH to bind within a nucleosome core (Kwon et al., 1994). This partially purified human SWI-SNF complex can also relieve the inhibition of the binding of TBP to a TATA-box DNA containing a specifically positioned nucleosome (Imbalzano et al., 1994).

Taken together, the above results demonstrated that activators increase transcription at several levels, ranging from alterations in chromatin structure to formation of the preinitiation complex by interacting with coactivators and with general transcription factors. The activators may also utilize different targets to regulate transcriptional initiation, promoter clearance, elongation and termination.
E. Regulation of transcription by HIV1 Tat.

Human immunodeficiency virus is the etiologic agent of acquired immunodeficiency syndrome (AIDS) and belongs to the lentivirus family of retroviruses. The HIV1 genome contains the three typical retroviral structural genes, *gag*, *pol* and *env*, which encode capsid proteins, reverse transcriptase, integrase, protease, and external glycoproteins. It also contains regulatory genes which encode six regulatory proteins controlling different aspects of the viral life cycle. Two regulatory proteins, Tat and Rev, are essential for viral replication in all cell types (Dayton et al., 1986; Fisher et al., 1986). The other regulatory proteins are not essential for viral growth and may be involved in efficient production of infectious viral particles or maintenance of the infection *in vivo*.

The life cycle of HIV-1 includes the formation of a provirus through reverse transcription, integration, and transcription from proviral DNA. Transcription of HIV-1 is an essential step for viral replication and is dependent on the transcription machinery of the host cell. The viral genome is transcribed from a single LTR containing the TATA box and binding sites for Sp1 and NF-κB as well as for other cellular activators. The cellular and viral proteins regulate transcription from the viral promoter mainly by controlling premature termination (Cullen, 1993; Jones and Peterlin, 1994). It is generally believed that the elucidation of the molecular mechanisms by which transcription of HIV1 is increased will provide new targets for AIDS therapy.

The viral regulatory protein Tat is expressed early after infection from three different spliced mRNAs. It is required primarily to increase the rate of transcription from the viral long terminal repeat (LTR) during viral growth. Tat can also increase the expression of certain cellular genes, including genes that encode tumor necrosis factors α and β (Sastry et al., 1990; Buonaguro et al., 1992), interleukin 2 (Westendorp et al., 1994) and glucose 6-phosphate dehydrogenase (Ursini et al., 1993). When Tat is released extracellularly by HIV-1 infected cells, it can be taken up by other cells and then enter the nucleus to transactivate genes (Ensoli et al., 1993). Studies have suggested that Tat is a multi-functional protein involved in the
pathogenesis of AIDS in many ways. Li et al. (1995) demonstrated that Tat can induce cell death by apoptosis in a T cell line and in cultured peripheral blood mononuclear cells from uninfected donors.

**Structure of Tat**

HIV-1 Tat ranges in size from 82 to 101 amino acids and has an apparent molecular mass of 16 kDa on an SDS gel (Schwartz et al., 1990). A 28 kDa protein, named Tev, encoded by the first exon of Tat, the middle of Env, and the second exon of Rev, has properties of both Tat and Rev (Benko et al., 1990). Mutational analysis and domain-swapping experiments have suggested that the amino-terminal portion of Tat, including the Cys-rich and core (FITKALGISY) domains of Tat, constitutes an activation domain. The basic domain of Tat, which is located near its carboxy-terminus, contains both a functional nuclear localization signal and an RNA binding motif which specifically recognizes the TAR element (Jones and Peterlin, 1994).

**Tat responsive element (TAR)**

Tat is unique among eukaryotic activators. It generally functions by binding to a RNA element termed the Tat responsive (TAR) element (Rosen et al., 1985) near the 5' ends of the nascent transcripts (Roy et al., 1990; Weeks et al., 1990), but can also function when bound to DNA as a Gal4 fusion protein (Southgate and Green, 1991; Kamine et al., 1991). TAR is present between nucleotides +19 to +44 in all viral mRNAs (Feng and Holland, 1988; Garcia et al., 1989). The HIV-1 TAR contains an upper stem, a trinucleotide bulge, and a structured hexanucleotide loop. Mutations that destabilize the hairpin structure eliminate the ability of the TAR element to respond to Tat protein in vivo (Jakobovits et al., 1988; Feng and Holland, 1988). This defect can be complemented by introducing compensatory mutations that restore the stem-loop structure (Feng and Holland, 1988). These results suggest that the hairpin structure and the sequence of the stem are not important for TAR function. Nevertheless, mutations in the TAR loop sequences do not affect binding of Tat to TAR in vitro. A three-nucleotide bulge and particular nucleotides surrounding the bulge are essential for Tat binding and transactivation (Roy
et al., 1990; Dingwall et al., 1990). The activity of TAR is highly position-dependent, since it becomes inactive when moved 50 to 100 bp distal to its normal location (Muesing et al., 1987; Jakobovits et al., 1988; Selby et al., 1989). TAR function is also orientation-dependent because it functions only in its native orientation. Unlike enhancers, multimerization of the TAR element does not synergistically increase Tat-activated transcription (Selby et al., 1989).

**Cellular TAR-binding proteins and cofactors**

Although Tat functions in many types of human cells, Tat functions poorly in rodent cells (Hart et al., 1989). The activity of Tat in rodent cells can be restored partially by introduction of human chromosome 12 into hamster cell lines (Newstein et al., 1990; Hart et al., 1989), suggesting that a human co-factor encoded by chromosome 12 is required for Tat-dependent activation. However, none of the cloned genes that were proposed to encode putative Tat-cofactors were localized to chromosome 12. In a reconstituted Tat-mediated transactivation system in vitro, a fraction containing a putative cofactor was found to be essential for Tat-activated transcription in reactions containing a set of purified general transcription factors, RNA polymerase II, Tat and Sp1 (Zhou and Sharp, 1995). This cofactor does not affect the basal level of transcription and enables Tat to function without TAFs.

The human-specific co-factor may facilitate the interaction of Tat with TAR because it is not required for activating transcription by a chimaeric Tat fused to the RNA-binding domain of bacteriophage MS2 in rodent cells (Alonso et al., 1992). Since mutations in the loop of TAR eliminate transcriptional activation by Tat in vivo, it was proposed that a cellular cofactor may bind to the TAR loop and cooperate with Tat to activate transcription. By using gel retardation assays and nuclease digestion of RNA-protein complexes, a number of proteins were shown to bind to TAR (Gatignol et al., 1989; Gaynor et al., 1989; Marciniak et al., 1990). Gel retardation assays and UV-cross-linking revealed that three proteins, TRP185, TRP140 and a 68 kDa protein bind to TAR (Sheline et al., 1991; Wu et al., 1991; Marciniak et al., 1990). The binding of the 68 kDa protein to TAR may not be specific since its interaction with TAR can be eliminated by adding nonspecific RNA or poly (I)-poly (C) into the binding reactions (Wu et al., 1991).
The binding of TRP185 to TAR requires both the TAR loop sequences and secondary structure as well as cofactors that are themselves not capable of binding TAR RNA directly (Sheline et al., 1991; Wu et al., 1991). The addition of TRP185 alone or TRP185 and the cofactor fraction stimulated the expression of the wild type HIV-LTR with minimal effects on the mutant template (Wu et al., 1991).

Gatignol et al. (1991) isolated a cDNA clone by using a uniformly $^{32}$P-labeled TAR RNA as a probe to screen a cDNA expression library. This cDNA encodes a 44-kDa protein, termed TRBP. TRBP was found to bind to a native TAR RNA probe as well as TAR RNA probes with changes in the loop and the bulge. TRBP can activate the HIV-1 LTR and increase gene expression from the HIV-1 LTR synergistically with Tat. TRBP was later found to bind to the Rev-responsive element (RRE) in the RNA of HIV-1 (Park et al., 1994). It was demonstrated to be a potent inhibitor of dsRNA-mediated activation of PKR (Bischoff and Samuel, 1989), which is a ribosome-associated protein kinase and phosphorylates eIF-2α. TRBP also complemented the growth and protein-synthesis defects of a vaccinia virus mutant lacking the expression of the dsRNA-binding protein E3L. Therefore, TRBP was proposed to represent a class of cellular RNA-binding protein that may antagonize interferon-mediated translational control mechanisms (Park et al., 1994).

A cellular co-factor may also interact directly with the activation domain of Tat in the absence of TAR. This was first suggested by the observation that transcriptional activation by Tat was inhibited by overexpression of the Tat activation domain. Several cellular proteins have been found to interact with Tat (Jones and Peterlin, 1994). A cDNA clone was isolated by using biotinylated Tat as a probe to screen a cDNA expression library (Nelbock et al., 1990). This cDNA encoded a protein, termed TBP1, that was able to specifically suppress Tat-mediated transactivation. TBP1 may be a cellular transactivator because a TBP1-GAL4 fusion protein can activate certain promoters in vivo (Ohana et al., 1993). A human protein, MSS1, identified as a suppressor of the growth defect of a yeast strain bearing a mutation in the SGV1 gene, which encodes a CDC28/CDC2 related kinase, was found to be a human homologue of TBP1 (Shibuya...
et al., 1992). Co-expression of MSS1 was found to increase Tat-mediated transactivation. It is still not clear that TBPI and MSS1 are able to interact directly with the activation domain of Tat.

Using Tat affinity chromatography, a 36 kDa human protein was identified as a Tat binding protein (TAP) (Desai et al., 1991). The gene encoding TAP was recently cloned (Yu et al., 1995b). TAP was demonstrated to interact with Tat in vitro and in vivo (Yu et al., 1995a; Yu et al., 1995b). The portion of Tat that interacts with TAP was mapped to a 17-amino acid conserved core domain within the activation region of Tat (Yu et al., 1995a). TAP can function as a transactivator when fused to the GAL4 DNA-binding domain. Its activation domain was localized to its carboxy-terminal portion. The activation domain of TAP was demonstrated to interact with the general transcription factor TFIIB. TAP was therefore proposed to be a coactivator that bridges Tat to the general transcription machinery via TFIIB.

A 42 kDa cellular protein was also reported to bind specifically to the activation domain of Tat but it did not bind to full-length HIV-1 Tat (Herrman and Rice, 1993). This protein displays a serine/threonine kinase activity that is able to phosphorylate Tat and the CTD of RNA polymerase II (Herrman and Rice, 1995). The effect of this kinase on Tat-mediated transactivation remains unknown.

Models for transcriptional activation by Tat

In most cellular assay systems and in cell free systems, the major effects of Tat are on transcription elongation, manifested as a suppression of transcriptional polarity (Cullen, 1993; Jones and Peterlin, 1994), but significant effects on the rate of initiation also have been reported (Jones and Peterlin, 1994). These and other studies have led to a model (Cullen, 1993; Jones and Peterlin, 1994) in which TAR-bound Tat on a paused promoter-proximal elongation complex "reaches back" to alter the elongation potential of the subsequently-formed transcription complex, either during assembly of the preinitiation complex or promoter clearance, and possibly alters the rate of initiation as well. The effect of Tat on initiation is supported by the demonstrated interactions of Tat with TBP (Kashanchi et al., 1994) and a specific TAF (Chiang et al., 1995). However, consistent with the idea that Tat mainly exerts its effects through elongation factors, in
vitro studies have shown that Tat acts synergistically with TFIIS and can mimic or enhance the action of a fraction containing TFIIF and TFIIH (Kato et al., 1992). It seems unlikely that Tat can influence elongation by interacting with TBP or a TAF. Of the several proteins that have been implicated as Tat cofactors, the relevance and mechanisms of action of the known TAR RNA-binding proteins (TRPs) and Tat-binding proteins (TBP-1 and TAP), all candidate cofactors, remain to be established (Jones and Peterlin, 1994).
Chapter II

Binding of basal transcription factor TFIIH to the acidic activation domains of VP16 and p53

A version of this chapter was published in Molecular and Cellular Biology Vol. 14 pp.7013-7024 in October, 1994. I performed all of the experiments presented in this chapter except those shown in figure 6 (a and b), which were performed by Angela Pearson.
Binding of basal transcription factor TFIIH to the acidic activation domains of VP16 and p53

Running title: Interactions of acidic activators with TFIIH

Hua Xiao¹,², Angela Pearson¹,², Benoit Coulombe²#, Ray Truant¹,², Sherry Zhang², Jeffrey L. Regier³, Steven J. Triezenberg³, Danny Reinberg⁴, Osvaldo Flores⁴, C. James Ingles¹,² & Jack Greenblatt¹,²*

¹Department of Molecular and Medical Genetics and
²Banting and Best Department of Medical Research, University of Toronto,
112 College Street, Toronto, Canada M5G 1L6
³Department of Biochemistry, Michigan State University,
East Lansing, Michigan 48824, USA
⁴Robert Wood Johnson Medical School, University of Medicine and
Dentistry of New Jersey, Piscataway, New Jersey 08854-5635, USA

# Present address: Département de Biologie, Université de Sherbrooke, Sherbrooke, Canada J1K2R1
* Correspondence Author:
Dr. J.F. Greenblatt
Banting and Best Department of Medical Research
University of Toronto
112 College Street, Rm. 210
Toronto, Ontario
Canada M5G 1L6
Tel: (416) 978-4141
FAX: (416) 978-8528
ABSTRACT

Acidic transcriptional activation domains function well in both yeast and mammalian cells and some have been shown to bind the general transcription factors TFIID and TFIIB. We now show that two acidic transactivators, herpes simplex virus VP16 and human p53, directly interact with the multi-subunit human general transcription factor TFIIH and its *Saccharomyces cerevisiae* counterpart, factor b. The VP16- and p53-binding domains in these factors lie in the p62 subunit of TFIIH and in the homologous subunit, TFB1, of factor b. Point mutations in VP16 that reduce its transactivation activity in both yeast and mammalian cells weaken its binding to both yeast and human TFIIH. This suggests that binding of activation domains to TFIIH is an important aspect of transcriptional activation.
INTRODUCTION

Accurate transcription in vitro by human RNA polymerase II involves the general transcription factors TFIIA, TFIIB, TFIID, TFII E, TFII F (RAP30 and RAP74), TFII H (also known as BTF2) and TFII J (reviewed in references 13 and 121). Beginning with TFIID, which recognizes the TATA boxes present in many promoters for RNA polymerase II, these factors and RNA polymerase II can be assembled in a defined order onto a promoter (5). TFII H and TFII J are the last of these factors to bind to an assembling initiation complex (12, 14, 26) and TFII H, also known as BTF 2, is the only factor known to possess associated enzymatic activities. These include an ATP-dependent DNA helicase activity (95, 96) and a protein kinase activity that can phosphorylate the carboxy-terminal heptapeptide repeat domain (CTD) of the largest subunit of RNA polymerase II (12, 21, 68). Evolutionarily related general initiation factors are utilized in the yeast Saccharomyces cerevisiae, and the yeast counterpart of human TFII H was originally called factor b (21, 29).

Regulation of transcription by RNA polymerase II involves DNA-binding activator proteins, negative regulators, and coactivators (reviewed in references 91 and 121). Activator proteins often contain separate domains for site-specific DNA binding and transcriptional activation. Activation domains, which function even when attached to the DNA-binding domain of a heterologous protein, interact with the basal transcription machinery via DNA looping (reviewed in references 32, 80, 81). Activation domains vary greatly in amino acid composition and sequence and include, among others, many that are highly acidic, others that are glutamine-rich, and some that are proline-rich (73). Acidic activation domains were originally identified in the yeast activators GCN4 (47) and GAL4 (69) and are also found in the C-terminal 78 amino acids of the herpes simplex virus transactivator VP16 (107) and in the N-terminal 73 amino acids of the mammalian tumor suppressor protein p53 (23). The p53 protein has a site-specific DNA binding domain in its carboxy-terminal portion (101). VP16 does not, by itself, bind specifically to DNA, but instead its amino-terminal portion binds DNA in association with mammalian factors, including the POU-homeodomain protein Oct-1 that recognizes octamer sequences in
DNA (28, 60, 102). When fused to the DNA-binding domain of GAL4, the VP16 and p53 activation domains stimulate transcription in yeast and human cells of a gene bearing GAL4-binding sites (9, 16, 23, 74, 87, 92). These observations imply that common mechanisms for transcriptional activation by acidic activators may exist in fungi and mammals.

Many activation domains have been shown to interact with TBP, the TATA-box-binding subunit of TFIID. These include the highly acidic activation domains in VP16 (50, 103), p53 (10, 67, 72, 84, 97, 108), c-myc (39), v-rel and c-rel (55, 118), and E2F-1 (37), as well as other kinds of activation domains found in the adenovirus activator E1A (48, 62), the Epstein-Barr virus proteins Zta (64) and R (70), the HTLV-1 activator Tax1 (8), the transactivator Tat of HIV-1 (53), and human c-fos and c-jun (85), PU-1 (38), and Sp1 (20). In certain cases, reduced binding of TBP by activation domains with point mutations that reduce transactivation in vivo has provided evidence for the biological relevance of these interactions (8, 50, 53, 62). The TBP-associated factors, known as TAFs, in the TFIID complex are required for transcriptional activation (82), and particular TAFs have also been shown to bind Sp1 (46) and VP16 (30). Although the ability to bind TAF40 has been attributed to a particular subdomain of VP16 (30), point mutations in the activation domain of VP16 that simultaneously affect both transcriptional activation and binding to TAFs have not yet been described.

Several transcriptional activation domains have also been shown to bind TFIIB directly. These include activation domains in VP16 (30, 66), Epstein-Barr virus R (70) and human c-rel (118) and CTF (58), as well as COUP-TF and other members of the nuclear receptor superfamily (1, 49). Consistent with these observations, activator proteins characteristically affect early steps in the assembly of the pre-initiation complex. Some experiments suggest that acidic activators accelerate a slow step involving TFIID and TFIIA (112, 114). Other experiments indicate that a direct interaction of the activator with TBP and/or TFIIB leads to recruitment of TFIIB into the pre-initiation complex (11, 58, 65, 66, 105). When the DNA template contains multiple activator-binding sites, acidic activators also stimulate a subsequent step in the assembly of the pre-initiation complex involving recruitment of the other general initiation factors and RNA
polymerase II (11). However, none of these studies on the mechanism of activation exclude the possibility that transactivation can also affect a later step in transcription by RNA polymerase II, such as formation of the open complex or chain elongation. Indeed, several types of mutations in VP16 that reduce transcriptional activation also reduce open complex formation (51). Moreover, the *Drosophila* heat shock factor (HSF) prevents pausing by RNA polymerase II about 20-40 bp downstream of the hsp70 initiation site (86) and both VP16 and the Tat protein of HIV-1 increase the processivity of chain elongation by RNA polymerase II (52, 61, 71, 119). Here we show that two acidic activators, VP16 and p53, interact directly and analogously with human TFIIH and its yeast counterpart, factor b. This interaction could be important for a late step in the pathway leading to production of a transcript by RNA polymerase II.
MATERIALS AND METHODS

Protein purification. Bacterially expressed human TBP was purified as previously described (45) with modifications. Strain BL21 containing plasmid pET27 (79) was grown in 2X YT medium at 37°C to an OD600 of 0.5. The expression of human TBP was induced with 0.4 mM isopropylthiogalactoside at 30°C for 30 min. Then 15 g of cells was harvested by centrifugation for 20 min at 4,000 rpm and stored frozen at -70°C. The frozen cell pellets were thawed and ground with 22.5 g of alumina. The mixture was resuspended in 22.5 ml HEMGN buffer containing 0.3 M KCl. The alumina and cell debris were removed by centrifugation at 40,000 rpm in a Beckman Ti70 rotor for 2 hours. HEMGN buffer was added to the supernatant to adjust its KCl concentration to 106 mM. Then 75 ml of diluted supernatant was loaded on a 125 ml DEAE-Sepharose column (Pharmacia) which was pre-equilibrated with HEMGN containing 100 mM KCl. Flow-through and wash fractions were collected as 50 ml fractions. 200 ml fractions containing the human TBP were loaded onto a 6 ml Heparin-Sepharose CL-6B (Pharmacia) column pre-equilibrated with HEMGN containing 100 mM KCl. The column was washed with 120 ml HEMGN buffer with 100 mM KCl and then eluted with a 120 ml gradient from 0.2 M KCl to 1 M KCl in HEMGN buffer. Fractions of 3 ml each were collected. The TBP protein was monitored by Western blotting using a cross-reacting antiserum raised against yeast TBP. The peak fractions were pooled and dialyzed against 0.1 M KCl in nuclear dialysis buffer (buffer D [98]). The aliquots were stored at -70°C.

Recombinant human RAP74 was expressed in the strain BL21 and purified as previously described (24). Purification of RAP74 was monitored by Western blotting using an antiserum against RAP74. The peak fractions were dialyzed against 0.1 M KCl in buffer D and stored at -70°C. Recombinant TFIIEα and TFIIEβ were purified as described previously (78). Recombinant RAP30 was purified as previously described (24). Highly purified calf thymus RNA polymerase II (99) was provided by S. McCracken. Recombinant p62 (plasmid pET11a-BTF2, generously provided by R. Roy and J.-M. Egly) was produced in a strain carrying the plasmid pET11a-BTF2. The procedure used for purification of p62 was similar to that used for
the purification of human TBP described above, except that the cells were grown in 2X YT medium to an OD600 of 0.5 at 30°C and induced with 0.4 mM isopropyl-thiogalacto-pyranoside for 3 hours. The DEAE-Sepharose flow-through fraction was dialyzed against ACB buffer containing 0.1 M NaCl.

The TFIIA/J fraction was purified as previously described (104). The TFIIH-mono-S fraction was purified as described (26). To purify TFIIH by VP16 chromatography, 80 ml of HeLa whole cell extract was chromatographed as described previously (103) on a 4 ml GST-VP16 affinity column (4 mg/ml). Bound proteins were eluted with 16 ml of 1 M NaCl in ACB buffer and dialyzed against 2 liters of buffer C (89). To further purify this TFIIH, 15 ml of bound proteins in 0.1 M KCl buffer C were loaded onto a 1 ml phosphocellulose column (89). The column was washed with 0.1 M KCl in buffer C and step eluted with 5 ml each of 0.3 M KCl, 0.5 M KCl and 0.85 M KCl in buffer C. Eluted fractions (1 ml each) were dialyzed against 0.1 M KCl in buffer D (98). The 0.85 M KCl eluate was found to have most of the TFIIH activity and no other general transcription factors were detected by activity assay and, in the cases of TBP and TFIIIB, also by western blotting (data not shown). This chromatographic behaviour on phosphocellulose of VP16-bound TFIIH was identical to that previously described for TFIIH (26). The peak fractions eluted with 0.85 M KCl were dialyzed against 0.1 M KCl in buffer D and 20 µl were loaded on 20 µl GST or GST-VP16 affinity columns (4 mg/ml) under identical conditions. The columns were washed with 20 µl of 0.1 M KCl in buffer D and 40 µl was collected as the flow-through fraction. The columns were washed again with 180 µl of 0.1 M KCl in buffer D and eluted with 80 µl of 1 M KCl in buffer D. The eluted fractions were dialyzed against 0.1 M KCl in buffer D prior to assaying for TFIIH activity.

Affinity chromatography. The GST, GST-VP16, GST-VP16^FP442, GST-VP16^A456, GST-VP16^A456^FP442 and GST-p53 (4 mg/ml each) affinity columns were prepared by immobilizing GST or GST fusion proteins on glutathione-Sepharose beads as described (65). The affinity columns coupled with PA-VP16, PA-VP16^FA442, PA-VP16^FP442, PA-VP16^FA442^FP473, PA-VP16^FA473^475, PA-VP16^FA442^475 or PA-VP16^FA442^473^475 (3 mg/ml
each) were prepared as described (103). For Fig. 5a, the GST-p53 column and control GST column were prepared as described (108). Affinity chromatography was performed as described previously (103). However, bound proteins were usually eluted with ACB buffer containing 1 M NaCl except as indicated otherwise. Eluates used in transcription reactions were dialyzed against 0.1 M KCl in buffer D.

**Antibodies and Western blotting.** Monoclonal antibody M. Ab3c9 (25) against p62 of BTF2 was provided by J-M. Egly. Anti-yeast TFB1 serum (22) was provided by J. Feaver and R. Komberg. Western blotting was performed using an alkaline phosphatase-coupled goat anti-rabbit-IgG or anti-mouse-IgG as described by the manufacturer (Bio-Rad). For Fig. 5d, the ECL Western blotting detection system (Amersham) was used.

**In vitro transcription.** Transcription in vitro on linear DNA templates and RNA analyses were performed essentially as described previously (15), except that 0.5 unit/ml RNaseT1 and 0.5 mM 3'-O-methyl-GTP (Pharmacia) were added to the reactions. Transcripts were initiated at the Ad2MLP on the plasmid pML(C2AT)-52 (93) in reactions containing bacterially expressed recombinant TFIIB (100 ng/ml), TBP (38 ng/ml), TFIIIEα (100 ng/ml), TFIIIEβ (100 ng/ml), RAP30 (50 ng/ml), and RAP74 (50 ng/ml), a TFIIA/J fraction (1 μl), TFIIH fraction (2 μl) and calf thymus RNA polymerase II (2.5 μg/ml).

**In vitro protein-binding assays**

$^{35}$S-labeled proteins were synthesized by using the plasmids encoding p62 (25), TFB1 (29), SSL1, SSL2 and RAD3 (22) and the TnT transcription/translation system (Promega) according to the protocol provided by Promega. The input fractions used for binding assays contained these $^{35}$S-labeled proteins diluted 10 times with ACB buffer containing 0.1 M NaCl, 0.5% NP40 and 1 mg of BSA per ml. Glutathione-Sepharose beads (20 μl) containing immobilized proteins were pre-incubated with ACB buffer containing 0.1 M NaCl, 0.5% NP40 and 1 mg/ml of BSA for two hours, then incubated with input (200 μl) at 4°C for two hours, and washed four times with ACB buffer containing 0.1 M NaCl and 0.5% NP40 (1 ml). The bound proteins were eluted with ACB buffer containing 1 M NaCl (80 μl) and the beads were then boiled in 80 μl of
SDS-PAGE sample buffer. For Fig. 6b, the $^{35}$S-labeled proteins were diluted 10 times with ACB buffer containing 0.1 M NaCl and 1 mg/ml of BSA and then used as inputs for affinity chromatography. Bound proteins were eluted with ACB buffer containing 1 M NaCl.

RESULTS

Interaction of the VP16 Activation Domain with TFIIH. To determine in a systematic way which general initiation factors bind to VP16, fusion proteins containing the acidic activation domain of VP16 (amino acids 412-490) fused to glutathione-S-transferase (GST) were produced in E. coli and used as ligands for affinity chromatography (66). The affinity columns were loaded with HeLa whole cell extracts, and bound proteins were eluted with salt and assayed for the general factors required for initiation by RNA polymerase II in vitro at the adenovirus major late promoter (MLP). Our reconstituted in vitro transcription system consisted of the recombinant human general factors TBP (79), TFIIIB (36), TFIIIE (α and β) (78), and TFIIIF (RAP30 and RAP74) (24, 99) all produced in E. coli; highly purified calf thymus RNA polymerase II (100); and the general initiation factor fractions TFIIA/TFIIJ (104) and TFIIH (26) partially purified from HeLa nuclear extracts. The TFIIA/TFIIJ fraction used in our experiments contained both TFIIA and TFIIJ, but only the TFIIJ activity is necessary for production of run-off transcripts from the adenovirus MLP in reactions containing recombinant TBP (14). As is shown in Figure 1a, no transcript was produced in the absence of TFIIH (lane 2) and the eluate from a GST-VP16 column (lane 4) contained an activity that could substitute for the TFIIH fraction (lane 1) to support transcription. In contrast, the eluate from a control GST column lacked TFIIH activity (lane 3). To confirm that the TFIIH activity binding to the VP16 column corresponded to the same factor, known as TFIIH or BTF2, that was described in previous experiments (26, 27) a monoclonal antibody (M. Ab3c9) (25) against the 62 kDa subunit (p62) of BTF2 was used. This monoclonal antibody inhibited the transcription activities of both TFIIH purified by a conventional procedure (26) (Fig. 1b, lanes 3-5) and the TFIIH activity that was eluted from a VP16 column (lanes 7-9). In contrast, an irrelevant monoclonal antibody against the retinoblastoma protein had no effect (lanes 16, 17). Moreover, the inhibition by M. Ab3c9
was specific because it could be reversed by the addition to the reaction of excess TFIIH that had been highly purified by VP16 affinity chromatography and subsequent phosphocellulose chromatography (Fig. 1b, lanes 10-13), but it could not be reversed by the addition of excess TFIIB (lanes 14, 15). Furthermore, when the affinity column eluates were subjected to Western blotting with M. Ab3c9, only the GST-VP16 column eluate, and not the control GST column eluate, contained the immunoreactive p62 polypeptide (Fig. 1c). Therefore, VP16 binds authentic TFIIH. This binding of TFIIH to VP16 was selective, since the binding of TFIIE and TFIIF polypeptides could not be detected by Western blotting with appropriate antisera (data not shown), and since no measurable amounts of TFIIE, TFIIF, or RNA polymerase II activity and only a trace amount of TFIIA/J activity bound specifically to the GST-VP16 column (Fig. 1d, lanes 3, 5, 7, 9). Previous studies have shown that the VP16 activation domain can bind TFIIB (65) and two subunits of TFIID, TBP (103) and TAF40 (30). Consistent with these studies, we found that, in addition to TFIIH activity, GST-VP16 columns also bound TFIID and TFIIB activities (data not shown).

The Interaction between VP16 and TFIIH is Related to Transcriptional Activation. To examine the specificity and biological relevance of the interaction between VP16 and TFIIH, we tested whether mutations in the VP16 activation domain (17, 88, 107) that affect its transcriptional activity also affect the binding of TFIIH. The structures of the fusion proteins used in these experiments are shown in Fig. 2. Truncation of the activation domain of VP16 to position 456 reduced the transactivation by VP16 to approximately 50% of maximal activity when a reporter gene containing the herpes simplex virus ICP4 promoter was used (17) and to an undetectable level when a GAL4-VP16 fusion protein was used for activation and the target gene had only one GAL4-binding site (111). This carboxy-terminal deletion (Δ456) had a large effect on the binding of TFIIH activity (Fig. 2a, lanes 4 and 6).

In the context of a full length VP16 activation domain, a Phe to Pro or Ala substitution at position 442 also moderately reduced the transactivation potency of VP16 with an ICP4 reporter gene (88). These mutations had a similar effect on activation by GAL4-VP16 in yeast (3) and a
stronger effect with reporters bearing single GAL4-binding sites in mammalian cells (111). These mutations in the amino-terminal portion of the activation domain of VP16 also substantially reduced the binding of TFIIH to the activation domain of VP16 (Fig. 2b, lanes 4, 5). Mutation of Phe residues at positions 473 and 475 within the carboxy-terminal portion of the activation domain by themselves have little effect on transactivation by VP16 (43) and only slightly reduced the binding of TFIIH to VP16 (Fig. 2b, lane 7). However, mutations at Phe473 and Phe475, when combined with a mutation at the amino-terminal 442 position, significantly reduced transactivation by VP16 (88) or GAL4-VP16 (111) and markedly reduced the ability of VP16 to bind to TFIIH (compare lanes 4, 5 with lanes 6, 8 and 9). In summary, our data show that both the amino-terminal (amino acids 412-456) and carboxy-terminal (amino acids 457-490) portions of the VP16 activation domain are very important for binding TFIIH. This correlation between transactivation activity in vivo and binding of VP16 to TFIIH in vitro implies that TFIIH is a target of VP16 for the purpose of transcriptional activation.

The VP16 Activation Domain Binds Directly to the 62 kDa Subunit (p62) of TFIIH. Since the activation domain of VP16 binds directly to TFIID (30, 103), TFIIB (66) and DNA replication factor A (RPA) (40, 63), it was possible that one of these proteins might act as a bridge between VP16 and TFIIH. To test this possibility, the TFIIH activity eluted from a GST-VP16 column was further purified by phosphocellulose chromatography (Fig. 3a). The resulting phosphocellulose fractions had TFIIH activity (lanes 6, 7 and 8), but undetectable amounts of the other general transcription factors as judged by the in vitro transcription assay and by Western blotting analysis with antibodies against TBP and TFIIIB (data not shown). Furthermore, this fraction did not contain detectable amounts of RPA as judged by Western blotting analysis with monoclonal antibodies against the 70 kDa and 32 kDa subunits of RPA (data not shown). When this highly purified preparation of TFIIH was tested for binding to VP16, 75-80% of the TFIIH activity could still bind to a GST-VP16 column (Fig. 3b, compare lane 4 to lane 2), while virtually all the activity again flowed through a control GST column (compare lane 1 to lane 3). Therefore, we concluded that interaction of the VP16 activation
domain with TFIIH might not require the presence of any other general transcription factor or RPA and could be direct.

Mammalian TFIIH and its yeast equivalent, factor b, have at least five subunits (12, 21, 22, 27). Since human cDNAs encoding the 62 kDa subunit have been cloned (25), we tested the possibility that p62 might mediate the interaction between VP16 and other subunits of TFIIH. For this purpose, agarose beads containing immobilized GST or GST-VP16 were incubated with 35S-labeled p62 that had been synthesized in vitro in a rabbit reticulocyte lysate. The beads were washed with low salt buffer and bound proteins were eluted successively with buffers containing 1 M NaCl and 1% SDS. The eluates were then analyzed by SDS-PAGE followed by autoradiography. As shown in Figure 4a, full length p62 made in vitro (lane 1) bound efficiently to the GST-VP16 beads (lanes 3 and 5), but only weakly to the control GST beads (lanes 2 and 4). About half of the p62 was eluted with 1 M NaCl (lane 3) and the rest was only eluted with buffer containing 1% SDS (lane 5). As was the case for binding of TFIIH activity from a HeLa extract (Fig 2b), the FP442 mutation in VP16 greatly reduced the ability of GST-VP16 to bind p62 made in vitro (Fig. 4b, compare lane 3 with lane 4), indicating that p62 contains a biologically important site in TFIIH for binding VP16. To rule out the possibility that mammalian proteins that are present in reticulocyte lysate might mediate the interaction between VP16 and p62, we also used a bacteriophage T7 RNA polymerase-based system to produce p62 in E. coli. In this case, soluble bacterial extract that was partially purified by passing it through a DEAE-Sepharose column was loaded onto affinity columns containing immobilized GST or GST-VP16. The bound proteins were eluted and analyzed by SDS-PAGE followed by Western blotting with M. Ab3c9. As shown in Fig. 4c (lane 1), the input fraction used for affinity chromatography contained a small amount of apparently intact bacterially produced p62 and large amounts of proteolytic fragments of p62 that were detected with M. Ab3c9. All of the polypeptides detected by M. Ab3c9 were derived from p62, because none were present in an extract derived from a control strain containing a plasmid vector lacking p62 sequences (data not shown). None of these p62-derived polypeptides bound to the GST column (lane 2), but a 30
kDa proteolytic fragment of p62 bound to the GST-VP16 column (lane 3). Therefore, VP16 binds directly to the 62 kDa subunit of TFIIH. Curiously, the largest form of p62 made in E. coli did not bind VP16 in these experiments, possibly because full length p62 made in E. coli, unlike full length p62 made in a reticulocyte lysate, is not properly folded or modified for binding VP16.

The Activation Domain of p53 also Binds TFIIH. The p53 activation domain (amino acids 1-73) is similar in size, net negative charge, and potency to that of VP16 (23, 74, 87). Previous experiments have shown that the p53 activation domain binds TBP (67, 72, 84, 97, 108), but not TFIIIB (67). To test whether the binding of VP16 to TFIIH might represent a phenomenon with general validity for acidic activation domains, we investigated whether p53 could also bind TFIIH. First, HeLa whole cell extracts were chromatographed on GST and GST-p53 (amino acids 1-73) columns (108). Bound proteins were eluted with buffer containing 0.5 M NaCl and assayed for the various general initiation factors by using the reconstituted in vitro transcription system described for Fig. 1. As is shown in Fig. 5a, eluate from the GST-p53 column (lane 10), but not the GST control column (lane 9), contained an activity that could substitute for TFIIH (lane 1). Consistent with this, Western blotting with M. Ab3c9 (Fig. 5b) showed that only the GST-p53 column (lane 3) and not the GST column (lane 2) bound the immunoreactive p62 found in the HeLa extract (lane 1). In agreement with previous experiments, the GST-p53 column, but not the GST control column, also bound TFIIID activity (Fig. 5a, lanes 1-4), but relatively little TFIIIB activity (lanes 5-8). The trace amount of TFIIIB bound by p53 may represent an indirect association mediated by TFIIID since we and others have not detected binding of recombinant TFIIIB to p53 (67).

We also used recombinant p62 made in reticulocyte lysates or in E. coli in p53 binding experiments. As shown in Fig. 5c, intact 35S-labeled p62 made in a reticulocyte lysate (lane 1) bound to GST-p53 beads (lane 3), but not to GST beads (lane 2). Similarly, as was the case for VP16 (Fig. 4c), only a 30 kDa fragment of p62 made in E. coli bound to a GST-p53 column (Fig. 5d, lane 3) and also, as before, it did not bind to a GST column (lane 2). Taken together,
these data imply that there are similar binding sites for VP16 and p53 in the p62 subunit of TFIIH.

The VP16 and p53 Activation Domains also Interact with Yeast Factor b. Since acidic activation domains, including those of VP16 and p53, generally function well in *Saccharomyces cerevisiae*, interactions of activators with the yeast analogue of TFIIH provide a critical test for the biological importance of activator-TFIIH interactions. As a first test of whether VP16 can bind factor b, yeast whole cell extract was chromatographed on affinity columns containing immobilized GST or GST-VP16. Bound proteins were eluted with buffer containing 1 M NaCl and analyzed by Western blotting with a rabbit antiserum against TFB1, a 73 kDa subunit of factor b (29). As shown in Fig. 6a, the GST-VP16 column (lane 2), but not the GST column (lane 1), bound some of the TFB1 in the yeast whole cell extract. Also, as expected from the results of previous experiments showing that VP16 directly binds yeast TBP (103), the GST-VP16 column bound a substantial amount of the TBP in the extract (data not shown).

TFB1 is the yeast homologue of human p62 (25, 29), and SSL1 (120) and RAD3 (44) have also been identified as subunits of yeast TFIIH (factor b) (22). As well, SSL2 (RAD25) (33, 83) is the yeast homologue of human XBP (ERCC3) (113 and references therein), which is associated with purified human TFIIH (BTF7) (95). Therefore, to identify the subunit of yeast TFIIH which binds to VP16, 35S-labeled TFB1, SSL1, RAD3, and SSL2 were synthesized in vitro in a rabbit reticulocyte lysate and subjected to chromatography on GST and GST-VP16 columns. Bound proteins were eluted with buffer containing 0.5 M NaCl and analyzed by SDS-PAGE and autoradiography. As shown in Fig. 6b, the full length TFB1 made in vitro, but not shorter fragments of TFB1 (see lane 1), bound to a GST-VP16 column (lane 3), but not a GST column (lane 2), while SSL1, RAD3, and SSL2 did not bind to either kind of column (lanes 4-12). Similarly, a GST-p53 (amino acids 1-73) column bound full length 35S-labeled TFB1 (Fig. 6c, lane 5), while a control GST column again did not (lane 2). Therefore, TFB1, the yeast counterpart of human p62, contains a binding site for VP16 and p53. The inability of shorter
forms of TFB1 made in vitro, probably as a consequence of transcriptional initiation at internal AUG codons, to bind VP16 or p53 suggested that the amino-terminal region of TFB1 might be important for binding acidic activators. Consistent with this, an amino-terminal fragment containing residues 1-526 of TFB1 (Fig. 6d, lane 1) also bound to GST-VP16 (lane 3), but not GST (lane 2), indicating that the carboxy-terminal portion of TFB1 is not needed for binding acidic activators. Moreover, the FP442 mutation in VP16 affects activation by GAL4-VP16 in Saccharomyces cerevisiae (3), as in mammalian cells (111), and greatly reduced binding to VP16 of 35S-labeled yeast TFB1 (Fig. 6c, compare lanes 3 and 4) just as it reduced binding to VP16 of p62 (Fig. 4b). Therefore, we conclude that acidic activators bind in very similar ways to human p62 and yeast TFB1.
DISCUSSION

Our data demonstrate that transcriptional activation domains can interact with the general transcription factor TFIIH. Binding of TFIIH by the VP16 activation domain was not detected previously (65, 103) because TFIIH activity had not yet been separated from fractions containing TFIID, TFIIIE and TFIIIF, and so TFIIH activity could not be assayed. Many binding studies with other activator proteins have been published since then, but these have generally focussed on TBP, TFII B, and TAFs and no systematic effort was made to examine the binding to activators of other general transcription factors. Depending on the salt and ligand concentrations and perhaps the particular extract used in the experiments, immobilized VP16 quantitatively removed either the TFIID activity (103) or the TFII B activity (65) from a HeLa nuclear extract, but apparently did not quantitatively remove the TFIIH activity. This may mean that there is an excess of TFIIH in an extract or that there is more than one form of TFIIH in an extract, only one of which contains its 62 kDa subunit and binds to VP16. Alternatively, since transcriptional initiation in vitro on supercoiled DNA does not require TFIIH (76, 109), transcription may not have been TFIIH-dependent in some of these experiments.

The correlation between the effects of the FP442 mutation and other VP16 point mutations on transactivation (3, 17, 88, 111) and their effects on binding to VP16 of human and yeast TFIIH (Figures 2, 4, and 6) is striking. Therefore, these interactions are likely to be important for the activation of transcription. Two other observations strengthen this conclusion: firstly, another acidic activation domain, that of p53, also binds to human p62 and its yeast homologue, TFB1; and secondly, the ability to bind both VP16 and p53 has been evolutionarily conserved even though p62 and TFB1 are only 21% identical in overall amino acid sequence (25, 29).

The effects of TFIIB mutations on activation by GAL4-VP16 in vitro (90) suggest that an activator-TFIIB interaction is also important for transactivation. One experiment with the FP442 mutation in VP16 led to the same conclusion (66). In other studies, neither the FP442 mutation in the amino-terminal portion of the VP16 activation domain nor mutations in important phenylalanine residues in the carboxy-terminal portion of the VP16 activation domain, which all
strongly reduce transactivation in vivo (17, 88, 111), affected the binding to VP16 of TFIIB (30, 34, 111). These mutations do, however, reduce the binding to VP16 of TFIIH (Figures 2, 4, 6), human TFIID activity (34), and yeast TBP (50). Since the critical F442 residue of VP16 is outside the portion of VP16 that binds TAF40 and point mutations in VP16 that affect both transactivation by VP16 and the binding to VP16 of TAF40 have yet to be described (30), the effects of VP16 point mutations on transactivation seem to correlate best so far with their effects on the binding to VP16 of TFIID (TBP) and TFIIH (p62). It is unclear whether VP16 point mutations that affect activation have similar effects on the binding to VP16 of TBP and TFIIH because these general factors bind to VP16 in similar ways or because the VP16 point mutations affect the folding of the VP16 activation domain.

Like VP16, the proline-rich activation domain of CTF binds TBP (117) and TFIIB (58) and recruits TFIIB into the initiation complex (58). However, since binding of p53 to human TFIIB is very weak (Fig. 5 and ref. 67), a strong interaction of an activator with TFIIB may not always be important for transactivation. Indeed, the glutamine-rich activation domains of Spl bind TBP (20) and recruit TFIIB into the pre-initiation complex (11) even though they have not yet been shown to bind TFIIB. One reasonable possibility is that even a weak interaction of an activator with TFIIB is sufficient to enable that activator to recruit TFIIB into the assembling preinitiation complex. Alternatively, it is possible that interaction of activators with TBP causes TBP to recruit TFIIB into the pre-initiation complex, while interaction of activators with TFIIB and/or TAFs leads to subsequent recruitment of RNA polymerase II and other factors (11). Consistent with this latter possibility, Kim et al. (57) recently described a mutation in TBP which prevents the binding to TBP of VP16 and results in a failure of GAL4-VP16 to recruit TFIIB into the preinitiation complex. Furthermore, in view of the recent discovery in yeast of an RNA polymerase II holoenzyme containing most of the general initiation factors and susceptible to transcriptional activation in the absence of TAFs (56, 59), it is also possible that an activator might simultaneously contact both TBP and TFIIB in order to assemble a preinitiation complex.
The interaction of acidic activators with multiple general initiation factors (e.g. TFIID, TFIIB, and TFIH) is compatible with the high degree of synergy that is observed in transcriptional activation (7, 11) as discussed by Greenblatt (32) and Herchlag (41). Formation of at least a partial open complex and transcriptional initiation on linear DNA do not require TFIIE, TFIH, or ATP, since RNA polymerase II can synthesize a promoter-specific trinucleotide in the absence of these factors (31). However, promoter escape by RNA polymerase II in the absence of stress generated by negative supercoiling does require TFIIE, TFIH and ATP (31). TFIH can assemble late into the preinitiation complex (12, 26) and could be involved in unwinding the DNA to convert a partial open complex to a complete open complex. Consistent with this, TFIH has an associated ATP-dependent DNA helicase activity (95, 96). In particular, the yeast TFIH subunit RAD3 (22) and its human counterpart XPD (ERCC2) (94) are DNA helicases (19, 106). Moreover, human XPB (ERCC3) is a helicase (19, 95), and XPB and its yeast homologue, SSL2, have helicase motifs in their amino acid sequences (33, 75, 113). One could therefore explain part of the synergism in transcriptional activation if interaction of acidic activators with TFIH stimulates melting of the DNA at the promoter. This phenomenon would have been difficult to detect in vitro because acidic activators also stimulate formation of the closed pre-initiation complex (65, 112, 114). However, such a model is consistent with observations that mutations which affect transcriptional activation by VP16 also affect open complex formation (51).

Phosphorylation of the CTD on RNA polymerase II occurs in the transition from an initiation complex to an elongation complex (6, 77). Since human and yeast TFIH have associated protein kinase activities that can phosphorylate the CTD (12, 21, 68), interaction of activators with TFIH may have a role in this process. Phosphorylation of the CTD reduces the affinity of the CTD for TBP (110) and might facilitate the escape of RNA polymerase II from the promoter region following transcriptional initiation. Indeed, this could explain how heat shock factor enables RNA polymerase II to escape a pause site about 20-40 nucleotides downstream.
from the *Drosophila* hsp70 promoter (86). No direct stimulation of the CTD kinase or DNA helicase activity of TFIIH has yet been reported, however, for any transcriptional activators.

Interaction of activators with TFIIH may also lead to more processive transcription by RNA polymerase II. For example, the transactivator protein Tat of HIV-1 binds TBP (53) and activates HIV-1 transcription by both stimulating initiation in the HIV-1 LTR (61 and reviewed in 18) and increasing the processivity of chain elongation by RNA polymerase II (52, 54, 61, 71). Similarly, other typical activators, including VP16, may also generally stimulate chain elongation by RNA polymerase II (119). In recent preliminary experiments we have observed that Tat, like VP16, also binds TFIIH (i.e. p62) (116). Effects of activators like Tat on chain elongation might be mediated, for example, via phosphorylation of the CTD on RNA polymerase II. Alternatively, although the XPD and XPB DNA helicases have not been found in association with elongating RNA polymerase II (19), interaction of activators with p62 may stabilize the association of p62 with the elongation complex. In that case, p62 might recruit XBP or XPD to unwind the DNA at pause sites ahead of the RNA polymerase II. Interaction of Tat with TFIIH might also indirectly stabilize the association with the elongation complex of TFIIF (54), which increases the rate of elongation by RNA polymerase II (2).

An additional possibility, which we cannot dismiss, is that interaction of activators with TFIIH also has a role in DNA repair. The *RAD3* and *SSL2* (*RAD25*) genes encoding subunits of yeast TFIIH are essential genes required for transcription by RNA polymerase II in *Saccharomyces cerevisiae* (35, 83). However, just as mutations in the human *XPB* gene lead to the DNA repair disorders xeroderma pigmentosum and Cockayne’s syndrome (113), subunits of yeast TFIIH are involved in DNA repair in *Saccharomyces cerevisiae* (22, 44, 75). Therefore, interaction of activators with TFIIH might help ensure that TFIIH is still associated with RNA polymerase II when it stalls at a site of DNA damage (discussed in reference 4). Sorting out whether interaction of activators with TFIIH is important for DNA repair as well as transcription will require further experimentation.
Figure 1. The acidic activation domain of VP16 interacts with the basal transcription factor TFIIB.

(a) Transcripts, indicated by the arrow, were initiated at the Ad2MLP on the plasmid pML(C2AT) -50 (93) in reactions containing bacterially expressed recombinant TFIIB, TBP, TFIIEα, TFIIEβ, RAP30, and RAP74, a TFIIA/J fraction, and calf thymus RNA polymerase II. The reactions were supplemented with either 1 μl of HeLa cell TFIIB mono-S fraction (26) (lane 1), or with buffer D (lane 2), or with 4 μl of the bound proteins from a GST column (lane 3), or a GST-VP16 column (lane 4).

(b) Inhibition of transcription by M. Ab3c9 against BTF2. Transcription reactions were as in Fig. 1a. TFIIB mono-S fraction (lanes 2-5 and 16) and TFIIB activity that had bound to a GST-VP16 column and was further purified by phosphocellulose chromatography as described in Fig. 2a (lanes 6-15 and 17) were incubated with M. Ab3c9 or the monoclonal antibody Ab-1 directed against Rb (Oncogene Science) as indicated. Some reactions, as indicated, were supplemented with additional VP16-bound TFIIB (lanes 10-13) or TFIIB (lanes 14, 15).
Figure 1. (c). Protein immunoblotting with M. Ab3c9 against p62 of BTF2. Lane 1, 20 µl of HeLa whole cell extract. Lane 2, 10 µl of TFIIH mono-S fraction. Lane 3, 40 µl of GST column eluate. Lane 4, 40 µl of GST-VP16 column eluate. Arrow indicates the p62 polypeptides.

(d) The binding of TFIIH to VP16 is selective. Transcription reactions missing TFIIH and the other indicated general transcription factor were supplemented with 4 µl of the bound proteins from either a GST-VP16 or a control GST column.
Figure 2. Effects of VP16 mutations on binding of TFIIH.
(a) The C-terminal portion of the VP16 activation domain is important for binding TFIIH. Top panel shows diagrams of the proteins used as affinity column ligands. The numbers indicate the positions of residues in VP16. HeLa cell extract was chromatographed on four 0.5 ml affinity columns containing GST, GST-VP16Δ456, GST-VP16Δ456FP442 or GST-VP16 wild type proteins (4 mg/ml). Transcription reactions containing all the general transcription factors, except TFIIH, were complemented with a TFIIH mono-S fraction (lane 1), buffer D (lane 2), or bound HeLa cell proteins eluted from affinity columns containing immobilized GST, GST-VP16Δ456, GST-VP16Δ456FP442, or GST-VP16 wild type (lanes 3-6).
Figure 2. (b) Point mutations in VP16 affect its ability to bind TFIIH. Top panel shows diagrams of the proteins used as affinity column ligands. The numbers indicate the residues of VP16. Transcription reactions containing the other general factors (see Fig. 1 legend) were supplemented with TFIIH mono-S fraction (lane 1), buffer D (lane 2), or bound HeLa cell proteins from 20 μl affinity columns containing 3 mg/ml immobilized PA-VP16, PA-VP16FA442, PA-VP16FP442, PA-VP16FA442FP473, PA-VP16FA473/475, PA-VP16FA442/475 or PA-VP16FA442/473/475 (lanes 3-9, respectively). Arrows indicate the position of the specific transcript.
Figure 3. Highly purified TFIIH binds to the acidic activation domain of VP16.
(a) Further purification by phosphocellulose chromatography of TFIIH that had been purified by GST-VP16 chromatography. Transcription reactions were used to assay 4 µl aliquots of phosphocellulose column fractions for TFIIH activity as described in Fig. 1a. Salt concentrations used for eluting the column are indicated. For each salt concentration, the two fractions with the highest protein concentrations were assayed.
(b) Transcription reactions lacking TFIIH (see Fig. 1a) were supplemented with phosphocellulose-purified TFIIH activity (0.85 M KCl eluate) that flowed through a GST column or a GST-VP16 column (lanes 1 and 2) or with TFIIH activity bound to the GST column or GST-VP16 column (lanes 3 and 4) (4 µl of each). Arrows indicate the position of the specific transcript. Note that equal volumes of the flow-through and eluted fractions were assayed, but that they represent 10% of the flow-through and 5% of the eluate, respectively.
Figure 4. The 62 kDa subunit of TFIIH mediates the interaction between TFIIH and VP16.  
(a) An autoradiogram of an SDS-polyacrylamide gel shows the binding of VP16 to p62. Lane 1 shows the input p62. The arrow indicates the position of full-length \(^{35}\)S-methionine-labeled p62 made in vitro. The bound proteins were eluted from GST beads and GST-VP16 beads first by treatment with 1 M NaCl in ACB buffer (lanes 2 and 3) and then by boiling in SDS-PAGE sample buffer (lanes 4 and 5).  
(b) An autoradiogram of an SDS-polyacrylamide gel shows the specific binding of p62 to VP16. Lane 1 is an aliquot of the in vitro translated p62 used as input. Salt eluates (1 M NaCl) from GST, GST-VP16 and GST-VP16 FP442 beads are shown (lanes 2, 3 and 4, respectively).  
(c) Binding of VP16 to p62 is direct. Soluble bacterial extract containing proteolyzed p62 was loaded on GST and GST-VP16 affinity columns. Aliquots of the input fraction (lane 1) and proteins eluted with salt (1 M NaCl) from the GST column (lane 2) or the GST-VP16 column (lane 3) were analyzed on an SDS-polyacrylamide gel followed by immunoblotting with the p62 monoclonal antibody M. Ab3c9.
Figure 5. The activation domain of p53 binds TFIIH.
(a) TFIIH activity is retained by a p53 affinity column. Affinity chromatography with HeLa whole cell extract and transcription assays were performed as described in Fig. 1a. Reactions contained all the basal transcription factors (lane 1) or were missing either recombinant TBP or TFIIIB, or TFIIH, as indicated at the top of the autoradiogram. Reactions were supplemented with 4 μl of buffer D (lanes 2, 5, 8) or HeLa proteins that were eluted with salt (0.5 M NaCl) from a GST column (lanes 3, 6, 9) or from a GST-p53 column (lanes 4, 7, 10). An arrow indicates the position of the specific transcript.
(b) Western blotting analysis shows that p53 binds authentic TFIIH. HeLa whole cell extract (lane 1) and bound proteins from the GST column (lane 2) or the GST-p53 column (lane 3) were analyzed by SDS-PAGE followed by immunoblotting with anti-p62 (M.Ab3c9).
Figure 5. The activation domain of p53 binds TFIIH.
(c) The activation domain of p53 binds p62. The binding assays contained $^{35}$S-methionine-labeled p62 made in vitro. Lane 1, input; lane 2, the salt eluate (1 M NaCl) from GST beads; lane 3, the eluate from GST-p53 beads.
(d) The binding of p62 to p53 is direct. Affinity chromatography with p62 made in E. coli and Western blotting analysis with M. Ab3c9 were performed as in Fig 4c. Lane 1, input; lane 2, eluted proteins from a GST affinity column; lane 3, eluted proteins from a GST-p53 column.
Figure 6. The activation domains of VP16 and p53 interact with yeast transcription factor b. 
(a) The activation domain of VP16 binds the transcription factor b in a yeast whole cell extract. 
Yeast whole cell extract that was prepared from strain BJ2168 as described previously (115) was 
chromatographed on a 20 μl GST affinity column and a 20 μl GST-VP16 affinity column. After 
washing with 10 column volumes of ACB buffer containing 0.1 M NaCl, the columns were 
eluted with ACB buffer containing 1 M NaCl. Equal volumes of eluates from the GST column 
(lane 1) and the GST-VP16 column (lane 2) were resolved by SDS-PAGE followed by 
immunoblotting with a rabbit polyclonal antiserum against a recombinant GST-TFB1 fusion 
protein. The arrow indicates the position of TFB1. The bracket indicates some GST-VP16 
column ligand that was eluted from the column with salt and also reacted with the antibody 
against GST-TFB1. 
(b). The activation domain of VP16 interacts with TFB1. 35S-labeled TFB1 (lane 1), SSL1 (lane 4), RAD3 (lane 7) and SSL2 (lane 10) made in vitro were chromatographed on 10 μl GST or 
GST-VP16 affinity columns. Bound proteins from GST columns (lanes 2, 5, 8, 11) or GST-
VP16 columns (lanes 3, 6, 9, 12) eluted with ACB buffer containing 0.5 M NaCl were resolved 
by SDS-PAGE followed by autoradiography.
Figure 6. The activation domains of VP16 and p53 interact with yeast transcription factor b.
(c) Mutant VP16 is deficient in binding TFB1, and p53 also interacts with TFB1. Binding assays used $^{35}$S-labeled TFB1 made in vitro (lane 1) as described in Fig 6b. Lane 2, GST eluate; lane 3, GST-VP16 eluate; lane 4, GST-VP16FP442 eluate; lane 5, GST-p53 eluate.
(d) The carboxy-terminus of TFB1 is not needed for the binding of VP16. A truncated $^{35}$S-labeled version of TFB1 was synthesized in vitro after cleavage of the TFB1 plasmid DNA with SalI within codon 527 of TFB1. The binding assays were done as in Fig 6b. Lane 1, input; lanes 2 and 3 are eluates from the GST and GST-VP16 columns, respectively.
ACKNOWLEDGMENTS

We thank J. Feaver and R. Kornberg for providing the plasmids encoding TFB1, SSL1, SSL2, and RAD3 and the antiserum against TFB1, Susan McCracken, Tim Westwood and Joyce Li for providing purified RNA polymerase II and recombinant TFIIE and RAP30, and Jean-Marc Egly and Richard Roy for the gift of M. Ab3c9. Thanks also to Rajesh Gupta and Johnson M.S. Wong for providing plasmids encoding some mutant PA-VP16 and GST-VP16 derivatives. HX also thanks Dr. Yuan Su for much encouragement and helpful discussion.

This work was supported by grants to both JG and CJI from the Medical Research Council of Canada and the National Cancer Institute of Canada. JG is an International Research Scholar of the Howard Hughes Medical Institute. SJT was supported by a grant from NIH and by a Junior Faculty Research Award from the American Cancer Society. AP was awarded a scholarship from the Fonds pour la formation de chercheurs et l'aide à la recherche. BC was awarded a fellowship from the Medical Research Council of Canada.
1. REFERENCES


70. Manet, E., C. Allera, H. Gruffat, I. Mikaelian, A. Rigolet and A. Sergeant. 1993. The acidic activation domain of the Epstein-Barr virus transcription factor R interacts in vitro with both TBP and TFIIB and is cell-specifically potentiated by a proline-rich region. Gene Express. 3: 49-59.


Chapter III

The activation domain of Tat binds to the general transcription factor TFIIH

Some of the data in this chapter was published in Molecular and Cellular Biology (Vol. 16, pp2044-2055) in 1996 as part of an article entitled "Three functional classes of transcriptional activation domain" by Justin Blau, Hua Xiao, Susan McCracken, Tony Kouzarides, Peter O'Hare, Jack Greenblatt, and David Bentley
A. Summary

Transcription by RNA polymerase II can be regulated at the level of both initiation and elongation. Recent experiments have revealed that some activators, including Sp1 and CTF, stimulate initiation, while others, including HIV-1 Tat, VP16, and p53, stimulate both initiation and elongation. I show here that only those activators which stimulate elongation bind to TFIIH. A single amino acid substitution in the activation domain of Tat that abolishes transactivation by Tat also reduces binding of the kinase subunit CDK7 of TFIIH to Tat. Thus, the TFIIH-activator interaction may be involved in stimulation of transcriptional elongation.
B. Introduction

Eukaryotic transcriptional activators often contain separate domains for site-specific DNA-binding and transcriptional activation. Based on the amino acid composition of their activation domains, some transcriptional activators can be classified into three types: acidic, glutamine-rich and proline-rich (Mitchell and Tjian, 1989). Growing evidence has suggested that different types of activators may act at different steps of the transcription process (Bentley, 1995). Some activators, including Sp1, CTF and GAGA, appear to affect only the rate of initiation (Zhou and Sharp, 1995; Blau et al., 1996; Lee et al., 1992). The HIV-1 transactivator protein Tat enhances predominantly the rate of elongation (Jones and Peterlin, 1994), although some effects on the rate of initiation have also been reported. Some activators, like the herpes simplex virus protein VP16 and adenovirus protein E1a, as well as cellular proteins E2F and p53, stimulate both initiation and elongation (Yankulov et al., 1994; Blau et al., 1996).

The Tat protein encoded by the human immunodeficiency virus 1 is a potent transactivator. Tat can increase HIV-1 LTR-directed gene expression by interacting with a short RNA which is located about 40 base pairs from the 5' end of the nascent transcript to the start site of transcription. The HIV-1 Tat protein is an 86-amino acid protein composed of a highly basic C-terminal region, which is important for nuclear localization and RNA binding, and a 48-amino acid N-terminal region, which is important for transactivation (Jones and Peterlin, 1994).

The function of an activation domain is to interact directly or indirectly with components of the transcription machinery to facilitate activated transcription. Activators may target one or more general transcription factors to recruit them to the preinitiation complex, to stabilize the preinitiation complex, or to alter functional activities of the preinitiation complex. Direct and indirect interactions of activators with the general transcription factors TFIIA, TFIIIB and TFIID, that are assembled early into the preinitiation complex (Buratowski et al., 1989; Buratowski, 1994), have been demonstrated (Stringer et al., 1990; Lin and Green, 1991; Chen et al., 1994; Ge and Roeder 1994; Kretzschmar et al., 1994). Biochemical evidence has suggested that these
interactions may result in accelerating the formation of a preinitiation complex on promoter DNA and the formation of an open complex. This effect subsequently increases the rate of transcription initiation (Zawel and Reinberg, 1992; Maldonado and Reinberg, 1995).

While substantial progress has been made in understanding the mechanisms of enhancing initiation of transcription by activators, the mechanism of controlling RNA polymerase II elongation is not as well understood. Tat has been used to study transcriptional elongation because it is able to stimulate elongation without any apparent effect on initiation in some cellular assay systems and in some cell free systems (Jones and Peterlin, 1994). Much evidence has led to a model for the action of Tat in which TAR-bound Tat on a paused promoter-proximal elongation complex "reaches back" to alter the elongation potential of the subsequently-formed transcription complex, either during assembly of the preinitiation complex or during promoter clearance (Cullen, 1993; Jones and Peterlin, 1994). This model is supported by the demonstrated interactions of Tat with TBP (Kashanchi et al., 1994) and a specific TAF (Chiang et al., 1995). However, the TBP-Tat interaction may only enhance the binding of TBP to the TATA-box DNA element of the HIV-1 promoter to stimulate the rate of initiation (Kashanchi et al., 1994). Our finding that VP16 and p53 interact with the general factor TFIIH (Chapter II), which appears to be involved in promoter clearance and subsequent elongation events (Goodrich and Tjian, 1994; Bentley, 1995), led us to propose that TFIIH may mediate the stimulation of elongation by activators. Here, we have examined binding of various activators to TFIIH. We demonstrate that the ability of activators to stimulate elongation correlates with binding of TFIIH.
C. Materials and methods

Plasmids.

The plasmid pGEX-3X-Tat was constructed by inserting a Tat gene that had been PCR-amplified from p tat72 (Frankel and Pabo, 1988) into the BamHI site of pGEX-3X (Pharmacia). pGEX-3X-Tat (1-48) was constructed by digesting pGEX-3X-Tat with EagI and EcoRI, filling in the ends with Klenow enzyme and recircularizing. pAE 38 and pAE 40 expressing GST-SpI A and GST-Sp1 B were described previously (Emili et al., 1994). pGX20T-CTF contains a DNA fragment encoding residues 399-499 of CTF. pET1Id-p89 was constructed by inserting ERCC3 cDNA, isolated from the plasmid pCD1 (Weeda et al., 1990), between the NcoI and BamHI sites of pET1Id (Novagen). pET11a-BTF2 was described previously (Fischer et al., 1992).

Affinity chromatography.

The affinity columns containing GST and GST fusion proteins were prepared as described previously (Chapter II). 800 µl aliquots of HeLa whole cell extracts in 0.1 M NaCl ACB buffer (Sopta et al., 1985) or 200 µl of HeLa nuclear extract in 0.1 M KCl buffer D (Chapter II) were chromatographed on affinity columns (40 µl) containing 4 mg/ml of GST or GST fusion proteins (except for GST-SpI A and GST-Sp1B columns, which contained 6 mg/ml). The columns were washed with 400 µl of 0.1 M NaCl in ACB buffer or 0.1 M KCl in buffer D. Bound proteins were eluted with 160 µl of 1 M NaCl in ACB buffer or 160 µl of 1 M KCl in buffer D and, subsequently, 160 µl of 1% SDS. Eluted samples (20 µl) were separated by SDS-PAGE, immunoblotted with antibodies, and visualized by ECL (Amersham). In Figure 3, a TFIIH fraction (100 µl) (Xiao et al., 1994) was passed through GST, GST-Tat (1-48) and GST-VP16 affinity columns (40 µl) in 0.1 M KCl buffer D. Aliquots of the flowthrough fractions (4 µl) were assayed for TFIIH activity in a reconstituted transcription system containing general transcription factors and RNA polymerase II, but lacking TFIIH (Xiao et al., 1994).

Antibodies and Western blotting.
Anti-p62 monoclonal antibody 3c9 was kindly provided by J.-M. Egly. Anti-human CDK7 antibody (C/9), Anti-TBP (SI-1) and Anti-p89 were purchased from Santa Cruz Inc.. Western blotting was performed using the ECL Western blotting detection system (Amersham).

**In vitro protein-binding assays.**

$^{35}$S-labeled p62 and p89 were synthesized by in vitro translation (Promega) using plasmids pET11a-BTF2 (Fischer et al., 1992) and pET11d-p89. Binding experiments were performed as described previously (Xiao et al., 1994).

**D. Results**

Previous studies suggested that the Tat, VP16, and p53 activation domains may stimulate elongation by targeting a common component of the transcription machinery which is not contacted by Sp1 or CTF, two factors which increase only the rate of initiation (Blau et al., 1996; Zhou and Sharp, 1995). To test this idea, we performed affinity chromatography with the immobilized activation domains of Tat and CTF1 and both activation domains (A and B) of Sp1 as well as with full length Tat. GST-fusion proteins were produced in *E. coli*, and purified by glutathione-Sepharose chromatography. The VP16 and p53 activation domains were previously shown to bind TFIIH and to make direct contact with its p62 subunit (Chapter II). As controls, we also performed affinity chromatography with GST and with the immobilized activation domain of VP16. HeLa whole cell extract was chromatographed on affinity columns under identical conditions. The high salt and SDS eluates from the columns were immunoblotted with a monoclonal antibody (3c9) against the p62 subunit of TFIIH. As shown in Figure 1, an immunoreactive 62 kDa polypeptide, identical in size to p62 in HeLa cell extract (lanes 1 and 6), was detected in the 1 M NaCl eluates from affinity columns containing full length Tat (lane 10) or the positive control VP16 (lanes 5 and 9). Moreover, p62 also bound to a GST-Tat column containing only the activation domain (amino acids 1-48) of Tat (lane 11). In contrast, no immunoreactive polypeptide was detected in the eluates from columns containing the activation domain A of Sp1 (lane 3), activation domain B of Sp1 (lane 4), the activation domain of CTF (lane 8), or the negative control GST (lanes 2 and 7).
Substitutions of Cys residues at positions 18, 22 and 37 in the Tat activation domain abolish transactivation by Tat in vivo and in vitro (Rice and Carlotti, 1990; Southgate and Green, 1991). To assess the significance of the interaction of TFIIH with the activation domain of Tat, we examined the effect on the binding of TFIIH of a single amino acid substitution (Cys to Ala at residue 22) in the activation domain of Tat. Protein affinity columns containing immobilized GST-Tat (1-48), GST-Tat (1-48, C22A), and GST were prepared, all containing the same ligand concentration. HeLa nuclear extract was chromatographed on the columns and the eluates were immunoblotted with monoclonal antibody 3c9 against p62, a polyclonal antibody against CDK7, the kinase subunit of TFIIH (Roy et al., 1994a; Shiekhattar et al., 1995; Serizawa et al., 1995) and anti-TBP antibody. In agreement with previous reports in which Tat was demonstrated to interact with TBP, TBP was indeed found in the eluate from the affinity column containing the activation domain of Tat, but not in the eluate from the column containing the mutated Tat activation domain (Figure 2, lanes 3 and 4). Similarly, the C22A mutation also reduced the binding of CDK7 to Tat. However, this mutation had a much more modest effect on the binding of p62 (compare lane 3 with lane 4). This result suggested that the Tat-TFIIH interaction may be biologically important. We propose that the defect of mutant Tat proteins in transactivation may result partly from defects in their binding to CDK7 and TFIID.

To determine whether other general transcription factors are required for the binding of TFIIH to Tat, we tested the ability of Tat to bind a highly purified preparation of TFIIH. A TFIIH fraction that lacks other general transcription factors and RNA polymerase II (Chapter II) was passed through affinity columns containing immobilized GST, GST-Tat (1-48) and GST-VP16. Flow-through fractions were assayed for the transcriptional activity of TFIIH. The transcriptional activity of TFIIH was partially depleted by the Tat and VP16 activation domain columns (Figure 3, lanes 3 and 4), but not by the control GST column (lane 2). This result suggested that Tat binds directly to TFIIH.

When the p62 and p89 subunits of TFIIH are made by translation in vitro in a rabbit reticulocyte lysate, both of them bind to the activation domain of Tat (Figure 4, lanes 3, 5, 8 and
9) and much less to GST control columns (lanes 2, 4, 7 and 9). This indicates that Tat could bind both p62 and p89 subunits of TFIIH. Therefore, these subunits may mediate an interaction of Tat with TFIIH.
TFIIH binds to the Tat and VP16 activation domains, but not the Sp1 and CTF activation domains. HeLa cell proteins were chromatographed on the indicated affinity columns and eluted with buffer containing 1 M NaCl (lanes 2-5, 7-11) and subsequently with buffer containing 1% SDS. Aliquots (20 µl) of the eluted samples were analyzed by SDS-PAGE and immunoblotting with monoclonal antibody 3c9 against p62. The input HeLa extracts are shown in lanes 1 and 6. Only the salt eluates are shown in the figure. No significant amount of p62 was detected in the SDS eluates from the GST, GST-Sp1A, GST-Sp1B and GST-CTF columns (data not shown).
Figure 2. Effect of a mutation in the activation domain of Tat on the binding of TFIIH. HeLa proteins (lane 1) were chromatographed on GST, GST-Tat (1-48) and GST-Tat (1-48 C22A) affinity columns and eluted from the columns with 1 M KCl in buffer D. The eluates were analyzed by SDS-PAGE and immunoblotted with antibodies against p62 (3c9), CDK7 (c19) and TBP(S1-1).
Figure 3. Binding of highly purified TFIIH to the activation domain of Tat. TFIIH was loaded on the indicated columns. The flow-through fractions from the GST, GST-Tat (1-48) and GST-VP16 affinity columns (lanes 2-4), or an equal volume of buffer D (lane 1), were assayed for TFIIH activity in a reconstituted transcription system.
Figure 4. Binding of the p62 and p89 subunits of TFIID to the activation domain of Tat. Lanes 1 and 6 show the input in vitro translated 35S-methionine-labeled p89 and p62 that were chromatographed on GST and GST-Tat (1-48) columns. Bound proteins were eluted from the GST and GST-Tat (1-48) beads first with 1M NaCl in ACB buffer (lanes 2, 3, 7, and 8) and then by boiling in SDS-PAGE sample buffer (lanes 4, 5, 9 and 10).
E. Discussion

We have demonstrated that TFIIH interacts with the activation domains of several activators including p53 and VP16 (Chapter II), E2F1 (Pearson and Greenblatt, unpublished data) and Tat. All of these activators stimulate transcriptional elongation (Yankulov et al., 1994, Blau et al., 1996; Zhou and Sharp, 1995). On the other hand, the glutamine-rich activation domains A and B of Sp1 (Courey and Tjian, 1988) and the proline-rich activation domain of CTF (Mermod et al., 1989), both of which do not stimulate elongation (Blau et al., 1996), did not bind TFIIH. This result provides strong evidence for TFIIH as a target for controlling elongation by activators. It also raises the question of how the interaction of TFIIH with activation domains could stimulate elongation. This interaction could facilitate assembly of an activated pre-initiation complex by stabilizing the binding of TFIIH. Alternatively, the binding of TFIIH to an activation domain could modulate one of the enzymatic activities of TFIIH. In particular, the Cdk component of TFIIH phosphorylates the RNA polymerase II CTD (Lu et al., 1992) and the large subunits of TFIIF and TFIIE (Ohkuma and Roeder, 1994).

TFIIF is not only an initiation factor but also an elongation factor (Greenblatt, 1991a). The phosphorylation of the large subunit of TFIIF (RAP74) was demonstrated to stimulate transcriptional elongation by RNA polymerase II (Kitajima et al., 1994). Moreover, Tat was shown to cooperate with TFIIF to stimulate the processivity of RNA polymerase II (Kato et al., 1992). It is, therefore, possible that Tat may stimulate elongation by interacting with TFIIH to enhance the phosphorylation of RAP74. It is conceivable that phosphorylation of RAP74 may relieve an inhibitory effect of the carboxy-terminal domain of RAP74 on the elongation-enhancing properties of TFIIF (Kitajima et al., 1994).

In several Drosophila genes, CTD phosphorylation is correlated with the release of polymerase stalled at the 5' ends of nascent transcripts (Weeks et al., 1993; O'Brien et al., 1994). As well, activation by the acidic activator protein GAL4-VP16 (Johnston and Dover, 1987), which stimulates elongation (Yankulov et al., 1994), depends on the CTD (Lieu et al., 1992). In contrast, the glutamine-rich activator protein Sp1, which does not stimulate elongation,
does not require the CTD for increasing transcription (Gerber et al., 1995; Zehring and Greenleaf, 1990). These observations suggest that phosphorylation of the CTD may play an important role in controlling elongation.

Dichlororibofuranosylbenzimidazole (DRB), an inhibitor of RNA polymerase II elongation (Chodosh et al., 1989; Marshall and Price, 1992; Robert and Bentley, 1992), inhibits CTD phosphorylation in vivo and also inhibits the processivity of RNA polymerase II and Tat-activated transcription (Marciniak and Sharp, 1991). Furthermore, DRB inhibits the CTD kinase activity of highly purified TFIIH with the same dose-response curve as for inhibition of elongation (Yankulov et al., 1995). This evidence, together with our evidence that various activators bind TFIIH, leads us to propose that the interaction of activators with TFIIH stimulates elongation by recruiting and/or stimulating the kinase activity of TFIIH which in turn phosphorylates the CTD. This model is consistent with recent evidence from the Reinberg laboratory that VP16 reduces the rate of abortive synthesis by facilitating promoter clearance via TFIIE and TFIIH (Maldonado and Reinberg, 1995).

Tat has been reported to induce apoptosis in uninfected lymphocytes through activation of Cdk activities (Li et al., 1995). It is not clear how Tat activates Cdns. Cdc2 and Cdk2 kinases can be activated by a CDK7/Cyclin H complex (Fesquet et al., 1993; Poon et al., 1993) which is a component of TFIIH (Roy et al., 1994a). Therefore, the interaction of Tat with TFIIH may be involved in Tat-mediated apoptosis. In this study we have shown that Tat binds p89, one helicase of TFIIH which is also involved in nucleotide excision repair and possibly apoptosis. Like Tat, the p53 tumor suppressor was previously demonstrated to bind p89 and p62 (Chapter II: Wang et al., 1994). The p53-TFIIH interaction was proposed to be involved in transcriptional activation (Chapter II) as well as p53-mediated apoptosis (Wang et al., 1996). Combining these and other data, we propose that Tat stimulates the activity of TFIIH, thereby inducing apoptosis in HIV-1-infected cells.
Chapter IV

Identification of human cellular proteins that interact with the activation domain of Tat

In this chapter, the immunofluorescent staining in figure 3 was performed by Pierre S. Tung
Summary

The HIV-1 Tat protein is a potent activator of transcription initiated at the viral LTR and is essential for viral growth. Transcriptional activation by HIV-1 Tat requires Sp1 and the general transcriptional machinery as well as at least one human-specific cofactor. This putative cofactor(s) is likely to interact with the amino-terminal portion of Tat, thereby mediating transactivation by Tat. By using Tat protein affinity chromatography, we have identified two human proteins (57 and 30 kDa) that bind to the activation domain of Tat (amino acids 1-38). One (57 kDa) is the human cellular protein CD46 and the other (30 kDa) is a novel human protein for which there is an existing cDNA. Point mutations in the amino-terminal portion of Tat that reduce its transactivation activity also weaken its binding of CD46 and the 30 kDa protein. Moreover, CD46 can bind to both of the glutamine-rich activation domains of Sp1. Coexpression of CD46 in human HeLa cells, which already contain CD46, further increased transcriptional activation by Tat of the HIV-1-LTR. We therefore propose that Tat and Sp1 form a complex mediated by CD46 to increase the efficiency of transcription.
Introduction

The human immunodeficiency virus 1 (HIV-1), the etiologic agent of acquired immunodeficiency syndrome (AIDS), encodes a number of regulatory proteins. One of them, called Tat, is required primarily to increase the rate of transcription from the viral long terminal repeat (LTR) during viral growth (Dayton et al., 1986; Fisher et al., 1986) and may also regulate expression of cellular genes and mediate apoptosis (Li et al., 1995). Tat can be taken up by HeLa cells and localizes rapidly to the nucleus (Green and Loewnstein, 1988; Frankel and Pabo, 1988).

Despite extensive study, the mechanism of Tat-responsive action is still poorly understood. In most cellular assay systems and in cell free systems the major effects of Tat are on transcription elongation, but effects on the rate of initiation have also been reported (Cullen, 1993; Greenblatt et al., 1993). Transactivation by Tat is critically dependent upon upstream Sp1 binding sites and on an RNA sequence, called TAR, which is located downstream of the site of transcription initiation (Rosen et al., 1985). The Sp1-binding sites interact with the cellular transactivator Sp1 that cooperates with Tat to stimulate transcription. Tat can also function when bound to DNA as a GAL4 fusion protein (Southgate and Green, 1991). Genetic and biochemical studies have suggested that a human-specific factor acts as a Tat cofactor to facilitate binding of Tat to TAR and mediates transcriptional activation (Madore and Cullen, 1993). This putative cofactor is probably encoded by a gene on human chromosome 12 because human chromosome 12 greatly stimulates Tat-responsive activation in rodent cells (Newstein et al., 1990; Hart et al., 1989). Furthermore, high levels of the amino-terminal portion of Tat can squelch transactivation by Tat and chimeric Tat proteins in vivo and in vitro, suggesting that at least one cofactor may interact with the amino-terminal portion of Tat. A partially purified cofactor can support Tat-activated transcription in an in vitro transcription system containing Sp1, highly purified general transcription factors and RNA polymerase II (Zhou and Sharp, 1995). Several candidate cofactors, including TBP1 (Nelbock et al., 1990), MSS1 (Shibuya et al., 1992), TRP-1 (Wu et al., 1991; Sheline et al., 1991), and a 42 kDa kinase (Hermann and Rice, 1993) have been
identified. Although these proteins have some of the expected characteristics of a coactivator, none has been determined to interact directly with the amino-terminal portion of Tat or to participate directly in the action of Tat as a cofactor. A cellular acidic activator, Tat-associated protein (TAP), has been shown to interact directly with the core domain of Tat and the general transcription factor TFIIB (Yu et al., 1995a; Yu et al., 1995b). However, since TAP is conserved among eukaryotes and is expressed in human and rodent cells, it can be excluded as a species-specific cofactor for Tat. In this study, we have identified two human proteins that bind to the amino-terminal 48 residues of Tat. Our data suggest that one of them, CD46, may act as a coactivator to mediate Tat-responsive activation.
MATERIALS AND METHODS

**Plasmids.** The plasmid pGEX-3X-Tat was constructed by inserting a Tat gene PCR-amplified from p tat72 (Frankel and Pabo, 1988) into the BamHI site of pGEX-3X (Pharmacia). pGEX-3X-Tat (1-48) was constructed by digesting pGEX-3X-Tat with EagI and EcoRI, filling in the ends with Klenow enzyme and recircularizing. The pGEX-3x-Tat (1-48K41T), pGEX-3x-Tat (1-48F38A), and pGEX-3x-Tat (1-48C22A) used for expression of mutant GST-Tat proteins were constructed by digesting the plasmids pGEX-3x-Tat K41T, pGEX-3x-TatF38A, and pGEX-3x-TatC22A with EagI and EcoRI, filling in and recircularizing. To construct the series of mutant GST-Tat plasmids, Tat coding sequences were first cloned into pBluescript (+) (Stratagene) and mutations were introduced at positions 41, 38 and 22 by using oligonucleotide-directed mutagenesis (Amersham). Afterwards, the Tat coding sequences containing mutations were cloned into pGEX-3X in-frame with the glutathione S-transferase gene. pAE 38 and pAE 40 expressing GST-Sp1A and GST-Sp1B were described previously (Emili et al., 1994). p167 was constructed by Rosen et al. (1985). pSV-HIV-tat was constructed by Emerman et al. (1987). pSVL was purchased from Pharmacia. pCEP4 was purchased from Invitrogen. pCEP4-CD46 was described previously (Dorig et al., 1993).

**Affinity chromatography.** Affinity columns (40 μl) containing GST, GST-Sp1A, GST-Sp1B, GST-VP16, GST-Tat, GST-Tat (1-48) and GST-Tat mutant proteins were prepared by binding GST or GST fusion proteins (4 mg/ml each) on glutathione-sepharose beads as described (Xiao et al., 1994). Affinity chromatography was performed as described previously (Stringer et al., 1990). However, bound proteins were eluted with ACB buffer containing 1 M NaCl. To purify Tat-binding proteins on a large scale, 100 ml of Hela whole cell extract (5 mg/ml protein) were chromatographed on a 5 ml GST-Tat (1-48) affinity column. The bound proteins were eluted with 20 ml of 1 M NaCl in ACB and dialyzed against 0.1 M NaCl in ACB. Aliquots of the dialyzed bound proteins were used as the input for Figure 5b. 15 ml of the dialyzed bound proteins were loaded onto a 0.5 ml phosphocellulose column and step eluted with 2.5 ml each 0.3, 0.5 and 0.85 M NaCl in ACB. 20 μl aliquots of the eluted fractions (0.5
were first analyzed by SDS-PAGE and stained with silver. The 57 kDa protein was present in the 0.5 M NaCl fraction and the the 30 kDa protein was present in the 0.3 M NaCl fraction. The 0.5 M NaCl fractions and 0.3 M NaCl fractions were pooled separately. The proteins were precipitated with TCA and loaded onto an SDS-polyacrylamide gel. The 57 kDa and 30 kDa bands were visualized by staining with Commassie blue and excised from the gel. The proteins were digested with endoproteinase C. The resulting peptides were isolated by HPLC and subjected to microsequencing by Ruiji Kobayashi (Cold Spring Harbor Laboratory).

**Antibodies and Western blotting.** Western blotting was performed with a monoclonal antibody against CD46 (Immunogene Inc.)

**Immunofluorescent staining.** Human cells grown on slides were incubated for a hour with a monoclonal antibody (1:50) against CD46 (Immunogene Inc.) followed by fluorescein isothiocyanate-conjugated goat anti-mouse IgG. The cells were visualized by immunofluorescence microscopy and photographed (P. S. Tung and M. F. Moran, unpublished procedure).

**Cotransfection and CAT assays.** HeLa cells were grown in α-MEM supplemented with 10% fetal calf serum in a 37°C, 5% CO2 incubator. Cells were split into 100 mm tissue culture dishes to about 50% confluency the night before transfection. Transfections were done by the calcium phosphate precipitation method as described (Chen and Okayama, 1988). Briefly, 1 µg of the plasmid p167 was coprecipitated with 2 to 4 µg of various plasmid DNAs. 2 µg of pRSVlacZ (Edlund et al., 1985) was included in each transfection to assess the transfection efficiency. pBluescript (Stratagene) was added to each transfection so that the final concentration of DNA being transfected remained the same. A DNA precipitate was allowed to form at room temperature for 25 minutes, then added to the cells. After 16-24 hours, cells were washed three times with phosphate-buffered saline (PBS), fresh medium was added and the cells were further incubated for 24-48 hours. Cells were harvested and lysed by freeze-thaw cycles and assayed first for β-galactosidase activity. Normalized quantities of cell extracts were then used to perform CAT assays and the products of the reactions were analyzed by thin-layer chromatography.
(Suen and Hung, 1991). Experiments were repeated three times and the standard error was less than 10%.
Results

Identification of a 57 kDa protein that binds selectively to the amino-terminal 48 residues of Tat.

The structure of the Tat protein has been studied intensively. Its TAR-binding domain and nuclear localization signal lie in the carboxy-terminal residues 49-58. The amino-terminal 48 amino acids constitute an activation domain sufficient to support transcriptional activation when fused to the DNA-binding domain of GAL4 (Southgate and Green, 1991). HIV-1 Tat (72 amino acids) and its amino-terminal 48 amino acids were fused to GST and expressed in E. coli. The fusion proteins were bound to Glutathione-Sepharose and used as ligands for protein affinity chromatography. A soluble HeLa whole cell extract was chromatographed on the columns. Bound proteins were eluted with buffer containing 1M NaCl, electrophoresed on a SDS-polyacrylamide gel, and stained with silver. Two polypeptides with the apparent molecular masses of 57 kDa and 30 kDa selectively bound to the GST-Tat (1-48) (Figure 1, lane 3) and GST-Tat (lane 2) columns, but not the GST column (lane 1). There were also many proteins that bound only to GST-Tat and not to GST-Tat (1-48) (Figure 1, lanes 2 and 3). These proteins bound to the basic RNA-binding domain of Tat and not to its activation domain. Since HIV-1 Tat is a multi-functional protein that regulates, not only the transcription of HIV-1 genes, but also other nontranscriptional processes (Huang et al., 1994), proteins that bind to intact Tat, but not its activation domain, may be involved in the other roles of Tat or may result from a nonspecific interaction with the highly basic region of Tat. When a similar affinity chromatography experiment was done with nuclear extract rather than whole cell extract (Figure 2A), 57 kDa and 30 kDa proteins again bound to the Tat activation domain, indicating that both might be nuclear proteins.

The 57 kDa polypeptide was isolated on a large scale and subjected to proteinase digestion. The small peptides were further isolated by HPLC and microsequenced. Two peptide sequences (KPYYEI and AEYATY) obtained from the 57 kDa polypeptide were identical to
sequences in the human membrane cofactor protein, known as CD46 (Cole et al., 1985; Lublin et al., 1988).

To confirm that the 57 kDa protein was CD46, we next immunoblotted the eluates from GST and GST-Tat (1-48) columns with a monoclonal antibody against CD46 (Figure 2B). An immunoreactive polypeptide, identical in size to CD46 in the nuclear extract (lane 1), was detected in the eluate from the affinity column containing the Tat activation domain (lane 3) but not in the eluate from the control column containing GST (lane 2).

CD46 was previously characterized as an integral membrane protein on the surface of most human cells. To localize CD46 in situ, a monoclonal antibody against CD46 and a fluorescein isothiocyanate-conjugated, polyclonal goat anti-mouse IgG antibody were used to stain human cells. By using immunofluorescence microscopy, CD46 was detected in the periphery of the nucleus and possibly in the nucleus of these human fibroblast cells (Figure 3). Since a soluble form of CD46 can be found in nuclear extracts and whole cell extracts, we conclude that a form of CD46 is localized in the perinuclear region of the cytosol and perhaps in the nucleus.

We also obtained five peptide sequences from the 30 kDa protein by the same procedures as the 57 kDa protein (data not shown). The five peptide sequences were identical to sequences in a novel protein encoded by a randomly cloned human cDNA existing in GeneBank (data not shown). This protein has not yet been further characterized.

**CD46 and the 30 kDa protein specifically interact with the activation domain of Tat.**

To determine whether these Tat-interacting proteins may be involved in Tat-responsive transactivation, we tested whether amino acid substitutions in the activation domain of Tat, which weaken transcriptional activation by GAL4-Tat (1-48), also reduce the binding of Tat-interacting proteins. Three GST-Tat (1-48) fusion proteins containing the point mutations K41T, F38A and C22A were prepared as ligands for affinity chromatography. As shown in Figure 4, all three mutations reduced the binding of CD46, while F41A and F38A reduced the binding of
kDa protein (compare lane 2 with lanes 3, 4 and 5). Thus, we conclude that interactions between the activation domain of Tat and these proteins may be biologically important.

**The activation domains of Sp1 also interact with CD46.**

We wished to determine whether other types of activation domains also interact with CD46. Because VP16 did not efficiently squelch transcriptional activation by a GAL4-Tat fusion protein, in contrast to Tat, and vice versa (Song et al., 1994, Madore and Cullen, 1993), the activation domain of VP16 is presumed not to bind a Tat-specific coactivator. In an initial experiment, we compared the HeLa cell proteins that bind to the glutamine-rich activation domains A and B of Sp1 with those that bind to the activation domains of VP16 and Tat. Each of those activation domains was produced as a GST fusion protein and attached to glutathione-Sepharose beads. As an arrow indicates in Figure 5A, both activation domains of Sp1 (lanes 2 and 3) bound a human protein identical in size to the CD46 that was bound by the activation domain of Tat (lane 5). No protein of a similar size was bound by the activation domain of VP16 (lane 4). To confirm that the activation domains of Sp1 can interact with CD46, we chromatographed the same amounts of bound fraction from a GST-Tat (1-48) affinity column on GST, GST-Sp1A and GST-Sp1B affinity columns (Figure 5B). CD46 was not detected in eluate from the GST column (lane 2), but was found to be retained on both GST-Sp1A and GST-Sp1B columns (lanes 3 and 4). This result demonstrated that the activation domains of Sp1 also interact with CD46.

**Involvement of CD46 in transactivation by Tat.**

The binding of CD46 to the activation domains of Tat and Sp1, and the correlated effects of mutations in the activation domain of Tat on transactivation by Tat and the binding of CD46, suggested a possible role of CD46 in transcriptional activation by Tat. To examine this possibility, the plasmid pCEP4-CD46, which contains a CD46 cDNA under the control of a CMV promoter, was cotransfected into HeLa cells with a plasmid containing a chloramphenicol acetyltransferase (CAT) gene under control of the HIV long terminal repeat (LTR) and another plasmid, pSV-tat, which expresses HIV-1 Tat. No stimulation of the expression of the CAT
gene was observed when CD46 was expressed alone in the absence of Tat (Figure 6, lane 7). In contrast, the expression of the CAT gene was increased more than 50-fold by cotransfection with the Tat expression plasmid (lane 3). This activity was further increased by about 4-fold when pCEP4-CD46 was cotransfected with pSV-tat (compare lanes 10 and 11). This modest increase in Tat-activated expression of the CAT gene upon cotransfection with the CD46 gene has been consistently observed in several experiments. By contrast, the expression vectors pSVL (lane 2) and pCEP4 (lanes 4, 8) had no any significant effect on CAT activity.

We also tested whether CD46 can increase Tat transactivation in mouse cell lines and Chinese hamster ovary cell lines (CHO). No more than a 2-fold increase in the expression of CAT from HIV-LTR-CAT was observed when pCEP4-CD46 was cotransfected with pSV-HIV-tat (data not shown). Rodent cells may be missing some factor other than CD46 which is needed for transactivation by Tat. Since we have shown that a 30 kDa protein also binds specifically to the activation domain of Tat, it is possible that the 30 kDa protein is also required for transcriptional activation by Tat.
Figure 1. The activation domain of Tat and intact Tat bind to a 57 kDa protein and other proteins. Aliquots of Hela whole cell extract were chromatographed on affinity columns containing GST, GST-Tat, and GST-Tat (1-48). The bound proteins were subjected to SDS-PAGE on a 10% gel and stained with silver. The arrow indicates the 57 kDa protein. The 30 kDa protein is barely visible in this experiment.
Figure 2. The activation domain of Tat interacts with CD46 from nuclear extracts. A. Aliquots of Hela nuclear extract were chromatographed on affinity columns containing GST or GST-Tat (1-48). The bound proteins eluted from the GST column (lane 1) and GST-Tat (1-48) column (lane 2) were subjected to SDS-PAGE on a 10% gel. The gel was stained with silver. The arrow indicates the 57 kDa protein. B. The bound proteins used for Figure 2A were also immunoblotted with a monoclonal antibody against CD46. 10 µl of nuclear extract was loaded in lane 1. An arrow indicates CD46.
Figure 3. Immunofluorescent staining of CD46 \textit{in situ}. Human cells were stained with a monoclonal antibody against CD46 and fluorescein isothiocyanate-conjugated goat anti-mouse IgG. The cells were visualized by immunofluorescence microscopy and photographed.
Figure 4. Mutations in the activation domain of Tat reduce the binding of CD46. Aliquots of Hela whole cell extracts were chromatographed on affinity columns containing GST, GST-Tat (1-48), or the various indicated GST-Tat (1-48) mutant proteins. The bound proteins eluted from the various columns, as indicated at the top of panel, were subjected to SDS-PAGE on 10% gels and stained with silver. Arrows indicate the 57 kDa CD46 and the 30 kDa protein.
GST
GST-Sp1A
GST-Sp1B
GST-VP16
GST-Tat1-48

Input
GST
GST-Sp1A
GST-Sp1B
Figure 5. CD46 also binds to the activation domains of Sp1. A. Aliquots of Hela whole cell extract were chromatographed on affinity columns containing GST, GST-Sp1A, GST- Sp1B, GST-VP16, and GST-Tat (1-48). The bound proteins eluted from the various columns, as indicated at the top of the panel, were subjected to SDS-PAGE on a 10% gel and stained with silver. An arrow indicates the 57 kDa CD46 protein. The 30 kDa Tat binding protein was not detected in this experiment. B. The bound proteins eluted from a GST-Tat (1-48) affinity column (lane 1) were rechromatographed on affinity columns containing GST, GST-Sp1A or GST-Sp1B. The bound proteins eluted from GST (lane 2), GST-Sp1A (lane 3) or GST-Sp1B (lane 4) were subjected to SDS-PAGE on a 10% gel and stained with silver. An arrow indicates CD46.
Figure 6. Effect of CD46 on transcriptional activation by Tat in transfection assays. The plasmid p167, which contains HIV1LTR-CAT (Rosen et al., 1985), was cotransfected with the indicated expression plasmids into HeLa cells. Cell lysates were first assayed for β-galactosidase activity. Normalized quantities of cell extracts were then analyzed by thin-layer CAT assays. The figure shows a typical set of CAT assays. Each transfection contained 1 µg of p167 and 4 µg of pBluescript (lane 1), 4 µg of pSVL (lane 2), 4 µg of pSV-tat (lane 3), 2 µg of pSV-tat and 2 µg of pCEP4 (lane 4), 2 µg of pSV-tat and 2 µg of pCEP4-CD46 (lane 5), 2 µg of pCEP4-CD46 and 2 µg of pBluescript (lane 6), 4 µg of pCEP4-CD46 (lane 7), and 4 µg of pCEP4 (lane 8). The CAT assay reactions in lanes 9, 10 and 11 were carried out for short times so that each resulting activity would be in the linear range: lane 9, 4 µg of pSVL; lane 10, 2 µg of pSV-tat and 2 µg of pCEP4; lane 11, 2 µg of pSV-tat and 2 µg of pCEP4-CD46.
Discussion

We have demonstrated that the activation domain of Tat specifically interacts with a human regulator of complement activation known as membrane cofactor protein (MCP) or CD46. CD46 is an integral membrane protein on the surface of most human cells. It functions as an inhibitor of deposition of complement proteins C3b and C4b on the host cells and as a receptor for Measles virus (Liszewski et al., 1991; Dörig et al., 1994). CD46 is expressed in cells as four isoforms with two different cytoplasmic tails of 16 (CYT-1) or 23 amino acids (CYT-2). Since one of the peptide sequences which we obtained was found to be identical to CYT-2, we conclude that Tat interacts with the CYT-2 isoform of CD46. The CYT-2 isoform was found to contain a nuclear localization signal and have a much slower processing rate (Liszewski et al., 1994). Although the biological significance of the Tat-CD46 interaction remains to be elucidated, we have considered two possible explanations for the interaction between Tat and CD46. One of the explanations is that CD46 may transport Tat into the nucleus. Tat protein can be taken up by cells and localized quickly to the nucleus (Frankel and Pabo, 1988). It has been proposed that the vitronectin-binding integrin, αvβ5, CD26, and a 90 kDa protein were the cell surface proteins that bind to Tat (Vogel et al., 1993; Gutheil et al., 1994; Weeks et al., 1993). These proteins may also play roles in Tat internalization by cells. Thus far, it remains unclear how Tat enters the nucleus. In this study, we have used a monoclonal antibody directed against CD46 to localize CD46 in human cells in situ. CD46 has been localized on the surface, in the perinuclear region of the cytosol and perhaps in the nucleus. Therefore, it is very possible that CD46 controls the transportation of Tat from the cytosol to the nucleus via associating with Tat. To prove this hypothesis would require more experiments to examine the possible co-localization of Tat and CD46 in situ and the possible enhancement of the entry of Tat into the nucleus by increasing the expression of CD46 in the cell.

An alternative explanation is that this interaction may be involved in directly controlling transcription. First, the binding site in Tat for CD46 is located in its activation domain; mutations that decrease the transactivation activity of Tat also reduce the binding of CD46. Secondly, the
Secondly, the cellular transcription factor Sp1, which is essential for Tat to stimulate transcription of the promoter in the HIV-1 LTR, also binds to CD46. The synergistic effects of Sp1 and Tat on HIV-1 transcription depend critically on the spacing between the Sp1 binding sites and the TATA element within HIV-1 promoter. This could imply that there is a physical interaction between Tat and Sp1 (Huang and Jeang, 1993). Although a direct interaction of Sp1 with Tat has been demonstrated (Jeang et al., 1993), it does not exclude the possibility that a cellular protein also mediates the interaction of Sp1 with Tat. The Sp1-binding domain of Tat lies in a region encompassing residues 30 to 60, while the region of Tat for binding CD46 is in the amino-terminal 48 residues. Therefore, CD46 may function as an adapter to stabilize the Sp1-Tat interaction. Alternatively, Sp1 and Tat may cooperate with one other to recruit CD46 into the transcriptional initiation complex. Finally, our cotransfection results showed that overexpression of CD46 in HeLa cells did not affect the function of the HIV-1 LTR in the absence of Tat but further increased stimulation of the HIV-1 LTR by Tat.

The fact that CD46 is expressed in most primate cells, but not in rodent cells, is consistent with the finding that Tat functions as a potent transactivator in primate cells, but not rodent cells (Newstein et al., 1990; Hart et al., 1989). However, CD46 is not the putative cofactor for Tat that is encoded by a gene localized on human chromosome 12 (Newstein et al., 1990; Hart et al., 1989) since the gene encoding CD46 is localized on human chromosome 1 (Dorig et al., 1993). It is possible that CD46 may function together with other Tat cofactor(s) to support transcription by Tat in HeLa, CHO or mouse cell lines. Recently, Sune and Garcia-Blanco (1995) demonstrated that a Tat affinity column depletes HeLa nuclear extracts of a factor(s) required for transcriptional activation by Tat in vitro. Transactivation by Tat can then be complemented by the addition of a fraction containing an activity with an apparent molecular mass of 200 kDa. If this fraction contains CD46, the CD46 may be in a complex with other polypeptides. These other protein(s) may act along with CD46 to function as a Tat cofactor. The 30 kDa protein which we found to also specifically interact with the activation domain of Tat may be another factor that is
also required for transcriptional activation by Tat. Determining whether CD46 and the 30 kDa protein are true cofactors for transactivation by Tat will require further experimentation.
Chapter V

Conclusions, perspectives and directions for future research
Eukaryotic gene expression is regulated primarily at the level of transcription by gene-specific transcriptional activators. The mechanisms for regulation of transcription can involve any of the steps in the process of transcription: initiation, promoter clearance, elongation and termination. Transcriptional activators, which act through the general transcription factors and RNA polymerase II, presumably function by promoting the assembly of functional preinitiation complexes on promoters to increase either initiation or a subsequent step such as promoter clearance or elongation. Both general initiation factors, including TBP, TFIIB, TFIIF and TFIIA, and various coactivators, such as TAFs and PC4, have been identified as activator targets for stimulation of initiation. The studies in this thesis have demonstrated that various activators, including HSV-1 VP16, HIV-1 Tat and the cellular tumor suppressor p53, interact with the general transcription factor TFIIF. The interaction of activators with TFIIF may be involved in the stimulation of transcription. This hypothesis can be tested by determining whether mutant forms of TFIIF which do not bind to activators are only defective in activated transcription. This will first require determining the exact domains of p62 and TFB1 that bind to activators. Then, the effects of mutations of TFB1 in yeast on initiation and elongation of transcription should be examined in vivo and in vitro.

Previous studies and the results presented in this thesis suggest that TFIIF is involved in enhancing transcriptional elongation by activators, but the underlying mechanism remains unknown. However, this could involve effects on phosphorylation of the CTD of the largest subunit of RNA polymerase II, since this event is correlated with the transition from initiation to productive elongation and since kinase inhibitors (DRB, H8) that block phosphorylation also block processive transcription. Studies have shown that TFIIF contains not only a helicase activity, which is involved in promoter clearance, but also a kinase activity, which phosphorylates the CTD in preinitiation complexes and is markedly stimulated by TFIIE. The kinase of TFIIF has been minimally identified as CDK7 and cyclin H. TFIIF can also phosphorylate TFIIF and TFIIE. While these observations suggest that TFIIF may be a potential target for activators that enhance processivity, the fact that presumptive CTD kinase inhibitors
block processive transcription only in crude extracts (and not in a system reconstituted with purified factors) predicts the involvement of additional negative factors that normally block elongation in the absence of CTD phosphorylation. Therefore, systems reconstituted with purified factors that are dependent upon both the CTD of RNA polymerase II and the TFIIH kinase for efficient elongation should be established. Establishing these systems will first entail the identification, purification and characterization of such a negative factor(s). The kinase inhibitors H8 and DRB, as well as antibodies specific for CDK7 and associated cyclin H, should be tested for their ability independently or jointly to inhibit TFIIH-mediated phosphorylation of the CTD and/or TFIIF and RNA polymerase processivity. The factor(s) which makes processivity dependent upon CTD/phosphorylation can be also tested for its effect on TFIIH-mediated phosphorylation of the CTD; and those elongation factors (e.g. TFIIS, SIII, TFIIF, or TFIIX) required for processive transcription can also be determined.

If phosphorylation is critical for RNA polymerase II processivity under physiological conditions, it is easy to imagine that activators might function by modulating phosphorylation. To explore this, the effects of Tat, p53 and VP16 on TFIIH-mediated phosphorylation of relevant factors (CTD, TFIIF and TFIIE) in both partially purified and highly purified systems, and both in the absence of and during preinitiation complex assembly and formation, can be tested. Special attention needs to be paid to possible cofactor requirements (e.g. Tat-associated factors, see below) and to the role of TFIIE and other TFIIH subunits (CDK7 and associated cyclin H as well as other subunits within the separable core TFIIH, Madonaldo and Reinberg, 1995). The relevance of the effects of activators on phosphorylation can be assayed using Tat and VP16 mutants deficient for enhancing processivity (Southgate and Green, 1991; Blau et al., 1996).

Studies presented in Chapter IV of this thesis demonstrate that CD46 binds specifically to the activation domain of Tat. The mutant Tat proteins that are deficient in transcriptional activation (Southgate and Green, 1991; Jones and Peterlin, 1994) reduce binding to CD46, suggesting that the interaction of the activation domain of Tat with CD46 is biologically relevant.
CD46 also interacts with the cellular factor Sp1, which is required for Tat-mediated activation. It is very possible that CD46 represents a bona fide Tat cofactor. Since one putative Tat cofactor can increase Tat-activated transcription in CHO cells, and CD46 is not expressed in CHO cells (Dörig et al., 1993), expression of CD46 could, in principle, enhance transcriptional activation by Tat in CHO cells. However, we have not been able to detect significant enhancement of the activity of Tat in CHO cells by co-transfecting the cells with a cDNA encoding CD46 (Dörig et al., 1993). Even if CD46 does contribute to activation by Tat in CHO cells, some other Tat cofactors may be missing in these cells.

The Tat co-factors were previously proposed to interact with Tat to facilitate binding to TAR and/or functional interactions with general factor targets. Further characterization of CD46 needs to be carried out in a reconstituted system using antibody raised against CD46. This would include a systematic analysis of physical and functional interactions with the other transcription components, especially TFIIB and TFIIF. Mobility shift assays can be performed to test the effect of CD46 on the binding of Tat to TAR. Possible regulation of cofactor function by covalent modification (e.g. phosphorylation) can also be analyzed by immunological and enzymatic assays.

Many oncogenes and antioncogenes encode transcription activators that control the processes of cell growth, differentiation and infection by viruses, and mutations in these genes often cause cancer. The tumor suppressor protein p53, oncogene product E2F1 and HIV-1 Tat have been shown to be capable of regulating both initiation and elongation of transcription (Bentley, 1995). This finding suggested that regulation of gene expression at the level of transcriptional elongation, which was previously implicated in control of expression of the proto-oncogenes c-myc, c-myb and c-fos (Bentley, 1995), is an important and perhaps general mechanism in control of gene expression. Therefore, the proposed study aimed at understanding the mechanism by which transcriptional activators enhance elongation will contribute to understanding the causes and mechanisms of cancer. Protein-protein interactions we have identified and protein modifications that increase the processivity of transcription elongation
might be potential targets for designing drugs to enhance or inhibit selected genes for cancer and AIDS therapies. An *in vitro* system with purified factors which duplicates the natural regulation of elongation by Tat and other activators can also be used for initial drug testing.
Chapter VI

References


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


Lane, D.P. and Crawford, L.V. (1979) T antigen is bound to a host protein in SV40-transformed cells. Nature 278: 261-263.


Liao, S.-M., Zhang, J., Jeffrey, D.A., Koleske, A.J., Thompson, C.M., Chao, D.M., Viljoen, M., van Vuuren, H.J.J., and Young,


Reinberg, D. and Roeder, R.G. (1987b) Factors involved in specific transcription by mammalian RNA polymerase


Workman, J.L. and Roeder, R.G. (1987) Binding of transcription factor TFIID to the major late promoter during in vitro
nucleosome assembly potentiates subsequent initiation by RNA polymerase II. Cell 51: 613-622.


IMAGE EVALUATION
TEST TARGET (QA-3)

1.0
1.1
1.25
1.4
1.6

1.0
1.1
1.25
1.4
1.6

150mm

6"