This is to certify that the

dissertation entitled

The Structure and Function of Human RAP30, the Small Subunit of General Transcription Factor TFIIF

presented by

Shi-Min Fang

has been accepted towards fulfillment
of the requirements for

Ph.D degree in Biochemistry

Major professor

Date 1/4/96

THE STRUCTURE AND FUNCTION OF HUMAN RAP30, THE SMALL SUBUNIT OF GENERAL TRANSCRIPTION FACTOR TFIIF

BY

Shi-Min Fang

A DISSERTATION

Sumitted to
Michigan State University
in partial fulfilment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Biochemistry

1995

ABSTRACT

THE STRUCTURE AND FUNCTION OF HUMAN RAP30, THE SMALL SUBUNIT OF GENERAL TRANSCRIPTION FACTOR TFIIF

BY

Shi-Min Fang

RAP30 is the small subunit of TFIIF (RAP30/74), a general initiation and elongation factor for transcription by RNA polymerase II. RAP30 has functions analogous to the functions of bacterial sigma factors. It binds to RNA polymerase II, recruits RNA polymerase II to the pre-initiation complex, prevents RNA polymerase II from binding nonspecifically to DNA, and greatly inhibits nonspecific transcription by RNA polymerase II. Applying a 2-dimensional sequence analysis method called Hydrophobic Cluster Analysis, we have aligned RAP30 to bacterial sigma factors. RAP30 contains sequences weakly similar and co-linear with conserved subregions 1.2, 2.1, 3.1, and 4.1 of sigma factors. RAP30 also has sequence similar to the N-terminal region of delta protein, a non-essential subunit of *B. sublilis* RNA polymerase.

In order to understand the relationship between the structure and the function of RAP30, a set of deletion mutants of RAP30 was constructed. Mutants were tested for accurate transcriptional activity, RAP74 binding, RNA polymerase II binding and TFIIB binding. Transcription assays indicate the importance of both N- and C-terminal regions for RAP30 function. RAP74 binds to the N-terminal region of RAP30 between amino acids 1-98. Two regions of RAP30, one near the N-terminus and one within the central region, are important for RAP30 binding. Multiple contacts within the RAP74-binding and RAP30 II-binding regions of RAP30 contribute to TFIIB

binding. Deletion of the N-terminal region of RAP30 abolishes RAP74-binding, RNA polymerase II-binding and TFIIB binding and activates the C-terminal region of RAP30 for DNA binding. The N-terminal region of RAP30, therefore, may be a domain that regulates the accessibility of central and C-terminal domains.

Several lines of evidence indicate that there may be significant coupling of TFIIF and TFIIB function in transcription. Both RAP30 and RAP74 bind independently to TFIIB. Analysis of deletion mutants of RAP74 shows that a C-terminal region between amino acids 358-517 binds directly to TFIIB, and this region of RAP74 also binds to RNA polymerase II. Interestingly, RAP74 antagonizes the interaction between TFIIB and RAP30, both by binding to RAP30 and by binding to TFIIB. RAP30, therefore, binds to TFIIB only in the absence of RAP74. When the TFIIF complex is intact, TFIIF-TFIIB contact most likely is maintained through the RAP74 subunit. If these binding relationships are maintained within functional transcription complexes, dynamic interactions between TFIIF subunits and TFIIB may be a mechanism to separate RAP30 and RAP74 functions during various stages of the transcription cycle.

To my parents and my homeland

On the River Bank Without A Boat

by Shi-Min Fang

After so many years,
So many rivers have dried up.
This one cannot last too long.

I will be able to cross it on foot
To look for the debris of paradise.
The paradise built with toy bricks
Collapsed on the other shore,
And decayed into those black black cobbles.

When I arrive at the middle of the river,
The river will rise quietly.
I will stay at the middle of the river,
The river will drown me slowly.

The river is rising quitely at this moment,
And all of cobbles are shining together.

ACKNOWLEDGMENTS

I would like to thank my mentor, Dr. Zachary Burton, for his guidance and friendship through the years. His contribution to this work has been invaluable. I also thank the members of my guidance committee, Drs. Jon Kaguni, Lee Kroos, John Linz and John Wang, for their advice and encouragement.

I acknowledge the supports of the former and present members of Dr. Burton lab, Dr. Ann Finkelstein, Dr. Wladyslaw Werel, David Chavez, Bo Qing Wang, Chun-hsiang Chang, Lei Lei, and Stephan Reymez. Special thanks to Ann for teaching me the first experiment in this lab, to Bo Qing and Chun-hsiang for providing RAP74 proteins and HeLa cell extracts that were used in this work. The friendship from my colleagues and friends has made East Lansing such a pleasant place to work and live.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PAGE

Chapter I

Chapter II

Chapter III

Chapter IV

LIST OF ABBREVIATIONS

a.a.	amino acid(s)
AdMLP	adenovirus major late promoter
ATP	adenosine triphosphate
bp	base pairs
BSA	bovine serum albumin
CTP	cytosine triphosphate
Da	Dalton
DNA	deoxyribonucleic acid
DTT	dithiothreitol
EDTA	ethylenediamine tetraacetic acid
EGTA	ethyleneglycol-bis-(β-aminoethyl ether) N,N,N',N'-tetraacetic acid
ELISA	enzyme-linked immunosorbent assay
GTP	guanine triphosphate
HCA	hydrophobic cluster analysis
HEPES	N-(2-hydroxyethyl)piperazine-N'-(2-ethanesulfonic acid)
IPTG	isopropyl-β-D-thiolgalactopyranoside
kb	kilobase pairs
kDa	kilodaltons
mRNA	messenger RNA
nt	nucleotides
NTP	nucleoside triphosphate
O.D. (490nm)	optical density at 490nm
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
RAP	RNA polymerase II associating proteins

RNA	ribonucleic acid
RNAP II	RNA polymerase II
rRNA	ribosomal RNA
SDS	sodium dodecyl sulfate
TAF	TBP associated factor
TBP	TATA box-binding protein
TCA	trichloroacetic acid
TFII	transcription factor of RNA polymerase II
tRNA	transfer RNA

Single letter abbreviations for the amino acids: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

CHAPTER I

INTRODUCTION

LITERATURE REVIEW

One of the most important ways to modulate gene expression in biological systems is to regulate transcription. The mechanism of gene transcription in eukaryotes is more complex and less well understood than in prokaryotes. In prokaryotes a single form of RNA polymerase, composed of three distinct subunits ($\alpha_2\beta\beta'$), performs all RNA synthesis. Selective recognition of promoters is determined by a σ initiation factor that associates with the $\alpha_2\beta\beta'$ "core". By contrast, there are three distinct nuclear RNA polymerases in eukaryotes. These three polymerases, each containing from 8 to 14 subunits, are responsible for transcribing different sets of genes: RNA polymerase I synthesizes ribosomal RNA precursors, RNA polymerase II transcribes protein-coding or class II genes, and RNA polymerase III synthesizes 5S ribosomal RNA and transfer RNA. Each of these enzymes requires a variety of auxiliary factors, known as "transcription factors", for selective promoter recognition and regulated transcription initiation, and cannot specifically initiate transcription on its own. In this review, only the synthesis of messenger RNA by RNA polymerase II is discussed.

Transcription by RNA Polymerase II

Transcription by RNA polymerase II (RNAP II) in eukaryotes involves multiple steps including: 1) template activation; 2) promoter recognition; 3) transcription initiation; 4) transcription elongation and 5) transcription termination. Each step requires a variety of transcription factors and DNA elements.

Transcription Factors and Promoters

Transcription is regulated by the interaction of specific proteins (trans-factors) with control signals in the DNA (cis-elements). The trans-factors that regulate transcription initiation of mRNA can be divided into three classes (Matsui et al. 1980; Davidson et al., 1983; Samuels et al., 1982; reviewed by Drapkin et al, 1993). The first class, termed general transcription factors (GTFs) or basal transcription factors, are required for basal level transcription initiation on almost all class II promoters (reviewed by Weinmann, 1992; Zawel and Reinberg, 1993; Conaway and Conaway, 1993; Buratowaki, 1994; Maldonado and Reinberg, 1995). The second class are sequence-specific transcription factors including activators and repressors, which regulate the rate of transcription initiation by binding to the specific DNA elements. They interact with basal transcription factors directly or indirectly (reviewed by Mitchell and Tjian, 1989; Ham et al., 1992; Tjian and Maniatis, 1994). The third class consists of cofactors, mediators or adaptors, which mediate the interactions between activators or repressors and basal factors to relay regulatory signals to RNAP II. Cofactors can also stimulate or repress transcription by themselves (Merino et al., 1993; Auble and Hahn, 1993; Ge and Roeder, 1994; Kretzschmar et al., 1994; reviewed by Zawel and Reinberg, 1995).

The promoter acts as a DNA signal that directs transcription factors and RNAP II to the initiation site (Weil et al., 1979; Manley et al., 1980). Cis-elements in the promoters of class II genes can be divided into four classes, and the combination of all these cis-elements gives a promoter its characteristic strength. The first two classes, TATA box and the initiator (Inr), are considered as the core promoter elements, and one or both of them appears to be present in all protein-coding genes. The TATA box, TATAAA, is located 25-30 base pairs upstream of the transcription start site, and is critical to determine the site of transcription initiation (Wasylyk et al., 1980; Grosschedl and Birnsteil, 1980; reviewed by Breathnach and Chambon, 1981). The initiator element consists of 17 nucleotides encompassing the transcription start site, and is present in most promoters, although its nucleotide sequence is not highly conserved (reviewed by Weis

and Reinberg, 1992). While TATA box and Inr can function cooperatively when present together (Smale and Baltimore, 1989; Nakatani et al., 1990; Conaway et al., 1990), either of them can potentially direct transcription initiation by itself (Myers et al., 1986; Smale et al., 1990; Nakatani et al., 1990). The basal transcription machinery, consisting of RNAP II and GTFs, functions through the TATA box and Initiator. The third class of cis-promoter element provides binding sites for transcription factors that recognize specific sequences such as CCAAT and the GC element. They are located within a few hundred base pairs of the transcription start site. (McKnight and Kingbury, 1982; Myers et al., 1986; Maniatis et al., 1987). The fourth class is enhancers, which may be located thousands of base pairs away from the start site (reviewed by Khoury and Gruss, 1983). The activators binding to enhancers can bind and stimulate the basal transcription machinery through DNA looping and protein-protein contacts to regulate transcription (reviewed by Ptashne, 1988; Ptashne and Gann 1990). Because their presence and arrangement are gene-specific, both consensus sequence elements and enhancers are considered as variable cis-elements.

Template Activation

Although biochemical studies of transcription by RNAP II have been carried out predominantly with naked DNA templates, DNA *in vivo* is packaged into a nucleoprotein complex known as chromatin. Packaging of naked DNA templates into chromatin causes transcription to be repressed. The binding of transcription activators and/or basal transcription factors, such as TFIID, to the DNA template prior to the assembly of chromatin can relieve this repression, suggesting that transcription activators function, at least in part, to antagonize chromatin-mediated inhibition of transcription (Croston et al., 1991; reviewed by Paranjape et al. 1994). A nuclear multiprotein complex, known as the SWI/SNF complex, was observed to induce changes in chromatin structure, and facilitate

dramatically the binding of transcription activators or GTFs to the nucleosomal DNA templates to allow transcription activation (Kwon et al., 1994; Imbalzano et al., 1994).

Promoter Recognition and Transcription Initiation

Accurate transcription by RNAP II is initiated from promoter sequences. RNAP II and GTFs (TFIIA, TFIID, TFIIB, TFIIE, TFIIF, TFIIH, and TFIIJ) assemble on the core promoter to form a pre-initiation complex, which is a highly ordered, stepwise process (reviewed by Weinmann, 1992, Zawel and Reinberg, 1993; Conaway and Conaway, 1993; Buratowaki, 1994; Maldonado and Reinberg, 1995). Once assembled, DNA strands are separated to expose the template for phosphodiester bond synthesis, and the pre-initiation complex is converted into a transcriptionally active open complex. This step requires hydrolysis of ATP at the β-γ bond position (Bunick et al., 1982; Sawadogo and Roeder, 1984). This energy requirement is unique to RNAP II transcription since RNA polymerase I, III, and bacterial RNA polymerase do not require ATP hydrolysis for transcription initiation. Although the mechanism involved in this requirement is still not clear, evidence suggests that ATP might serve as a substrate for a helicase activity that carries out the DNA unwinding reaction at the transcription start site (W. Wang, et al., 1992). Alternatively, it was shown that ATP hydrolysis was required not during the formation of the first few phosphodiester bonds in the RNA, but subsequently for promoter escape (Goodrich and Tjian, 1994). Promoter escape, also called promoter clearance, is the transition stage from abortive initiation to productive elongation. Newly synthesized short RNA transcripts, 5 to 10 nucleotides in length, are not stably associated with RNAP II and are released before productive elongation occurs (Luse and Jacob, 1987; Linn and Luse, 1991; Jacob et al., 1991; 1994). After transcription initiation, the structure of the initiation complex is altered, so RNAP II can escape the promoter and enter the elongation phase. It has been shown that TFIIE, TFIIH and ATP hydrolysis are

required for promoter escape in a reconstituted transcription system (Goodrich and Tjian, 1994). In cell extract systems the RAP74 subunit of TFIIF (Chang et al., 1993) and P-TEFb (Marshall and Price, 1995) are additionally required for promoter escape. In the reconstituted system, after promoter escape, TBP remained bound to the promoter, whereas TFIIB, TFIIE, TFIIF, and TFIIH were released from the complex (Zawel et al., 1995). *In vivo*, in the presence of a full complement of elongation factors, some of these factors may remain associated with elongating RNAP II.

Elongation

Once RNAP II escapes the promoter, it commences highly processive elongation until it encounters a pause, arrest or termination site. Elongation may be discontinuous since RNAP II may stall without releasing the transcript. The barriers to elongation include DNA sequences recognized by RNAP II, DNA-binding proteins and nucleosomes. Protein factors that affect the DNA template structure and topology or directly act on the ternary elongation complex (RNA-DNA-protein) can regulate elongation. Elongation factors can be divided into at least two classes. The first class includes SII/TFIIS, which helps RNAP II to read-through the template when it is arrested in ternary complexes (reviewed by Kane, 1994; by Reines, 1994). TFIIS can stimulate the cleavage of nascent RNA from the 3' end by RNAP II. This cleavage reaction causes RNAP II to move backwards and releases a short RNA fragment of 7-17 nucleotides (Rudd et al., 1994). Once the transcription is restarted from the newly exposed 3' end of the transcript, RNAP II can move through the pause. The second class of elongation factors includes TFIIF, which increases the overall rate of elongation and promotes more rapid read-through of some blocks to elongation (Flores et al., 1989). It was proposed that the stimulation of elongation by TFIIF is mediated by the phosphorylation of its RAP74 subunit (Kitajima et al., 1994). Another elongation factor, IIS, also stimulates the rate of

elongation and promotes the read-through (Bradsher et al., 1993; Bradsher et al., 1993b). Transcription activators can also stimulate elongation by RNAP II *in vivo* (Yankulov et al., 1994).

Termination

Transcription termination results when the RNA polymerase not only stops transcription but also releases its transcript (reviewed by Kerppola and Kane, 1991). The termination regions of DNA templates are downstream from the 3'-end processing signals. RNAP II releases both its transcript and the DNA template when it crosses the termination regions (Proudfoot, 1989). The position of termination regions is gene specific and can span several hundred nucleotides (Wahle and Keller, 1992). The sequences sufficient for polyadenylation are not sufficient for termination. Also required is a region several hundred nucleotides downstream, where the termination occurs (Tantravahi et al., 1993).

The mechanism controlling RNAP II to terminate transcription is poorly understood. RNAP II recognizes specific sites in the DNA template which serve as termination signals. This recognition appears to be intrinsic to the polymerase as it occurs in the absence of accessory factors. The intrinsic termination site usually consists of T-rich region in the nontranscribed strand (Reines et al., 1987; Kerppola and Kane, 1988). Sequences flanking the termination site have been suggested to influence intrinsic termination efficiency. RNA secondary structures do not appear to control selection of these sites, but DNA structure affects the efficiency of intrinsic termination, probably by bending the template (Kerppola and Kane, 1990). RNA-binding proteins, which are involved in accurate 3' end formation of RNA transcripts, may also influence the termination reaction (Proudfoot, 1989).

Transcription Initiation by RNA Polymerase II

Transcription initiation by RNAP II is a complex process. Two working models of pre-initiation complex formation on TATA-containing promoters have been proposed: the ordered multistep assembly model and the holoenzyme assembly model (Fig. 1).

Multistep Assembly Process

In this model, the RNAP II and GTFs assemble on the TATA-containing promoter in a highly ordered fashion through protein-DNA and protein-protein interactions (reviewed by Zawel and Reinberg, 1993). In addition to RNAP II, at least seven GTFs (TFIID, TFIIA, TFIIB, TFIIF, TFIIE, TFIIH, and TFIIJ) modulate basal transcription. This process begins with the binding of TFIID to the TATA element. TFIID is the only GTF that binds DNA specifically, and this function is intrinsic to one of its subunits, the TATA binding protein (TBP) (Lewin, 1990). Other subunits of TFIID were termed TAFs (TBP associated factors). TFIIA appears to have no essential role in basal transcription (Cortes et al. 1992), but can stabilize the binding of TFIID to the TATA box (Buratowski et al., 1989; Maldonado et al, 1990; Lee et al., 1992). The addition of TFIIB to the DA complex (complex including promoter DNA, TFIIA and TFIID) results in a DAB complex (Buratowski et al., 1989; Maldonado et al, 1990). TFIIB serves as a bridge between TBP and RNAP II, and this interaction is critical for determination of the transcription start site (Pinto et al., 1992; Berroteran et al., 1994; Li et al., 1994). Although both TFIIB (Tschochner et al., 1992; Ha et al., 1993) and TBP (Usheva et al., 1992) can bind directly to RNAP II, stable binding of RNAP II to the DAB complex requires prior association of TFIIF with RNAP II (Flores et al., 1991; Buratowsk et al., 1991). The RAP30 subunit of TFIIF reduces the affinity of RNAP II for non-specific DNA (Killeen and Greenblatt, 1992), and by itself is capable of delivering

Figure 1. Preinitiation complex formation: the ordered multistep assembly model and the holoenzyme assembly model

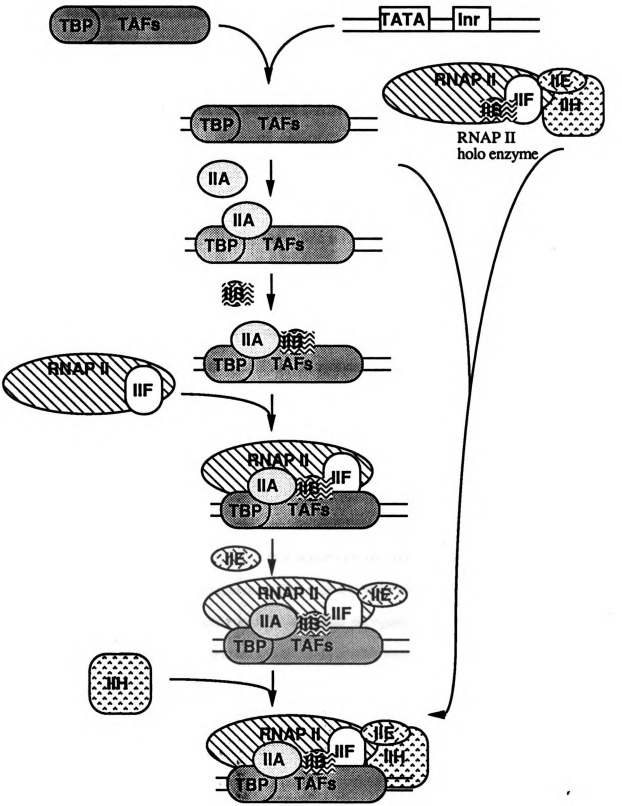

TBP
TAFs
TATA
Inr
RNAP II
IIE
IIF
IIH
RNAP II
holo enzyme
IIA
IIB

Figure 1

RNAP II to the DAB complex (Flores et al. 1991; Killeen et al., 1992). The RAP74 subunit of TFIIF likely stabilizes the complex (Flores et al., 1991; Tan et al., 1994; Coulombe et al., 1994). After formation of the DABPolF complex, TFIIE, TFIIH and TFIIJ sequentially join and form the complete DABPolFEHJ pre-initiation complex.

TFIIH contains several enzymatic activities, including an ATPase, a helicase and a kinase specific for the carboxy-terminal domain (CTD) of the largest subunit of RNAP II, and is essential to form a transcriptionally active complex (reviewed by Drapkin and Reinberg, 1994). TFIIE is required for the stable association of TFIIH with the complex (Flores et al., 1992), and stimulates the CTD kinase activity of TFIIH (Serizawa et al., 1994; Ohkuma and Roeder, 1994). TFIIJ was proposed as the last GTFs to bind and complete assembly of the pre-initiation complex (Zawel and Reinberg, 1993). The requirement, identity and function of TFIIJ, however, are not clearly established, and TFIIJ may be a component of native TFIID (Zawel and Reinberg, 1993).

Minimal Transcription Machinery

The multistep assembly model was derived from studies using a highly purified reconstituted system *in vitro*. It appears that under certain conditions, some GTFs can be dispensible for accurate transcription. Killeen et al. showed that TBP, TFIIB and the RAP30 subunit of TFIIF are sufficient for promoter recognition by RNAP II (Killeen et al., 1992). Tyree et al. further demonstrated that recombinant TBP, TFIIB, and RAP30, along with purified, native RNAP II, are able to accurately transcribe a subset of promoters using supercoiled templates (Tyree et al., 1993). Therefore TBP, TFIIB, RAP30 and RNAP II constitute a central, catalytic core of the transcription machinery. Negative supercoiling of the template removes the requirement for DNA strand separation by TFIIH and TFIIE (Goodrich and Tjian, 1994). The requirement for certain GTFs is also promoter dependent. When a supercoiled template was used, transcription of

the IgH promoter required only TBP, TFIIB, and RNAP II (Parvin and Sharp, 1993). Only TFIIB, RNAP II, and the YY1 transcription factor were required for initiation from supercoiled adeno-associated virus P5 promoter (Usheva and Shenk, 1994). Thus, it appears that only TFIIB, RNAP II, and a factor (TBP or YY1) that binds DNA at the TATA box or Initiator are absolutely required for accurate, basal level initiation *in vitro*.

Holoenzyme Assembly Model

Although it has been generally accepted that the assembly of the pre-initiation complex is a highly ordered multistep process, more recent evidence suggests that there is another pathway for assembly, in which RNAP II bound with a much more complex set of GTFs and accessory proteins binds to the promoters. This more complex form of RNAP II is refered to as the RNAP II "holoenzyme". TFIIB, TFIIF, TFIIE, and TFIIH assemble with RNAP II to form a transcriptionally active multifunctional complex capable of binding directly to TBP at the promoter in a single step (Serizawa et al., 1994b). A yeast RNAP II holoenzyme containing TFIIB, TFIIF, TFIIH, and many other unidentified polypeptides has been isolated from cell extracts by Young and co-workers (Koleske and Young, 1994; reviewed by Koleske and Young, 1995). This holoenzyme is capable of accurate transcription initiation when supplemented with TBP and TFIIE. Using another approach, Kornberg and co-workers have isolated another yeast RNAP II holoenzyme, containing RNAP II, SRB (suppressor of RNA polymerase B) proteins, Gal11, sug1, TFIIF and other unidentified polypeptides (Kim et al., 1994). This holoenzyme stimulates basal transcription and also affects the CTD kinase activity of TFIIH. This preparation can be further separated into two fractions: one consists of 12 core subunits of RNAP II and can support basal transcription, the other one, which stimulates both basal and activated transcription and was termed the "mediator", contains SRBs, Gal11, Sug1, TFIIF and other unidentified polypeptides (Kim et al., 1994).

Taken together, these results suggest that the preinitiation complex may be able to assemble by different pathways on different promoters. Several forms of RNAP II holoenzyme may exist in cells for transcription of different genes. Since the multistep assembly pathway was established based on the minimal requirement for assembly *in vitro*, the holoenzyme pathway is expected to more closely resemble the physiological function of RNAP II.

Transcription Activators

Transcription activators are a group of regulatory proteins that bind to the specific sequences located upstream or downstream of the core promoter region and increase the rate of transcription initiation. Transcription activators contain at least two domains: a DNA-binding domain that recognizes specific sequences, and a separate and independently acting domain required for activation (Ptashne, 1988). Based on the properties of their activation domains, activators have been loosely classified into several groups, such as acidic (including the yeast activator Gal4 and herpes simplex virion protein VP16), glutamine-rich (including Sp1) and proline-rich (including CTF) (reviewed by Mitchell and Tjian, 1989; by Tjian and Maniatis, 1994). Of these classes, acidic activators are unique in that they can function universally in all eukaryotes.

It is widely believed that activators function to facilitate some rate-limiting step in complex assembly or initiation by interacting with basal transcription factors and mediators. Some of the GTFs have been shown to be the target of upstream activators. VP16 can bind to the TBP component of TFIID (Stringer et al., 1990). Mutations that affect the strength of the activator also affect its ability to bind TFIID (Ingles et al., 1991). TFIIB also binds to the activation domain of VP16. A critical mutation in the activation domain of VP16 that dramatically reduces the activation potential eliminates binding to TFIIB (Lin and Green, 1991). Mutations in TFIIB that disrupt the interaction

with VP16 also cause defects in activated transcription but function normally in basal transcription (Roberts et al., 1993).

TFIIB has two functional domains: an N-terminal domain that interacts with TFIIF and RNAP II, and a C-terminal domain that interacts with TBP (Ha et al., 1993). It was proposed that these domains are engaged in an intramolecular interaction, and the binding of VP16 to the C-terminal domain of TFIIB disrupts this interaction and exposes binding sites for TFIIF and RNAP II. In this way interaction with an activator promotes pre-initiation complex formation (Roberts and Green, 1994) .

The 62kDa subunit of TFIIH is also the target of several activators, including VP16 and the tumor suppressor p53 (Xiao et al., 1994). Since TFIIH is involved in transcription, nucleotide excision repair, and perhaps cell cycle progression (Drapkin and Reinberg, 1994), the interactions between TFIIH and activators could be important not only in transcription, but also in other cellular process.

Coactivators

Although transcription activators can directly interact with GTFs, addition of transcription activators to highly purified reconstituted transcription systems does not substantially stimulate transcription. Activation, therefore, may require additional components to fully transduce the activation signal from the activator to the basal machinery. Indeed, several of these factors, termed coactivators, mediator or adaptors, have been identified. The coactivators can either enhance (PC1, PC2, PC3, PC4, ACF) or repress (NC1, NC2/Dr1, Dr2/PC3/Topoisomerase I (Topo I), MOT1/ADI, NOT1-4, TUP) transcription, and their functions appear to be general, in the sense that these factors do not appear to bind specifically to promoter DNA (reviewed by Zawel and Reinberg, 1995).

Of these coactivators, the best known ones are Dr2/PC3/Topo I and PC4. Topo I interacts with TBP and interferes with the ability of TFIID to bind to the TATA box, repressing basal transcription. However, Topo I can also enhance the response of the basal transcription complex to activators, and since repression is overcome by activators, the net fold-activation is greater in the presence of Topo I (Merino et al., 1993; Kretzschmar et al, 1993). Indeed, Topo I was first isolated by tracking an activity that enhanced the response to activators (Merion et al., 1993), and was also considered as a positive cofactor, PC3 (Kretzschmar et al., 1993). Among positive cofactors PC1-PC4, PC4 is the best characterized. PC4 binds to single stranded DNA and its transcriptional activity is negatively regulated by phosphorylation by casein kinase II (Kretzschmar et al., 1994; Ge et al, 1994). PC4 does not affect basal transcription, but rather enhances the response to multiple activator types. Consistent with this function, PC4 was found to interact with both the VP16 activation domain and DNA-TBP-TFIIA complex (Ge and Roeder, 1994; Ge et al., 1994).

RNA Polymerase II

RNAP II has been purified from several eukaryotic species, and generally has 10-12 subunits with sequences and structures that are similar between species. HeLa cell RNAP II contains 10 subunits ranging in size from 240 to 10 kDa. The yeast RNAP II, which appears to have 10 subunits with apparent molecular sizes from 220 to 10 kDa in SDS-PAGE, is actually composed of 12 subunits, RPB1-RPB12, and the genes encoding all these 12 subunits have been cloned and shown to be essential for wild type growth (reviewed by Young, 1991; by Woychik and Young, 1994). The three largest subunits, RPB1, RPB2, and RPB3, have counterparts in the bacterial RNA polymerase. Amino acid sequence comparison showed that RPB1 and RPB2 are clearly related to the two largest subunits of *E. coli* RNA polymerase, β' and β, respectively (Allison et al., 1985;

Sweetser et al., 1987). RPB1 and RPB2 are also functional homologs of their bacterial counterparts. RPB1, like β', is involved in DNA-binding whereas both RPB2 and β bind nucleotide substrates (reviewed by Young, 1991). The third largest subunit, RPB3, is related to the α subunit of *E. coli* RNA polymerase (Kolodziej and Young, 1991; Young, 1991). The two largest subunits of RNAP II are also related in size and sequence to the two largest subunits of RNA polymerases I and III, whereas RPB3 and RPB11 are homologous to the AC40 and AC19 subunits of RNA polymerases I and III (Mann et al., 1987; Woychik and Young, 1994). In addition, all three nuclear RNA polymerases share five identical subunits, the RPB5, RPB6, RPB8, RPB10, and RPB12 (Woychik et al, 1990; Carles et al., 1991; Treich et al., 1992).

The three eukaryotic nuclear and RNA polymerases have significant structural and functional relationships with bacterial RNA polymerase. The largest subunit of RNAP II, however, contains a unique carboxy terminal domain (CTD) that is not present in its prokaryotic counterpart or in RNA polymerases I or III (reviewed by Dahmus, 1994). The CTD consists of multiple repeats of the consensus sequence YSPTSPS. This heptapeptide sequence is repeated 26-27 times in yeast, 42-44 times in *Drosophila*, and 52 times in mouse and human. The CTD can be highly phosphorylated at serine and threonine sites in the heptapeptide repeats. The highly phosphorylated form of RNAP II, which is 240 kDa, is referred to as the IIO form, and the unphosphorylated or dephosphorylated form of 215 kDa as the IIA form (Cadena and Dahmus, 1987; Kim and Dahmus, 1989). The third form, IIB, which is 180 kDa and lacks most or all of CTD, is only observed *in vitro* and is considered to be a proteolytic product of purification (Kim and Dahmus, 1988; Kolodziej et al., 1990).

Genetic analysis indicates that the CTD has an essential role *in vivo*. Deletion mutations that remove most or all of the CTD are lethal in yeast (Nonet et al., 1987; Allison et al., 1988), *Drosophila* (Zehring et al., 1988) and mouse (Bartolomei et al., 1988). However, the dependency on the CTD appears to be promoter-specific. In yeast

truncations of the CTD from 27 to 11 repeats do not significantly reduce transcription from the *HIS4* gene whereas the inducible transcription of the *INO1* and *GAL10* genes is eliminated (Scafe et al, 1990). *In vitro* transcription studies also showed that CTD is unnecessary for Sp1- or Ad-2 MLTF-mediated activation (Buratowski and Sharp, 1990; Zehring and Greenleaf, 1990). The simplest interpretation of these results is that RNAP II is recruited to the promoter by a variety of interactions and only some of these involve the CTD. Several models for CTD function have been proposed, including involvement in transcription initiation, enzyme localization, DNA-binding, removal of chromatin proteins from DNA, and general regulation of enzyme activity (Young, 1991). A longer CTD seems to complement Gal4 deletion mutations (Allison and Ingles, 1989), and the CTD can directly interact with the TBP of TFIID (Usheva et al., 1992; Conaway et al., 1992). Thus it seems likely that CTD functions to facilitate interactions between upstream transcription activators and components of the basal transcription machinery, either directly or through additional factors.

The interaction of RNAP II with the preinitiation complex is dramatically influenced by the state of phosphorylation of the CTD. The nonphosphorylated IIA form stably associates with the pre-initiation complex, whereas the phosphorylated IIO form associates with the actively elongating complex. The conversion of RNAP IIA to RNAP IIO occurs prior to the formation of the first phosphodiester bond (Cadena and Dahmus, 1987; Payne et al., 1989; Laybourn and Dahmus, 1989; Laybourn and Dahmus, 1990; Lu et al., 1991). The postulated interaction between the acidic activation domains of a transcription activator and the CTD could be disrupted by phosphorylation of the CTD, providing a switch from initiation to elongation (Suzuki, 1990; Peterson et al., 1991). A model was presented that phosphorylation of CTD modulates the transition from initiation to elongation (Laybourn and Dahmus, 1990; Buratowski and Sharp, 1990). The IIA form of RNAP II would associate with assembly preinitiation complexs, then CTD kinases would trigger the elongation phase by phosphorylating the CTD.

The CTD phosphatase acts to reset the cycle. CTD kinase is an intrinsic activity of TFIIH (Feaver et al., 1991; Serizawa et al., 1992; Lu et al., 1992). A CTD phosphatase has also been identified (Chambers and Dahmus, 1994). This phosphatase interacts with a region of RNAP II distinct from CTD, and will not dephosphorylate the CTD unless this domain is covalently linked to RNAP II. The RAP74 subunit of TFIIF stimulates CTD phosphatase activity, and this stimulation can be suppressed by TFIIB (Chambers et al., 1995).

General Transcription Factors

TFIID (TBP and TAFs)

TFIID was first identified as an activity in nuclear extracts that can specifically bind a TATA-containing promoter (Parker and Topol, 1984). Further characterization showed that native TFIID is a multiprotein complex composed of the TATA-box-binding protein (TBP) and a number of tightly bound proteins called TBP-associated factors (TAFs) (reviewed by Hernandez, 1993; Goodrich and Tjian, 1994b; Moncollin et al., 1994). The data from biochemical and genetic studies suggest that TBP is involved in transcription by all three RNA polymerases (reviewed by Hernandez, 1993; Rigby, 1993). TBP is a subunit of the RNA polymerase I transcription factor SL1 (Comai et al., 1992) and a subunit of the RNA polymerase III transcription factor TFIIIB (Huet and Sentenac, 1992; Kassavetis et al., 1992; Lobo et al., 1992; Simmen et al., 1992; Taggart et al., 1992; White and Jackson, 1992). In yeast, mutations in TBP affect transcription by all three RNA polymerases (Cormack and Struhl, 1992; Schultz et al., 1992). Thus, TBP is a universal eukaryotic transcription factor involved in the formation of the initiation complexes on TATA-containing and TATA-less RNA polymerase I, II, and III promoters.

cDNAs encoding TBP have been cloned from a number of species, including yeast (Cavallini et al., 1989; Hahn et al., 1989; Horikoshi et al., 1989; Schmidt et al., 1989), *Drosophila* (Hoey et al., 1990; Muhich et al., 1990) and human (Hoffman et al., 1990; Kao et al., 1990; M.G. Peterson et al., 1990). TBP is a small protein of 27 kDa in yeast, and 38 kDa in *Drosophila* and human. Amino acid sequence comparison of TBP from different species shows that the N-terminal region is highly divergent and the C-terminal region is highly conserved. The C-terminal region invariably contains two copies of a long imperfect repeat of 61-62 amino acids and short basic repeats (Hernandez, 1993). Although yeast and human TBP are highly conserved at the C-terminal domain and can substitute for each other in basal transcription *in vitro*, human TBP is unable to functionally replace yeast TBP *in vivo* (Cormack et al., 1991; Gill and Tjian, 1991). The analysis indicates that the species specificity is due to many minor differences within the conserved C-terminal domain, instead of differences in the highly divergent N-terminal domain. Mutagenesis demonstrates that the conserved C-terminal domain is sufficient for binding to the TATA box and basal transcription activities (Horikoshi et al., 1990; Hoey et al., 1990; M.G. Peterson et al., 1990), whereas the function of the divergent N-terminal region is unclear, although there are some indications that it may mediate some interactions with coactivators (Pugh and Tjian, 1990; Zhou et al., 1991).

TBP interaction with its TATA box recognition sequence has been analyzed in serveral ways. Kinetic analysis of yeast TBP-TATA box complex formation suggests a two-step pathway. The first step includes the binding of TBP to the TATA box, followed by the formation of a stable TBP-TATA complex likely dependent on conformational changes in the TBP, promoter, or both (Hoopes et al., 1992). TBP, as a monomer (Horikoshi et al., 1990), contacts DNA through interactions with the minor groove of the TATA element (Starr and Hawley, 1991; Lee et al., 1991), and induces a 100° bend around the TATA element (Horikoshi et al., 1992; Y. Kim et al., 1993; J.L. Kim et al.,

1993). X-ray crystallography studies predict that TBP resemble a saddle in which the C-terminal direct repeats straddled the DNA (Nikolov et al., 1992; Y. Kim et al., 1993; J.L. Kim et al., 1993). The orientation of TBP binding to DNA exposes the entire upper surface of TBP, and provides a target for many protein-protein interactions. Indeed, a diverse set of proteins, including GTFs, transcription activators and coactivators, have been reported to bind TBP specifically.

TBP is tightly associated with TAFs, which can only be dissociated under denaturing conditions. At least 7 TAFs have been identified from human (Pugh and Tjian, 1991; Zhou et al., 1993), 8 from *Drosophila* (Dynlacht et al., 1991). 9 yeast TAFs, which are more weakly associated with TBP, have also been identified (Reese et al., 1994). Monoclonal antibodies against the CTD of RNAP II block transcription initiation by native TFIID, but not by TBP, suggesting that TAFs may function to establish interactions between RNAP II and TFIID (Conaway et al., 1992). One or more TAFs may interact directly with TFIIH to stabilize its interaction with the preinitiation complex (Conaway et al., 1992b). However, the major function of TAFs appear to serve as coactivators that provide a link between the activation domains of enhancer-binding activators and the basal transcription machinery to direct gene-specific transcriptional activation (Chiang et al., 1993; Chen et al., 1994; Jacq et al., 1994). Specific interactions between TAFs and activators have been demonstrated, e.g., dTAF40 interacts with VP16 (Goodrich et al., 1993), dTAF110 with Sp1 (Hoey et al., 1993), and dTAF150 with NTF-1 (Chen et al., 1994). Different activators require different subsets of TAFs, e.g., a complex containing only TBP, TAF150 and TAF250 is sufficient to mediate a response to NTF-1 (Chen et al., 1994). Recent studies suggest that TAFs may play an important role not only as coactivators, but also as core promoter recognition factors. TAF150 was found to have sequence-specific DNA-binding activity at select core promoters (Verrijzer et al., 1994). Further investigation found that TAF250 and TAF150 can confer promoter selectivity by interacting specifically with downstream elements that include the initiator,

increasing the stability of TBP binding to TATA box, and mediating a higher level of transcription (Verrijzer et al., 1995).

TFIIA

TFIIA was first identified as a heterotrimer composed of three subunits of 37 kDa (α), 19 kDa (β), and 13 kDa (γ) from mammalian cells (Cortes et al., 1992), and a heterodimer composed of two subunits of 32 kDa and 13 kDa from yeast, which are encoded by the *TOA1* and *TOA2* genes, respectively (Ranish et al., 1992). The isolation of cDNA encoding human and *Drosophila* TFIIA revealed that the α and β subunits are actually encoded by a single cDNA, and therefore are probably generated by protein processing (Ma et al., 1993; DeJong and Roeder, 1993; Yokomori et al., 1993). Amino acid sequence analysis showed that the *Drosophila* and human α and β subunits share sequence similarity with the amino- and carboxy terminal regions of the large subunit of yeast TFIIA, and the γ subunit is homologous to the small subunit of yeast TFIIA (Ma et al., 1993; DeJong and Roeder, 1993; Yokomori et al., 1993; Sun et al., 1994; Yokomori et al., 1994; Ozer et al., 1994; Bernstein et al., 1994). Human and Yeast TFIIAs are functionally exchangable in basal transcription (Ranish et al., 1992).

TFIIA functions at an early stage of pre-initiation complex formation (Reinberg et al., 1987; Flores et al., 1992), by stimulating the binding of TFIID to DNA through direct interaction with TBP (Buratowski et al., 1989; Maldonado et al., 1990; Lee et al., 1992; Yamamoto et al., 1992; Buratowski and Zhou, 1992). TFIIA stimulates basal transcription when native TFIID is used, but has no effect when TBP is used. It was proposed that TFIIA functions as an anti-repressor, through interaction with TBP, to remove the negative components present in the crude TFIID fraction (Cortes et al., 1992). Further investigation demonstrated that TFIIA indeed counteracts the activity of Dr2, a negative regulator of basal transcription associated with TFIID (Merino et al., 1993).

However, transcription reactions useing highly purified TFIID, which does not contain repressors, are still significantly stimulated by recombinant TFIIA, suggesting that TFIIA directly stimulates transcription, perhaps through a TAF (Sun et al., 1994; Yokomori et al., 1994; Ozer et al., 1994; Bernstein et al., 1994). Indeed, TFIIA interacts directly with *Drosophila* TAF110 (Yokomori et al., 1993). TFIIA can also stimulate activated transcription by several activators, including Sp1, NTF-1, and VP16 (Yokomori et al., 1994). TFIIA binds directly to Zta and supports transcription activation by this activator (Ozer et al., 1994). Positive coactivator PC4 interacts with VP16, as well as TFIIA within the DNA-TBP-TFIIA complex (Ge and Roeder, 1994; Ge et al., 1994).

TFIIB

TFIIB is a single polypeptide of 33 kDa in human (Ha et al., 1991), 34.5 kDa in Drosophila (Wampler and Kadonaga, 1992), and 35 kDa in yeast (Pinto et al., 1992). The structure of TFIIB can be divided into two domains: a compact, protease-resistant C-terminal core, and a protease-susceptible N-terminal region (Barberis et al., 1993; Malik et al, 1993). Mutagenesis showed that the entire C-terminal core, which contains two imperfect direct repeats and a region with the potential to form an amphipathic helix flanked by the repeats (Ha et al, 1991; Pinto et al., 1992), is required and sufficient to interact with the promoter-bound TBP, and mutations in this region abolish the ability of TFIIB to support basal transcription (Buratowski and Zhou, 1993; Ha et al, 1993; Hisatake et al, 1993; Yamashita et al., 1993). The N-terminal region, which contains a putative zinc finger motif (Ha et al, 1991; Pinto et al., 1992), is important to recruit RNAP II to the promoter (Barberis et al., 1993; Buratowski and Zhou; Ha et al., 1993; Hisatake et al., 1993), and to bind the RAP30 subunit of TFIIF (Ha et al., 1993). Sequences comparison of TFIIB from different species revealed a highly conserved

region immediately distal to the zinc finger (Na and Hampsey, 1993), which is required for accurate start-site selection *in vivo* (Pinto et al., 1994).

TFIIB plays a critical role in basal transcription initiation. It is a component of minimal transcription machinery (Tyree et al., 1993; Parvin and Sharp, 1993; Usheva and Shenk, 1994). In the absence of other transcription factors, TFIIB is capable of binding stably to the TBP-promoter complex to form the DB complex (Buratowski et al., 1989). TFIIB also directly interacts with RNAP II (Wampler and Kadonaga, 1992; Tschochner et al., 1992; Ha et al., 1993). Thus, TFIIB serves as a bridge between RNAP II and TBP. In initiation complex assembly, the association of TFIIB is a rate-limiting step (Lin and Green, 1991). The crystal structure of TATAbox/TBP/TFIIB was recently solved, and the results strongly suggest that the N-terminal region of TFIIB is presented as a scaffold for subsequent assembly of RNAP II/TFIIF into the pre-initiation complex (Nikolov et al., 1995). Nuclear magnetic resonance structure of TFIIB also supports this model (Bagby et al., 1995). Genetic studies suggest that the interaction between RNAP II and TFIIB is critical in determining the transcription start site. Mutations in the yeast gene encoding TFIIB (*SUA7*) shifted the transcription start site (Pinto et al., 1992). Mutations in the largest subunit of RNAP II similarly affected initiation, suggesting that there is a functional interplay between TFIIB and the largest subunit of RNAP II in determining the transcription start site (Berroteran et al., 1994). Biochemical data are consistent with this suggestion. In a transcription system reconstituted with *Saccharomyces cerevisiae* GTFs, only pair-wise exchange of RNAP II and TFIIB from *S. cerevisiae* with counterparts from *Schizosaccharomyces pombe* shifted transcription initiation to an initiation pattern characteristic of *S. pombe* (Li et al., 1994). Thus, RNAP II and TFIIB are the sole determinants of the transcription start site.

TFIIB is also involved in activated transcription. The acidic activators GAL4-VP16 and GAL4-AH, and the proline-rich activator GAL4-CTF can function to recruit TFIIB and/or stabilize its association with TFIID (Lin and Green, 1991; T.K. Kim and

Roeder, 1994). A direct interaction between TFIIB and VP16 is required for transcriptional activation, and mutations in TFIIB that disrupt this interaction reduced activated, but not basal, transcription (Roberts et al., 1993). The N-terminus of TFIIB was proposed to bind to the C-terminus via an intramolecular interaction, and the binding of VP16 to the C-terminal region disrupts this interaction, changing the conformation of TFIIB in favor of formation of the pre-initiation complex (Roberts and Green, 1994). Recently, a glutamine-rich activator GAL4-*ftz*Q was demonstrated to directly interact with the N-terminus of TFIIB to mediate transcriptional activation, suggesting that different activators can function by contacting distinct regions of TFIIB (Colgan et al., 1995).

TFIIE

TFIIE is a heterotetramer composed of two 56 kDa (α) and two 34 kDa (β) subunits (Ohkuma et al., 1990; Inostroza et al., 1991; M.G. Peterson et al., 1991). cDNAs encoding both subunits of TFIIE haves been isolated (M.G. Peterson et al., 1991; Ohkuma et al., 1991; Sumimoto et al., 1991). The N-terminal region of TFIIE-α shows similarity to the σ subunit of bacterial RNA polymerase and several interesting structural motifs: a leucine repeat, a zinc finger, and a helix-turn-helix (Ohkuma et al., 1991). Recent mutational analysis suggests that the N-terminal half of TFIIE-α is essential for transcription and phosphorylation. The regions containing the leucine repeat and helix-turn-helix are important for interaction with TFIIE-β, and the zinc finger domain may be important for interaction with RNAP II (Ohkuma et al., 1995). The acidic C-terminal half is a direct binding site for TFIIH and mutation of this site also reduces basal transcription dramatically (Ohkuma et al., 1995). TFIIE-β contains regions similar to bacterial σ factors and the RAP30 subunit of TFIIF, and a portion of the basic region-helix-loop-helix motif found in several enhancer-binding proteins (Sumimoto et al., 1991).

TFIIE binds stably to RNAP II in solution and was purified as one of the RAPs from an RNAP II affinity column (Flores et al., 1989; Buratowski et al., 1991). Further study showed that it is TFIIE-α that binds to RNAP IIA (Maxon et al., 1994). The association of TFIIE with the pre-initiation complex requires the DABPolF complex, and form the DABPolFE complex, which is necessary for subsequent recruitment of TFIIH (Flores et al., 1992). Many lines of evidence indicated that TFIIE functionally interacts with TFIIH. TFIIE stimulates CTD phosphorylation by TFIIH (Lu et al., 1992; Ohkuma and Roeder, 1994). TFIIE-α also mediates the helicase activity of TFIIH either negatively (Drapkin et al., 1994) or positively (Serizawa et al., 1994). TFIIE has also been implicated in promoter clearance in conjunction with TFIIH (Goodrich and Tjian, 1994). TFIIE from *Saccharomyces cerevisiae* could not replace the *Schizosaccharomyces pombe* TFIIE in an *S. pombe* system unless the *S. cerevisiae* TFIIE and TFIIH were swapped together (Li et al., 1994). Indeed, TFIIE has been demonstrated to physically interact with TFIIH as well as TFIIF (Maxon et al., 1994).

TFIIH

TFIIH is the most complex of the GTFs and the only GTF that has been shown to contain enzymatic activities. The human TFIIH consists of 8 subunits of 34, 38, 41, 44, 50, 62, 80, and 89 kDa (Schaeffer et al., 1994). Its yeast homolog, factor b, contains 5 subunits (Feaver et al., 1993). TFIIH has several enzymatic activities, including an ATPase, a helicase, and a kinase specific for the CTD of the largest subunit of RNAP II (reviewed by Drapkin and Reinberg, 1994).

Most of subunits of TFIIH have been cloned from human, rat and yeast cells (reviewed by Drapkin and Reinberg, 1994; by Maldonado and Reinberg, 1995). The largest subunit of TFIIH, p89 in human, is identical to the nucleotide excision repair protein XPB/ERCC3 , which contains ATP-dependent DNA unwinding helicase activity

(Schaeffer et al., 1993). Later, other subunits of TFIIH were identified as the products of genes *XPD/ERCC2* and *SSL1*, which are also involved in nucleotide excision repair (Feaver et al., 1993; Drapkin et al., 1994; Humbert et al., 1994). Like ERCC3, ERCC2 also contains ATP-dependent helicase activity. Purified TFIIH was found to be able to restore nucleotide excision repair activity in a cell-free extract defective in nucleotide excision repair (Drapkin et al., 1994; Wang et al., 1994; VanVuuren et al., 1994). Recently, a reconstituted system for nucleotide excision repair with purified human components was developed for the first time and it was completely dependent on TFIIH (Mu et al., 1995). These results clearly demonstrate that TFIIH plays a role in nucleotide excision repair. It was proposed that there are two forms of TFIIH, a holo-TFIIH containing CTD-kinase activity and involved in transcription, and a repariasome lacking CTD-kinase activity and active in nucleotide excision repair (Svejstrup et al., 1995).

Genetic and biochemical analyses of Rad3 (ERCC2) and Rad25 (ERCC3) in yeast have defined the role of the ATPase and helicase domains of these proteins. The ERCC2 and ERCC3 helicase activities are required for nucleotide excision repair. However, although both proteins are absolutely required for transcription, the ERCC2 helicase is dispensable for this process, whereas the ERCC3 helicase is essential (Sung et al., 1988; Feaver et al., 1993; Guzder et al., 1994). The helicase activity of TFIIH is involved in promoter escape (Goodrich and Tjian, 1994). It was recently reported that two polypeptides of TFIIH are the MO15/Cdk7 kinase and cyclin H subunits of the Cdk (cyclin-dependent kinase)-activating kinase Cak, which phosphorylates cdc2, Cdk2 and Cdk4, and which is necessary for cell-cycle progression (Serizawa et al., 1995; Shiekhattar et al., 1995). Thus, Cak is the CTD kinase intrinsic to TFIIH. In addition to phosphorylating the CTD, TFIIH also specifically phosphorylates TBP, TFIIE-α and the RAP74 subunit of TFIIF (Ohkuma and Roeder, 1994). The presence of Cak in TFIIH suggests that there could be a link between transcription and the cell cycle machinery. The functional role of the CTD kinase of TFIIH, however, is not yet completely clear.

Since TFIIH plays multiple roles in transcription, nucleotide excision repair, and perhaps cell cycle progression, it is likely that TFIIH is a target of regulators. Indeed, the p63 subunit of TFIIH interacts with the activation domains of VP16, and the tumor suppressor p53 (Xiao et al., 1994).

TFIIF (RAP30/74)

RAP30/74 was first identified and purified from HeLa cell extract by affinity chromatography on a column containing immobilized RNA polymerase II, among a group of proteins termed RAPs (RNA polymerase II-Associated Proteins), capable of binding to immobilized RNAP II (Burton et al., 1986). Later it was purified to homogeneity from rat liver as initiation factor βγ (Conaway and Conaway, 1989), from *Drosophila* as factor 5 (Price et al., 1989), from human cells as TFIIF (Flores et al., 1988; 1990) and FC (Kitajima et al., 1990), and from yeast as factor g (Henry et al., 1992). Human RAP30/74 is composed of two subunits with apparent relative molecular masses of 30,000 and 74,000 (RAP30 and RAP74 subunits, respectively) (Burton et al., 1988). Its activity elutes from a gel filtration column with an apparent molecular weight of 220 kDa, suggesting that it exists in solution as an $\alpha_2\beta_2$ structure (Flores et al., 1988). Different from other TFIIF homologs, purified yeast factor g comprises three polypeptides, with apparent masses of 105, 54, and 30 kDa (Henry et al., 1992).

The human cDNAs encoding RAP30 (Sopta et al., 1989) and RAP74 (Finkelstein et al., 1992; Aso et al., 1992) have been cloned. RAP30 and RAP74 produced in *E. coli* can replace natural human RAP30/74 to direct accurate transcription *in vitro* (Finkelstein et al., 1992). The sequence of RAP30 cDNA encodes a protein of 249 amino acids with a predicted relative molecular mass of 28,378 and is very basic (pI=10.4), with a region of dense positive charge at amino-acid residues 168-182. It was suggested that the central portion of RAP30 (residues 93-165) is weakly similar to two noncontiguous regions of E.

coli σ^{70}, the regions 2 and 1b. These two regions are the regions conserved among bacterial and bacteriophage σ factors. They are postulated to contain the site of interaction of σ^{70} with the core component of bacterial RNA polymerase (Sopta et al., 1989). Another research group also suggested that the carboxy terminus of the RAP30 has weak sequence similarity with region 4 of bacterial σ factors; in particular, it exhibits significant similarity to the carboxyl-terminal regions 4.1 and 4.2 of SpoIIIC (*Bacillus subtilis* σ^{K}) (Garrett et al., 1992). Region 4 is believed to direct recognition and binding of RNA polymerase at the -35 element of bacterial promoters. These similaries, however, are very weak.

The cDNA of RAP74 encodes a protein of 517 amino acids with relative molecular mass of 58, 254. RAP74 is a highly charged protein. The amino acid sequence suggests that RAP74 has a globular N-terminal domain (residues 1-179), a charged central domain (residues 180-356), and a globular C-terminal domain (residues 357-517) (Finkelstein et al., 1992). Encompassing the end of the central domain and the beginning of the C-terminal domain of RAP74 is a sequence, DSSEES, which is repeated three times. The function of this sequence is unknown. It might serve as a target for a serine kinase (Karlin, 1993).

Homologues of human RAP30 have been cloned from rat (Kobayashi et al., 1992; Garrett et al., 1992), *Xenopus* (Gong et al., 1992), *Drosophila* (Gong et al., 1995; Frank et al., 1995), and yeast (Tfg2) (Henry et al., 1994). Sequence comparison reveals that compared with human RAP30, rat RAP30 shows 97.6% identity, *Xenopus* RAP30 83% identity and 91% similarity, *Drosophila* RAP30 50% identity and 63% identity, yeast Tfg2 30.6% identity and 50.8% similarity. The homologues of human RAP74 have also been cloned from *Xenopus* (Gong et al., 1992b), from *Drosophila* (factor 5) (Kephart et al., 1993), and from yeast (SSU71/Tfg1) (Henry et al., 1994; Sun and Hampsey, 1995). Compared with human RAP74, Xenopus RAP74 shows 76% identity and 86% similarity, *Drosophila* RAP74 43% identity and 65% similarity, and yeast Ssu71/Tfg1 26.6%

identity and 49.9% similarity. The conserved amino acids are found particularly in regions near the amino and carboxyl termini of RAP30 or RAP74.

Yeast TFIIF is composed of three subunits, encoded by genes *TFG1*, *TFG2*, and *TFG3*, respectively. While Tfg1 and Tfg2 are the homologs of mammalian RAP74 and RAP30, Tfg3, which is less tightly associated and at most stimulatory to transcription and dispensable for cell viability, is one of TAFs, TAF30, and its gene is identical to a known gene, *ANC1*, which is implicated in cytoskeletal function (Welch and Drubin, 1994; Henry et al., 1994). Tfg3 has no known counterpart in mammalian TFIIF, but has sequence similarity to human proteins ENL and AF-9 (Henry et al., 1994).

Antibodies against RAP30 co-immunoprecipitate RAP74, indicating that RAP30 and RAP74 are normally associated in a complex (Burton et al., 1988). The N-terminus of RAP30 is essential and sufficient for RAP74 binding (Yonaha et al., 1992; Frank et al., 1995; Tan et al., 1995; Chapter III), and the N-terminus of RAP74 interacts with RAP30 (Yonaha et al., 1992; Wang and Burton, 1995). Both subunits of RAP30/74 are required for accurate transcription in a HeLa cell extract, from which RAP30/74 has been depleted by antibody against RAP30 (Burton et al., 1986, 1989). This result was confirmed with systems reconstituted from purified components (Flores et al., 1989). Both RAP30 and RAP74 are physical components of the pre-initiation complex (Flores et al., 1991). A variety of evidence suggests that RAP30/74 promotes transcription initiation by a mechanism similar to that of bacterial σ factors. First, RAP30/74 promotes binding of RNAP II to the DAB complex (Flores et al., 1991). Second, RAP30/74 binds stably to RNAP in solution (Reinberg and Roeder, 1987). Third, RAP30/74 prevents non-specific binding of RNAP II to free DNA (Conaway and Conaway, 1990; Killeen and Greenblatt, 1992).

It was proposed that RAP30/74 binds RNAP II through RAP30. Indeed, RAP30 appears not only structurally, but also functionally, homologous to bacterial σ factors. It was found that mammalian RNAP II specifically protected a serine residue in the σ^{70}-

related region of RAP30 from phosphorylation *in vitro*. In addition, human RAP30/74 bound to *E. coli* RNA polymerase and was displaced by σ^{70}. These results suggest that RAP30 and σ^{70} have functionally related RNA polymerase-binding regions and are able to bind to the same region of RNA polymerase (McCracken and Greenblatt, 1991). RAP30 alone can prevent RNAP II from binding nonspecifically to DNA and greatly inhibit nonspecific transcription by RNAP II (Killeen and Greenblatt, 1992). Furthermore, recombinant RAP30 is sufficient for recruitment of RNA polymerase II to the DAB complex (Flores et al., 1991; Killeen et al., 1992). This ability of RAP30 to recruit RNA polymerase to a promoter is also a characteristic of σ factors in prokaryotes. The findings that the RAP30 C-terminus is a cryptic DNA-binding domain (Tan et al., 1994) and that RAP30 can be crosslinked in the pre-initiation complex to promoter sequences between the TATA box and the transcription start site (Coulombe et al., 1994) suggest that TFIIF stabilizes the pre-initiation complex by nonspecifically binding to promoter DNA through its RAP30 subunit. This region of the promoter between the TATA box and Initiator is not known to contain specific promoter elements.

Recombinant RAP30 by itself cannot release RNAP II from non-promoter DNA once it has bound, although RAP30/74 has this capacity (Killeen et al., 1992; Killeen and Greenblatt, 1992). Other studies also indicate that RAP74 participates in the interaction of RAP30/74 with RNAP II by stabilizing either the interaction between RAP30 and RNAP II, between RNAP II and DNA, or between RAP30 and DNA (Garrett et al., 1992; Tyree et al., 1993; Coulombe et al., 1994). Indeed, RAP74 can directly interact with RNAP II through its C-terminus, which is masked by its N-terminus and central region (Wang and Burton, 1995). Although RAP30 can recruit RNAP II to the DAB complex without RAP74 (Flores et al., 1991; Killeen et al., 1992), positioning of RAP30 close to the promoter within the pre-initiation complex requires RAP74 (Coulombe et al., 1994). RAP74 can enhance the binding of bacterial RNA polymerase to a promoter and stimulates transcription *in vitro* by bacterial RNA polymerase (Chibazakura et al., 1994).

Evidence from both genetic and biochemical studies indicate that there may be significant coupling of TFIIB and TFIIF function in transcription. Both TFIIB and TFIIF are components of a minimal core transcription machinery (Tyree et al., 1993). The *SUA7* gene in yeast encodes TFIIB, and the RAP74 homologue in yeast is encoded by *SSU71/TFG1*. An *ssu71* mutant supresses a cold-sensitive *sua7* mutant, which alters the initiation start site (Sun and Hampsey, 1995). RAP74 stimulates CTD phosphatase activity, apparently through its interaction with RNAP II. TFIIB, however, suppresses stimulation of phosphatase activity by RAP74 (Chambers et al., 1995). In addition, RAP30 interacts physically with the N-terminal region of TFIIB (Ha et al., 1993; Chapter III). Thus, TFIIF interacts genetically, biochemically, and physically with TFIIB. Both RAP30 and RAP74 have also been shown to bind TFIIE (Maxon et al., 1994).

TFIIF also functions in activated transcription. A low amount of TFIIF was sufficient for basal transcription, whereas a higher quantity of TFIIF was required for activated transcription by serum response factor (SRF) and GAL4-VP16, but not Sp1. TFIIF could relieve squelching by SRF *in vitro*, and RAP74 bound to DNA interacted with either SRF or GAL4-VP16 (Zhu et al., 1994). Another study identified TFIIF as a stoichiometric component of a multiprotein mediator complex that responds to both GAL4-VP16 and GCN4 activators for transcription *in vitro* (Y.J. Kim et al., 1994).

In addition to being essential for initiation, RAP74 has also been implicated as being required for promoter escape at least under some circumstances (Chang et al., 1993). TFIIF also stimulates transcription elongation (Price et al., 1989). Mechanistic studies suggest that TFIIF increases the overall rate of RNA chain elongation by RNAP II by suppressing transient pausing of RNAP II at many sites on DNA templates (Flores et al., 1989; Izban and Luse, 1992; Bradsher et al., 1993; Kephart et al., 1994; Tan et al, 1994). TFIIF stimulates the basal level of elongation, but not the activated level of elongation by the HIV trans-acting protein Tat. Antiserum against RAP74 preferentially suppresses activation by Tat, suggesting that Tat may simulate elongation through

RAP74 (Kato et al., 1992). It was proposed that the stimulation of elongation by TFIIF is mediated by the phosphorylation of its RAP74 subunit (Kitajima et al., 1994). Other studies suggest that both RAP30 and RAP74 function in stimulation of the elongation by RNAP II (Tan et al., 1994; Tan et al., 1995).

OVERVIEW

Our laboratory has focused on functions of TFIIF, which has essential functions during the initiation and elongation stages of the transcription cycle by RNA polymerase II. A HeLa cell extract depleted of TFIIF has been used to assay TFIIF function *in vitro* (Burton et al., 1986, 1988). Cloning of cDNAs encoding the RAP30 and RAP74 subunits of TFIIF (Sopta et al., 1989; Finkelstein et al., 1992) and bacterial production of recombinant RAP30 and RAP74 proteins (Wang et al., 1993, 1994) has greatly facilitated further studies. A set of deletion mutants of RAP74 has also been constructed and assayed (Wang and Burton, 1995).

Since RAP30 has sigma-like function, several research groups have tried to identify the regions of RAP30 similar to the conserved regions of bacterial sigma factors (Sopta et al., 1989; McCracken and Greenblatt, 1991; Gong et al., 1992; Garrett et al., 1992). These previously reported sequence comparisons were based on linear alignment methods, and may not be reliable. In Chapter II, we applied a 2-dimensional sequence analysis method named Hydrophobic Cluster Analysis to compare amino acids sequences of bacterial sigma factors, *B. subtilis* delta protein, human and *Drosophila* RAP30, and yeast Tfg2p and Cdc73p. Our analysis indicates that RAP30 contains sequences similar and co-linear with conserved regions of sigma factors. The N-terminal region of RAP30 is similar to subregion 1.2 of bacterial sigma factors. The central region of RAP30 has been aligned to the core binding domain of sigma factors, subregion 2.1. The C-terminal region of RAP30 is similar to subregions 3.1 and 4.1 which are implicated in DNA binding. The C-terminal end of RAP30 may have similarity to the N-terminal region of delta protein. These results suggest that human RAP30 is structurally related to sigma and delta factors, RNA polymerase binding proteins from bacteria. Based on the sequence comparison, we suggest that yeast Cdc73p may be similar to RAP30, sigma and delta

factors. Cdc73p may be a subunit of an alternative version of TFIIF in yeast. The work described in this chapter is in preparation for publication.

In Chapter III, we report the experiments to map the functional domains of RAP30. A set of deletion mutants of human RAP30 was constructed with 6 histidine tags at the C-termini to facilitate the purification and binding reactions. The C-terminal histidine tag causes freshly renatured RAP30 to behave as a monomer in gel filtration, whereas the unmodified RAP30 is a dimer. Each of these mutants was tested for the ability to stimulate accurate runoff transcription from the Adenovirus major late promoter. Transcription activity of RAP30 was very sensitive to deletion of N-terminal or C-terminal sequences, indicating the importance of both N-and C-terminal sequences for RAP30 function. The RAP74 binding site on RAP30 was mapped by immobilizing histidine-tagged RAP30 mutants on nickel-chelate resin and testing the ability of this affinity matrix to retain RAP74. The results indicate that RAP74 binds to the N-terminal region of RAP30 between amino acids 1-98. The surfaces of RAP30 that interact with RNA polymerase II were determined using affinity beads containing covalently immobilized RNA polymerase II. The results demonstrate that RAP30 N-terminal sequences between 1-50 and central sequences between 131-159 and 152-176 are important for polymerase binding. These results are consistent with the sequence analysis described in Chapter II and indicate the significance of our RAP30 alignments with sigma subregions 1.2 and 2.1. The region of RAP30 that binds to TFIIB was mapped using two methods that may differ in the sensitivity of detection of protein-protein interactions. Using an affinity bead procedure in which TFIIB was covalently immobilized, an extensive surface of RAP30, localized to amino acids 1-176, was found to contribute to TFIIB binding. A minimal interaction domain was located between amino acids 27-118 by a more sensitive ELISA assay. Apparently, multiple contacts within the RAP74-binding and RNA polymerase II-binding regions of RAP30 contribute to TFIIB binding. It has been reported that deletion of the N-terminal region of RAP30 activates the DNA-

binding activity of the C-terminal region (Tan et al., 1994). We demonstrate here that deletion of the N-terminal region of RAP30 reduces RAP74, RNA polymerase II and TFIIB binding. The N-terminal region of RAP30, therefore, might regulate the function of central and C-terminal domains. Part of the work described in this chapter has been submitted to the Journal of Biological Chemistry for publication.

There are several lines of evidence indicating that there may be significant coupling of TFIIF and TFIIB function in transcription (Ha et al., 1993; Sun and Hampsey, 1995; Chambers et al., 1995). In Chapter IV, we investigated the relationship between TFIIF and TFIIB. Using an affinity bead procedure in which TFIIB was immobilized, TFIIB is shown to bind directly and independently to both RAP30 and RAP74. This result confirms the previous report of a direct interaction between TFIIB and RAP30 (Ha et al., 1993) and demonstrates for the first time that TFIIB also interacts directly with RAP74. A set of RAP74 mutants was tested for binding to TFIIB. The TFIIB binding site on RAP74 appears to be located within the C-terminal region between amino acids 358-517, and is masked by the N-terminal region and central region. Sequences within the N-terminal region and central region of RAP74 mask the C-terminal region for RNA polymerase II binding (Wang and Burton, 1995) and CTD phosphatase stimulation (Chambers et al., 1995). Using an affinity bead procedure and ELISA assays, we found that RAP74 antagonizes the interaction between TFIIB and RAP30 through two mechanisms. By binding the N-terminal region of RAP30, the N-terminal region of RAP74 can block interaction between RAP30 and TFIIB. Also, by binding TFIIB, the C-terminal region of RAP74 can block this interaction. RAP30, therefore, binds to TFIIB only in the absence of the RAP74 subunit. When the TFIIF complex is intact, TFIIF-TFIIB contact is most likely maintained through the C-terminal region of RAP74. We present three models to describe alternate pathways for initiation and productive elongation based on dynamic interaction between TFIIB and TFIIF subunits. The dynamic interaction between TFIIB and TFIIF subunits may be a

mechanism to separate RAP30 and RAP74 functions during various stages of the transcription cycle. The importance of dynamic interactions between RAP30 and RAP74 subunits of TFIIF is indicated by the work of Chang et al. (1993) in which it was shown that RAP30 and RAP74 have partially separable activities in accurate initiation and early elongation. Work described in Chapter IV has been submitted to the Journal of Biological Chemistry for publication.

REFERENCES

Allison, L.A., and Ingles, C.J. (1989). Proc. Natl. Acad. Sci. USA 86, 2794-2798.

Allison, L.A., Moyle, M., Shales, M., and Ingles, C.J. (1985). Cell 42, 599-610.

Allison, L.A., Wong, J.K.-C., Fitzpatrick, V.D., Moyle, M., and Ingles, C.J. (1988). Mol. Cell. Biol. 8, 321-329.

Aso, T., Vasavada, H.A., Kawaguchi, T., Germino, F.J., Ganguly, S., Kitajima, S., Weissman, S. M., and Yasukochi, Y. (1992). Nature 355, 461-464.

Auble, D.T., and Hahn, S. (1993). Genes Dev. 7, 844-856.

Bagby, S., Kim, S., Maldonado, E., Tong, K.I., Reinberg, D., and Ikura, M. (1995). Cell 82, 857-867.

Barberis, A., Muller, C.W., Harrison, S.C., and Ptashne, M. (1993). Proc. Natl. Acad. Sci. USA 90, 5628-5632.

Bartolomei, M.S., Halden, N.F., Cullen, C.R., and Corden, J.L. (1988). Mol. Cell. Biol. 8, 330-339.

Bengal, E., Flores, O., Krauskopf, A., Reinberg, D., and Aloni, Y. (1991). Mol. Cell. Biol. 11, 1195-1206.

Bernstein, R., DeJong, J., and Roeder, R.G. (1994). J. Biol. Chem. 269, 24361-24366.

Berroteran, R.W., Ware, D.E., Hampsey, M. (1994). Mol. Cell. Biol. 14, 226-237.

Bradsher, J.N., Tan, S., McLaury, H.-J., Conaway, J.W., and Conaway, R.C. (1993). J. Biol. Chem. 268, 25594-25603.

Bradsher, J.N., Jackson, K.W., Conaway, R.C., and Conaway, J.W. (1993b). J. Biol. Chem. 268, 25587-25593.

Breathnach, R., and Chambon, P. (1981). Ann. Rev. Biochem. 50, 349-383.

Bunick, D., Zandomeni, R., Ackerman, S., and Weinmann, R. (1982). Cell 29, 877-886.

Buratowski, S. (1994). Cell 77, 1-3.

Buratowski, S., and Sharp, P.A. (1990). Mol. Cel. Biol. 10, 5562-5564.

Buratowski, S., and Zhou, H. (1992). Science 255, 1130-1132.

Buratowski, S., and Zhou, H. (1993). Proc. Natl. Acad. Sci. USA 90, 5633-5637.

Buratowski, S., Hahn, S., Guarent, L., and Sharp, P.A. (1989). Cell 56, 549-561.

Buratowski, S., Sopta, M., Greenblatt, J., and Sharp, P.A. (1991). Proc. Natl. Acad. Sci. USA 88, 7509-7513.

Burton, Z.F., Ortolan, L.G., and Greenblatt, J. (1986). EMBO J. 5, 2923-2930.

Burton, Z.F., Killeen, M., Sopta, M., Ortolan, L.G., and Greenblatt, J. (1988). Mol. Cell. Biol. 8, 1602-1613.

Cadena, D.L., and Dahmus, M.E. (1987). J. Biol. Chem. 262, 12468-12474.

Carles, C., Triech, I., Bouet, F., Riva, M., and Sentenac, A. (1991). J. Biol. Chem. 266, 24092-24096.

Cavallini, B., Faus, I., Matthes, H., Chipoulet, J.M., Winsor, B., Egly, J.M., and Chambon, P. (1989). Proc. Natl. Acad. Sci. USA 86, 9803-9807.

Chambers, R.S., and Dahmus, M.E. (1994). J. Biol. Chem. 269, 26243-26248.

Chambers, R.S., Wang, B.Q., Burton, Z.F., and Dahmus, M.E. (1995). J. Biol. Chem. 270, 14962-14969.

Chang, C., Kostrub, C.F., and Burton, Z.F. (1993). J. Biol. Chem. 268, 20482-20489.

Chen, J.-L., Attardi, L., Verrijzer, C.P., Yokomori, K., and Tjian, R. (1994). Cell 79, 93-105.

Chiang, C.-M., Ge, H., Wang, Z., Hoffmann, A., and Roeder, R.G. (1993). EMBO J. 12, 2749-2762.

Chibazakura, T., Kitajima, S., Yonaha, M., and Yasukochi, Y. (1994). Biochimica et Biophysica Acta 1219, 592-600.

Colgan, J., Ashali, H., and Manley, J. (1995). Mol. Cell. Biol. 15, 2311-2320.

Comai, L., Tanese, N., and Tjian, R. (1992). Cell 68, 965-976.

Conaway, J.W., and Conaway, R.C. (1989). J. Biol. Chem. 264, 2357-2362.

Conaway, J.W., and Conaway, R.C. (1990). Science 248, 1550-1553.

Conaway, R.C., and Conaway, J.W. (1993). Annu. Rev. Biochem. 62, 161-190.

Conaway, J.W., Travis, E., and Conaway, R.C. (1990). J. Biol. Chem. 265, 7564-7596.

Conaway, R.C., Bradsher, J.N., and Conaway, J.W. (1992). J. Biol. Chem. 267, 8464-8467.

Conaway, R.C., Bradsher, J.N., and Conaway, J.W. (1992b). J. Biol. Chem. 267, 10142-10148.

Cormack, B.P., and Struhl, K. (1992). Cell 69, 685-696.

Cormack, B.P., Strubin, M., Ponticelli, A.S., and Struhl, K. (1991). Cell 65, 341-348.

Cortes, P., Flores, O., and Reinberg, D. (1992). Mol. Cell. Biol. 12, 413-421.

Coulombe B., Li, J., and Greenblatt, J. (1994). J. Biol. Chem. 269, 19962-19967.

Croston, G.E., Kerrigan, L.A., Lira, L.M., Marshak, D.R., and Kadonaga, J.T. (1991). Science 251, 643-649.

Dahmus, M.E. (1994). Transcription Mechanisms and Regulation, ed. by Conaway, R.C. and Conaway, J.W., New York: Raven Press, 243-262.

Davidson, B.L., Egly, J.-M., Mulvihill, E.R., and Chambon, P. (1983). Nature (London) 301, 680-686.

DeJong, and Roeder, R.G. (1993). Genes Dev. 7, 2220-2234.

Drapkin, R., and Reinberg, D. (1994). Trends Biochem. Sci. 19, 504-508.

Drapkin, R., Merino, A., and Reinberg, D. (1993). Curr. Opin. Cell Biol. 5, 469-476.

Drapkin, R., Reardon, J., Ansari, A., Huang, J.-C., Zawel, L., Ahn, K., Sancar, A., and Reinberg, D. (1994). Nature 368, 769-772.

Dynlacht, B.D., Hoey, T., and Tjian, R. (1991). Cell 66, 563-576.

Feaver, W.J., Gileadi, O., Li, Y., and Kornberg, R.D. (1991). Cell 67, 1223-1230.

Feaver, W.J., Svejstrup, J.Q., Bardwell, L., Bardwell, A.J., Buratowski, S., Gulyas, K.D., Donahue, T.F., Friedberg, E.C., and Konrberg, R.D. (1993). Cell 75, 1379-1387.

Finkelstein, A., Kostrub, C.F., Li, J., Chavez, D.P., Wang, B.Q., Fang, S.M., Greenblatt, J., and Burton, Z.F. (1992). Nature 355, 464-467.

Flanagan, P.M., Kelleher, R.J. III, Feaver, W.J., Lue, N.F., LaPointe, J.W., and Kornberg, R.D. (1990). J. Biol. Chem. 265, 11105-11107.

Flores., O., Maldonado, E., Burtonz, Z.F., Greenblatt, J., and Reinberg, D. (1988). J. Biol. Chem. 263, 10812-10816.

Flores, O., Maldonado, E., and Reinberg, D. (1989). J. Biol. Chem. 264, 8913-8921.

Flores, O., Ha, I., and Reinberg, D. (1990). J. Biol. Chem. 265, 5629-5634.

Flores, O., Lu, H., Killeen, M.T., Greenblatt, J., Burton, Z.F., and Reinberg, D. (1991). Proc. Natl. Acad. Sci. USA 88, 9999-10003.

Flores, O., Lu, H., and Reinberg, D. (1992). J. Biol. Chem. 267, 2786-2793.

Frank, D.J., Tyree, C.M., George, C.P., and Kadonaga, J.T. (1995). J. Biol. Chem. 270, 6292-6297.

Garrett, K.P., Serizawa, H., Hanley, J.P., Bradsher, J.N., Tsuboi, A., Arai, N., Yokota, T., Arai, K., Conaway, R.C., and Conaway, J.W. (1992). J. Biol. Chem. 267, 23942-23949.

Ge, H., and Roeder, R.G. (1994). Cell 78, 513-523.

Ge, H., Zhao, Y., Chait, B.T., and Roeder, R.G. (1994). Proc. Natl. Acad. Sci. USA 91, 12691-12695.

Gill, G., and Tjian, R. (1991). Cell 78, 513-523.

Gong, D.-W., Hashimoto, S., Wada, K., Roeder, R.G., Nakatani, Y., and Horikoshi, M. (1992). Nucleic Acids Res. 20, 6414.

Gong, D.-W., Hasegawa, S., Wada, K., Roeder, R.G., Nakatani, Y., and Horikoshi, M. (1992b). Nucleic Acids Res. 20, 6736.

Gong, D.-W., Mortin, M.A., Horikoshi, M., and Nakatani, Y. (1995). Nucleic Acids Res. 23, 1882-1886.

Goodrich, J.A., and Tjian, R. (1994). Cell 77, 145-156.

Goodrich, J.A., and Tjian, R. (1994b). Curr. Opin. Cell Biol. 6, 403-409.

Goodrich, J.A., Hoey, T., Thut, C., Admon, A., Tjian, R. (1993). Cell 75, 519-530.

Grosschedl, R., and Birnsteil, M.L. (1980). Proc. Natl. Acad. Sci. USA 77, 1432-1436.

Guzder, S.N., Sung, P., Bailly, V., Prakash, L., and Prakash, S. (1994). Nature 369, 578-381.

Ha, I., Lane, W.S., and Reinberg, D. (1991). Nature 352, 689-695.

Ha, I., Roberts, S., Maldonado, E., Sun, X., Kim, L., Green, M., and Reinberg, D. (1993). Genes Dev. 7, 1021-1032.

Hahn, S., Buratowski, S., Sharp, P.A., and Guarente, L. (1989). Cell 58, 1173-1181.

Ham, J., Steger, G., and Yaniv, M. (1992). FEBS 307, 81-86.

Henry, N.L., Sayre, M.H., and Kornberg, R.D. (1992). J. Biol. Chem. 267, 23388-23392.

Henry, N.L., Campbell, A.M., Feaver, W.J., Poon, D., Weil, P.A., and Kornberg, R.D. (1994). Genes Dev. 8, 2868-2878.

Hernandez, N. (1993). Genes Dev. 7, 1291-1308.

Hisatake, K., Roeder, R.G., and Horikoshi, M. (1993). Nature 363, 744-747.

Hoey, T., Dynlacht, B.D., Peterson, M.G., Pugh, B.F., and Tjian, R. (1990). Cell 61, 1179-1186.

Hoey, T., Weinzierl, R., Gill, G., Chen, J.-L., Dynlacht, B.D., and Tjian, R. (1993). Cell 72, 247-260.

Hoffmann, A., Sinn, E., Yamamoto, T., Wang, J., Roy, A., Horikoshi, M., and Roeder, R.G. (1990). Nature 346, 387-390.

Hoopes, B.C., LeBlanc, J.F., and Hawley, D.K. (1992). J. Biol. Chem. 267, 11539-11547.

Horikoshi, M., Wang, C.K., Fujii, H., Cromlish, J.A., Weil, P.A., and Roeder, R.G. (1989). Nature 341, 299-303.

Horikoshi, M., Yamamoto, T., Ohkuma, Y., Weil, P.A., and Roeder, R.G. (1990). Cell 61, 1171-1178.

Horikoshi, M., Bertuccioli, C., Takada, R., Wang, J., Yamamoto, T., and Roeder, R. (1992). Proc. Natl. Acad. Sci. USA 89, 1060-1064.

Huet, J., and Sentenac, A. (1992). Nucleic Acids Res. 20, 6451-6454.

Humbert, S., Van Vuuren, H., Lutz, Y., Hoeikmakers, J.H., Egly, J.M., and Mancollin, V. (1994). EMBO J. 13, 2393-2398.

Imbalzano, A.N., Kwon, H., Green, M.R., and Kingston, R.E. (1994). Nature (London) 370, 481-485.

Ingles, C.J., Shales, M., Cress, W.D., Triezenberg, S., and Greenblatt, J. (1991). Nature 351, 588-590.

Inostroza, J., Flores, O., and Reinberg, D. (1991). J. Biol. Chem. 266, 9304-9308.

Izban, M.G., and Luse, D.S. (1992). J. Biol. Chem. 267, 13647-13655.

Jacob, G.A., Luse, S.W., and Luse, D.S. (1991). J. Biol. Chem. 266, 22537-22544.

Jacob, G.A., Kitzmiller, J.A., and Luse, D.S. (1994). J. Biol. Chem. 269, 3655-3663.

Jacq, X., Brou, C., Lutz, Y., Davidson, I., Chambon, P., and Tora, L. (1994). Cell 79, 107-117.

Kane, C.M.(1994). Transcription Mechanisms and Regulation, ed. by Conaway, R.C. and Conaway, J.W., New York: Raven Press, 279-296.

Kao, C.C., Lieberman, P.M., Schmidt, M.C., Zhou, Q., Pei, R., and Berk, A.J. (1990). Science 248, 1646-1650.

Karlin, S. (1993). Proc. Natl. Acad. Sci. USA 90, 5593-5597.

Kassavetis, G.A., Joazeiro, C.A.P., Pisano, M., Geiduschek, E.P., Colbert, T., Hahn, S., and Blanco, J.A. (1992). Cell 71, 1055-1064.

Kato, H., Sumimoto, H., Pognonec, P., Chen, C., Rosen, C.A. and Roeder, R.G. (1992). Genes Dev. 6, 655-666.

Kephart, D.D., Price, M.P., Burton, Z.F., Finkelstein, A., Greenblatt, J., and Price, D.H. (1993). Nucleic Acids Res. 21, 1319.

Kephart, D.D., Wang, B.Q., Burton, Z.F., and Price, D.H. (1994). J. Biol. Chem. 269, 13536-13543.

Kerppola, T.K, and Kane, C.M. (1988). Mol. Cell. Biol. 8, 4389-4394.

Kerppola, T.K., and Kane, C.M. (1990). Biochemistry 29, 269-278.

Kerppola, T.K., and Kane C.M. (1991). FASEB J. 5, 2833-2842.

Khoury, G., and Gruss, P. (1983). Cell 33, 313-314.

Killeen, M., and Greenblatt, J. (1992). Mol. Cell. Biol. 12, 30-37.

Killeen, M., Coulombe, B., and Greenblatt, J. (1992). J. Biol. Chem. 267, 9463-9466.

Kim, J.L., Nikolov, D.B., and Berley, S.K. (1993). Nature 365, 520-552.

Kim, T.K., and Roeder, R.G. (1994). Proc. Natl. Acad. Sci. USA 91, 4170-4147.

Kim, W.Y., and Dahmus, M.E. (1988). J. Biol. Chem. 264, 11511-11520.

Kim, W.Y., and Dahmus, M.E. (1989). J. Biol. Chem. 264, 3169-3176.

Kim, Y., Geiger, J.H., Hahn, S., and Sigler, P.B. (1993). Nature 365, 512-519.

Kim, Y-J, Bjorklund, S., Li, Y., Sayre, M., and Kornberg, R. (1994). Cell 77, 599-608.

Kitajima, S., Tanaka, Y., Kawaguchi, T., Nagaoka, T., Weissman, S.M., and Yasukochi, Y. (1990). Nucleic Acids Res. 18, 4843-4849.

Kitajima, S., Chibazakura, T., Yonoha, M., Yasukochi, Y. (1994). J. Biol. Chem. 269, 29970-29977.

Kobayashi, Y., Kitajima, S., and Yasukochi, Y. (1992). Nucleic Acids Res. 20, 1994.

Koleske, A.J., and Young, R.A. (1994). Nature (London) 368, 466-469.

Koleske, A.J., and Young, R.A. (1995). Trends Biochem. Sci. 20, 113-116.

Kolodziej, P.A., and Young, R.A. (1991). Mol. Cell. Biol. 11, 4669-4678.

Kolodziej, P.A., Woychik, N.A., Liao, S.-M., and Young, R.A. (1990). Mol. Cell. Biol. 10, 1915-1920.

Kretzschmar, M., Meisterernst, M., and Roeder, R.G. (1993). Proc. Natl. Acad. Sci USA 90, 11508-11512.

Kretzschmar, M. Kaise, K., Lottspeich, F., and Meisterernst, M. (1994). Cell 78, 525-534.

Kwon, H., Imbalzano, A.N., Khavari, P.A., Kingston, R.E., and Green, M.R. (1994). Nature (London) 370, 477-481).

Laybourn, P.J., and Dahmus, M.E. (1989). J. Biol. Chem. 264, 6693-6698.

Laybourn, P.J., and Dahmus, M.E. (1990). J. Biol. Chem. 265, 13165-13173.

Lee, D.K., Horikoshi, M., and Roeder, R.D. (1991). Cell 67, 1241-1250.

Lee, D.K., Dejong, J., Hashimoto, S., Horikoshi, M., and Roeder, R.G. (1992). Mol. Cell. Biol. 12, 5189-5196.

Lewin, B. (1990). Cell 61, 1161-1164.

Li, Y., Flanagan, P.M., Tschochner, H., Kornberg, R.D. (1994). Science 263, 805-807.

Lin, Y.S., and Green, M.R. (1991) Cell 64, 971-981.

Linn, S.C., and Luse, D.S. (1991). Mol. Cell. Biol. 11, 1508-1522

Lobo, S.M., Tanaka, M., Sullivan, M.L., and Hernandez, N. (1992). Cell 71, 1029-1040.

Lu, H. Flores, O., Weinmann, R., and Reinberg, D. (1991). Proc. Natl. Acad. Sci. USA 88, 10004-10008.

Lu, H., Zawel, L., Fisher, L., Egly, J.-M., and Reinberg, D. (1992). Nature 358, 641-645.

Luse, D.S., and Jacob, G.A. (1987). J. Biol. Chem. 262, 14990-14997.

Ma, D., Watanabe, H., Mermelstein, F., Admon, A., Oguri, K., Sun, X., Wada, T., Imai, T., Shiroya, T., Reinberg, D., and Handa, H. (1993). Genes Dev. 7, 2246-2257.

Maldonado, E., Ha, I., Cortes, P., Weis, L., and Reinberg, D. (1990). Mol. Cell. Biol. 10, 6335-6347.

Maldonado, E. and Reinberg, D. (1995). Curr. Opio. Cell Biol. 7, 352-361.

Malik, S., Lee, D.K., and Roeder, R.G. (1993). Mol. Cell. Biol. 13, 6253-6259.

Maniatis, T., Goodbourn, S., and Fischer, J.A. (1987). Science 236, 1237-1244.

Manley, J.L., Fire, A., Cano, A., Sharp, P.A., and Gefter, M. (1980). Proc. Natl. Acad. USA 87, 3624-3628.

Mann, C., Buhler, J.M., Treich, I., and Sentenac, A. (1987). Cell 48, 627-637.

Marshall N.F., and Price, D.H. (1995). J. Biol. Chem. 270, 12335-12338.

Matsui, T., Segall, J., Weil, A.P., and Roeder, R.G. (1980). J. Biol. Chem. 255, 11992-11996.

Maxon, M.E., Goodrich, J.A., and Tjian, R. (1994). Genes Dev. 8, 515-524.

McCracken, S., and Greenblatt, J. (1991). Science 253, 900-902.

McKnight, S.L., and Kingbury, R. (1982). Science 217, 616-624.

Merino, A., Madden, K.R., Lane, W.S., Champoux, J.J., and Reinberg, D. (1993). Nature (London) 365, 227-232.

Mitchell, P.J., and Tjian, R. (1989). Science 245, 371-378.

Moncollin, V., Roy, R.D.W., and Egly, J.-M. (1994). Transcription Mechanisms and Regulation, ed. by Conaway, R.C. and Conaway, J.W., New York: Raven Press, 45-62.

Muhich, M.L., Lida, C.T., Horikoshi, M., Roeder, R.G., and Parker, C.S. (1990). Proc. Natl. Acad. Sci. USA 87, 9148-9152.

Myers, R.M., Tilly, K., and Maniatis, T. (1986). Science 232, 613-618.

Na, J.G., and Hampsey, M. (1993)., Nucleic Acids Res. 21, 3413-3417.

Nakatani, Y. Horikoshi, M., Brenner, M., Yamamoto, T., Besnard, F., Roeder, R.G., and Freese, E. (1990). Nature (London) 348, 86-88.

Nikolove, D.B., Hu, S.H., Lin, J., Gasch, A., Hoffmann, A., Horikoshi, M., Chua, N-H., Roeder, R.G., and Burley, S.K. (1992). Nature 360, 40-46.

Nikolove, D.B., Chen, H., Halay, E.D., Usheva, A.A., Hisatake, K., Lee, D.K., Roeder, R.G., and Burley, S.K. (1995). Nature 377, 119-128.

Nonet, M., Sweetser, D., and Young, R.A. (1987). Cell 50, 909-915.

Ohkuma, Y. and Roeder, R.G. (1994). Nature 368, 160-163.

Ohkuma, Y., Sumimoto, H., Horikoshi, M., and Roeder, R.G. (1990). Proc. Natl. Acad. Sci. USA 87, 9163-9167.

Ohkuma, Y., Sumimoto, H., Hoffmann, A., Shimasaki, S., Horikoshi, M., and Roeder, R.G. (1991). Nature 354, 398-401.

Ohkuma, Y., Hashimoto, S., Wang, C.K., Horikoshi, M., and Roeder, R.G. (1995). Mol. Cell. Biol. 15, 4856-4866.

Ozer, J., Moore, P.A., Bolden, A.H., Lee, A., Rosen, C.A., Lieberman, P. (1994). Genes Dev. 8, 2324-2335.

Parker, C.S., and Topol, J. (1984). Cell 36, 357

Payne, J.M., Laybourn, P.J., and Dahmus, M.E. (1989). J. Biol. Chem. 264, 19621-19629.

Paranjape, S.M. Kamakaka, R.T., and Kadonaga, J.T. (1994). Annu. Rev. Biochem. 63, 265-297.

Parvin, J.D., and Sharp, P.A. (1993). Cell 76, 1115-1121.

Peterson, C.L., Kruger, W., and Herskowitz, I. (1991). Cell 64, 1135-1143.

Peterson, M.G., Tanese, N., Pugh, B.F., and Tjian, R. (1990). Science 248, 1625-1630.

Peterson, M.G., Inostroza, J., Maxon, M.E., Flores, O., Admon, A., Reinberg, D., and Tjian, R. (1991). Nature 354, 369-373.

Pinto, I., Ware, D.E., Hampsey, M. (1992) Cell 68, 977-988.

Pinto, I., Wu, W.-H., Na, J.G., and Hampsey, M. (1994). J. Biol. Chem. 269, 30569-30573.

Price, D.H., Sluder, A.E., and Greenleaf, A.L. (1989). Mol. Cell. Biol. 9, 1465-1475.

Proudfoot, N.J. (1989). Trends Biochem. 14, 105-110.

Ptashne, M. (1988). Nature (London) 335, 683-689.

Ptashne, M., and Gann., A.A.F. (1990). Nature (London) 346, 329-331.

Pugh, B.F., and Tjian, R. (1990). Cell 61, 1187-1197.

Pugh, B.F., and Tjian, R. (1991). Genes Dev. 5, 1935-1945.

Ranish, J.A., Lane, W.S., and Hahn, S (1992). Science 255, 1127-1129.

Reese, J., Apone, L., Walker, S.S., Griffin, L.A., and Green, M.R. (1994). Nature 371, 523-527.

Reinberg, D., and Roeder, R.G. (1987). J. Biol. Chem. 262, 3310-3321.

Reinberg, D., Horikoshi, M., and Roeder, R.G. (1987). J. Biol. Chem. 262, 3322-3330.

Reines, D., Wells, D., Chamberlin, M.J., and Kane, G.M. (1987). J. Mol. Biol. 196, 299-312.

Reines, D. (1994). Transcription Mechanisms and Regulation, ed. by Conaway, R.C. and Conaway, J.W., New York: Raven Press, 263-278.

Rigby, P.W.J. (1993). Cell 72, 7-10.

Roberts, S., and Green, M.R. (1994). Nature 371, 717-720.

Roberts, S., Hal, I., Maldonado, E., Reinberg, D., and Green, M. (1993). Nature 363, 741-744.

Rudd, M.D., Izban, M.G., Luse, D.S. (1994). Proc. Natl. Acad. Sci. USA 91, 8057-8061.

Samuels, M.A., Fire, A., and Sharp, P.A. (1982). J. Biol. Chem. 257, 14419-14427.

Sawadogo, M., and Roeder, R.G. (1984). J. Biol. Chem. 257, 14419-14427.

Sayre, M.H., Tschochner, H., and Kornberg, R.D. (1992). J. Biol. Chem. 267, 23383-23387.

Scafe, C., Chao, D., Lopes, J., Hirsch, J.P., Henry, S., and Young, R.A. (1990). Nature 347, 491-494.

Schaeffer, L., Roy, R., Humbert, S., Moncollin, V., Vermeulen, W., Hoeijmakers, J.H., Chambon, P., Egly, J.-M. (1993). Science 260, 58-63.

Schaeffer, L., Moncollin, V., Roy, R., Staub, A., Mezzina, M., Sarasin, A., Weeda, G., Hoeijimaker, J.H.J., and Egly, J.-M. (1994). EMBO J. 13, 2388-2392.

Schmide, M.C., Kao, C.C., Pei, R., and Berk, A.J. (1989). Proc. Natl. Acad. Sci. USA 86, 7785-7789.

Schultz, M.Z., Reeder, R.H., and Hahn, S. (1992). Cell 69, 697-702.

Serizawa, H., Conaway, R.C., and Conaway, J.W. (1992). Proc. Natl. Acad. Sci. USA 89, 7476-7480.

Serizawa, H., Conaway, J.W., and Conaway, R.C. (1994). J. Biol. Chem. 269, 20750-20756.

Serizawa, H., Conaway, J.W., and Conaway, R.C. (1994b). Transcription: Mechanisms and Regulation, ed. by Conaway, R.C. and Conaway, J.W., New York: Raven Press, 27-43.

Serizawa, H., Makela, T.P., Conaway, J.W., Conaway, R.C., Weinberg, R.A., and Young, R.A. (1995). Nature 374, 280-282.

Shiekhattar, R., Mermelstein, F., Fisher, R.P., Drapkin, R., Dynlacht, B., Wessling, H.C., Morgan, D.O., and Reinberg, D. (1995). Nature 374, 283-287.

Simmen, K.A., Bernues, H., Lewis, J.D., and Mattaj, I.W. (1992). Nucleis Acids Res. 20, 5889-5898.

Smale, S.T., and Baltimore, D. (1989). Cell 57, 103-113.

Smale, S.T., Schmidt, M.C., Berk, A.J., and Baltimore, D. (1990). Proc. Natl. Acad. Sci. USA 87, 5268-5272.

Sopta, M., Burton, Z.F., and Greenblatt, J. (1989). Nature 341, 410-414.

Starr, D.B., and Hawley, D.K. (1991). Cell 67, 1231-1240.

Stringer, K.F., Ingles, C.J., and Greenblatt, J. (1990). Nature 345, 783-786.

Sumimoto, H., Ohkuma, Y., Sinn, E., Kato, H., Shimasaki, S., Horikoshi, M., and Roeder, R.G. (1991). Nature 354, 401-404.

Sun, Z.W., and Hampsey, M. (1995). Proc. Natl. Acad. Sci. USA 92, 3127-3131.

Sun, X., Ma, D., Sheldon, M., Yeung, K., and Reinberg, D. (1994). Genes Dev. 8, 2336-2348.

Sung, P., Higgins, D., Prakash, L., and Prakash, S. (1988). EMBO J. 7, 3263-3269.

Suzuki, M. (1990). Nature 344, 562-565.

Svejstrup, J.Q., Wang, Z., Feaver, W.J., Wu, X., Bushnell, D.A., Donahue, T.F., Friedberg, E.C., and Kornberg, R.D. (1995). Cell 80, 21-28.

Sweetser, D., Nonet, M., and Young, R.A. (1987). Proc. Natl. Acad. Sci. USA 84, 1192-1196.

Taggart, A.K.P., Fisher, T.S., and Pugh, B.F. (1992). Cell 71, 1015-1028.

Tan, S., Garrett, K.P., Conaway, R.C., and Conaway, J.W. (1994). Proc. Natl. Acad. Sci. USA 91, 9808-9812.

Tan, S., Conaway, R.C., and Conaway, J.W. (1995). Proc. Natl. Acad. Sci. USA 92, 6042-6046.

Tantravahi, J., Alvira, M., and Falck-Pedersen, E. (1993). Mol. Cell. Biol. 13, 578-587.

Treich, I., Charles, C., Riva, M., and Sentenac, A. (1992)., Gene Expr. 2, 31-37.

Tjian, R., and Maniatis, T. (1994). Cell 77, 5-8.

Tschochner, H., Sayre, M.H., Flanagan, P.M., Feaver, W.J., and Kornberg, R.D. (1992). Proc. Natl. Acad. Sci. USA 89, 11291-11296.

Tyree, C.M., George, C.P., Lira-DeVito, L.M., Wampler, S.L., Dahmus, M.E., Zawel, L., and Kadonaga, J.T. (1993). Genes Dev. 7, 1254-1265.

Usheva, A., and Shenk, T. (1994). Cell 76, 1115-1121.

Usheva, A., Maldonado, E., Goldring, A., Lu, H., Houbavi, C., Reinberg, D., and Aloni, Y. (1992). Cell 69, 871-881.

VanVuuren, A.J., Vermeulen, W., Ma, L., Weeda, G., Appeldoon, E., Jaspers, N.J.G., Var der Eb, A.J., Bootsma, D., Hoeijmakers, J.H.J., Humbert, S. et al. (1994). EMBO J. 13, 1645-1653.

Verrijzer, P., Yokomori, K., Chen, J.L., and Tjian, R. (1994). Science 264, 933-941.

Verrijzer, P., Chen, J-L., Yokomori, K., and Tjian, R. (1995). Cell 81, 1115-1125.

Wahle, E., and Keller, W. (1992). Annu. Rev. Biochem. 61, 419-440.

Wampler, S.L., and Kadonaga, J.T. (1992). Genes Dev. 6, 1542-1552.

Wang, B.Q., Kostrub, C.F., Finkelstein, A., and Burton, Z.F. (1993). Protein Expr. Purif. 4, 207-214.

Wang, B.Q., Lei, L., and Burton, Z.F. (1994). Protein Expr. Purif. 5, 476-485.

Wang, B.Q., and Burton, Z.F. (1995). J. Biol. Chem. 270, 27035-27044.

Wang, W., Carey, M., and Gralla, J.D. (1992). Science 255, 450-453.

Wang, Z., Svejstrup, J.Q., Feaver, W.J., Wu, X., Kornberg, R.D., and Friedberg, E.C. (1994). Nature 368, 74-75.

Wasylyk, B., Derbyshire, R., Guy, A., et al. (1980). Proc. Natl. Acad. Sci. USA 77, 7024-7028.

Weil, P.A., Luse, D.S., Segall, J., and Roeder, R.G. (1979). Cell 18, 469-484.

Weinmann, R. (1992). Gene Expression 2, 81-91.

Weis, L., and Reinberg, D. (1992). FASEB J. 6, 3300-3309.

Welch, M.D., and Drubin, D.G. (1994). Mol. Biol. Cell. 5, 617-632.

White, R.J., and Jackson, S.P. (1992). Cell 71, 1041-1053.

Woychik, N.A., Lane, W.S., and Young, R.A. (1990). Genes Dev. 4, 313-323.

Woychik, N.A., and Young, R.A. (1994). Transcription: Mechanisms and Regulation, ed. by Conaway, R.C. and Conaway, J.W., New York: Raven Press, 227-242.

Wu, D., Park, C.-H., Matsunaga, T., Hsu, D., Reardon, J.T., and Sancar, A. (1995). J. Biol. Chem. 270, 2415-2418.

Xiao, H., Pearson, A., Coulombe, B., Truant, R., Zhang, S., Regier, J.L., Triezenberg, S.J., Reinberg, D., Flores, O., Ingles, C.J., and Greenblatt, J. (1994). Mol. Cell. Biol. 14, 7013-7023.

Yamamoto, T., Horikoshi, M., Wang, J. Hasegawa, S., Weil, P.A., and Roeder, R.G. (1992). Proc. Natl. Acad. Sci. USA 89, 2844-2848.

Yamashita, S., Hisatake, K., Kobubo, T., Doi, K., Roeder, R.G., Horikoshi, M., and Nakatani, Y. (1993). Science 261, 463-466.

Yankulov, K., Blau, J., Purton, T., Roberts, S., and Bentley, D.L. (1994). Proc. Natl. Acad. Sci. USA 91, 5237-5241.

Yokomori, K., Admon, A., Goodrich, J.A., Chen, J.-L., and Tjian, R. (1993). Genes Dev. 7, 2235-2245.

Yokomori, K., Zeidler, M.P., Chen, J.-L., Verrijzer, C.P., Mlodzik, M., and Tjian, R. (1994). Genes Dev. 8, 2313-2323.

Yonaha, M., Aso, T., Kobayashi, Y., Vasavada, H., Yasukochi, Y., Weissman, S.M., and Kitajima, S. (1992). Nucleic Acids Res. 21, 273-279.

Young, R.A. (1991). Annu. Rev. Biochem. 60, 689-715.

Zawel, L., and Reinberg, D. (1993). Prog. Nucl. Acids Res. 44, 67-108.

Zawel, L., and Reinberg, D. (1995). Annu. Rev. Biochem. 64, 533-561.

Zawel, L., Kumar, K.P., and Reinberg, D. (1995). Gene. Dev. 9, 1479-1490.

Zehring, W.A., and Greenleaf, A.L. (1990). J. Biol. Chem. 265, 8351-8353.

Zehring, W.A., Lee, J.M., Weeks, J.R., Jokerst, R.S., and Greenleaf, A.L. (1988). Proc. Natl. Acad. Sci. USA 85, 3698-3702.

Zhou, Q., Schmidt, M., and Berk, A. (1991). EMBO J. 10, 1843-1852.

Zhou, Q., Boyer, T.G., and Berk, A.J. (1993). Gene Dev. 7, 180-187.

Zhu, H., Joliot, V., and Prywes, R. (1994). J. Biol. Chem. 269, 3489-3497.

CHAPTER II

SEQUENCE ANALYSIS OF RAP30

ABSTRACT

The RAP30 subunit of general transcription factor TFIIF has functions analogous to the functions of bacterial sigma factors. Using a 2-dimensional sequence analysis method, named Hydrophobic Cluster Analysis, RAP30 was found to be structurally related to bacterial sigma factors. RAP30 was found to contain sequences weakly similar and co-linear with conserved subregions 1.2, 2.1, 3.1, and 4.1 of sigma factors. RAP30 also has sequence similar to the N-terminal region of delta protein, a non-essential subunit of *B. sublilis* RNA polymerase. These proposed alignments are different from those reported previously based on linear alignment methods. We also found that the overall structure of yeast Cdc73p is similar to RAP30, suggesting that Cdc73p may be an alternative subunit of yeast TFIIF.

INTRODUCTION

Transcription in bacteria is catalyzed by a multisubunit RNA polymerase, $\alpha_2\beta\beta'$. This core polymerase is unable to recognize promoter sequences and does not accurately initiate transcription from promoter sites in the absence of a σ factor. Sigma factors are a family of relatively small, dissociable subunits of RNA polymerase, which bind to the core polymerase to form the holoenzyme. The holoenzyme then can recognize and bind to a promoter, permitting transcription initiation from a specific site. Each holoenzyme specifically recognizes a set of related promoter elements with a consensus sequence recognized by the sigma factor. After transcription initiation, the sigma factor is released and the core polymerase elongates an RNA chain (Reviewed by Helmann and Chamberlin, 1988; Helmann, 1994).

Each sigma factor has at least two functions: core binding and activation of promoter recognition. Some sigma factors also have DNA-melting activity and inhibit nonspecific transcription. The interaction of sigma factor with the core enzyme involves sites on both the β and β' subunits (Ishihama, 1992). Since there is a competition between different sigma factors for core binding (Zhou et al., 1992), it is likely that different sigma factors bind to the same region of the core polymerase. Therefore, different sigma factors may have similar core binding motifs. The binding of sigma factor to core reduces the affinity of the core for non-promoter DNA about 10^4 fold (Hinkle and Chamberlin, 1972). Genetic analyses indicate that sigma factors play a direct role in determining promoter recognition by contacting -35 and -10 consensus regions of promoter DNA (Zuber et al., 1989; Siegele et al, 1989; Gardella et al., 1989). Indeed, biochemical studies revealed two cryptic promoter DNA-binding domains of sigma 70 (Dombroski et al., 1992).

Sigma factors can be divided into two unrelated groups: the sigma 70 family, which consists of the primary and many of the alternative sigma factors, and the sigma 54 family. Sequence comparisons of members of the sigma 70 family shows that there are four highly conserved regions, and some of these regions can be further divided into subregions (Gribskov and Burgess, 1986; Helmann and Chamberlin, 1988; Lonetto et al., 1992). Regions 2 and 4 are the most highly conserved regions, indicating that these two regions are the most crucial for sigma factor functions. Indeed, genetic and biochemical studies suggest that most of sigma functions require portions of regions 2 and 4. Subregion 2.1, the most highly conserved region, is involved in core binding (Lesley and Burgess, 1989; Shuler et al., 1995). Subregion 2.3 is involved in DNA melting, which helps to establish the open promoter complex (Jones and Moran, 1992; Jones et al, 1992). Subregion 2.4 recognizes the -10 region of the promoter (Siegele et al., 1989; Zuber et al., 1989; Daniels et al., 1990; Waldurger et al., 1990; Dombroski et al., 1992), whereas subregion 4.2, which contains a helix-turn-helix DNA-binding motif, interacts with the -35 region of the promoter (Siegele et al., 1989; Gardella et al., 1989; Dombroski et al., 1992). Regions 1 and 3 are missing in many members of the sigma 70 family. Region 1 masks both the promoter and nonpromoter DNA-binding domains of sigma 70 (Dombroski et al., 1992). The function of region 3 is unknown. It may be involved in DNA-binding (Dombroski et al., 1992).

Bacillus subtilis RNA polymerase, in contrast to *E. coli* RNA polymerase, contains an additional non-essential subunit, the 21 kDa δ protein, which is encoded by the *rpoE* gene (Pero et al., 1975; Lampe et al., 1988). Together with sigma, the delta protein functions as an initiation subunit of polymerase to allow site-selective interaction with DNA templates (Achberger and Whiteley, 1981). Both delta and sigma bind to core simultaneously but with negative cooperativity, and promoter recognition may involve a ternary complex of Eδσ (Hyde et al., 1986). Delta inhibits transcription from non-specific template (Tjian et al., 1977), reduces the affinity of RNA polymerase for nonpromoter

DNA (Dickel et al., 1980; Achberger and Whiteley, 1981), and decreases the association of RNA polymerase with both promoter and non-promoter DNA (Achberger et al., 1982; Hilton and Whiteley, 1985). Delta participates in both the initiation and core recycling phases of the transcription cycle by inhibiting open complex formation at the promoter and enhancing core enzyme recycling (Juang and Helmann, 1994).

The RAP30 subunit of general transcription factor TFIIF has sigma-like and delta-like functions. It can not only bind to eucaryotic RNAP II, but it can also compete with sigma 70 to bind *E. coli* RNA polymerase (McCracken and Greenblatt, 1991). RAP30 can prevent RNAP II from binding nonspecifically to DNA and greatly inhibits nonspecific transcription by RNAP II (Killeen and Greenblatt, 1992). Furthermore, recombinant RAP30 is sufficient for recruitment of RNA polymerase II to the pre-initiation complex (Flores et al., 1991; Killeen et al., 1992). Like sigma 70, RAP30 also contains a cryptic DNA-binding domain at its C-terminus (Tan et al., 1994). Therefore, RAP30 appears to be a member of the sigma family.

It was proposed that RAP30 binds to RNAP II through its central region which has weak structural similarity to subregions 2.1 and 2.2 of sigma (McCracken and Greenblatt, 1991). The C-terminus of RAP30 was also suggested to contain weak similarity to subregion 4.2 of sigmas (Garrett et al., 1992). These sequence alignments, using linear alignment methods, do not show strong similarity and may be unreliable. Here we use a 2-dimensional sequence analysis method, named Hydrophobic Cluster Analysis (HCA), to compare amino acids sequences of sigma factors, delta protein, human and *Drosophila* RAP30, and yeast Cdc73p. Cdc73p was previously described as an extragenic suppressor of an α-factor receptor (*STE2*) mutant (Reed et al., 1988) and was recently identified as a novel yeast RNA polymerase II-associating protein (Wade et al., 1995). HCA is a powerful method for comparing amino acids sequences of distinct proteins, and can detect similarities in the 3D folding of proteins with very low sequence identity (Gaboriaud et al., 1987; Lemesle-Varloot et al., 1990). HCA can also be used to

search homologous domains that are separated by variable segments (Henrissat et al., 1988). Our analysis from HCA indicates that RAP30, as well as Cdc73p, may have sequence similarities to conserved subregions 1.2, 2.1, 3.1, and 4.1 of sigma factors and the N-terminal region of delta protein. These proposed alignments differ from those reported previously based on linear alignment methods. Based on the sequence comparison, we also suggest that Cdc73p may be an alternative subunit of yeast TFIIF.

METHODS

Hydrophobic Cluster Analysis

Amino acid sequences were obtained from GCG database. HCA plots were drawn using a Macintosh plot program from Doriane S.A. (France). The amino acid sequence is displayed as a classical α-helix projected in two dimensions and duplicated. Amino acids are represented by the standard one-letter code, except the following amino acids are denoted by symbols: Proline by "star", glycine by "diamond", cysteine by "c" with a circle, threonine by "square", and serine by "square" filled with a spot. V, I, L, F, W, M, and Y are considered hydrophobic. A and C can be considered hydrophobic in a hydrophobic environment. Hydrophobic clusters composed of adjacent hydrophobic residues that are not broken by prolines are circled.

Quality of Alignments

The quality of HCA alignments was calculated as the percentage of matched amino acids after linear alignment following the indication from HCA. HCA alignment scores were calculated as follows:

$$\text{HCA alignment score} = \frac{2CR \times 100}{RC_1 + RC_2}\%$$

Where RC_1 (RC_2) is the number of hydrophobic residues in cluster 1 (cluster 2). CR is the number of hydrophobic residues in cluster 1 that are in correspondence with hydrophobic residues in cluster 2 . If the HCA score is more than 60% along wide segments of their sequences (more than 100 residues), 3D homologous proteins are

detected even if their sequence identity is as low as 10%. When the hydrophilic segments localized between hydrophobic clusters are not conserved or differ in length, they can be considered to correspond to loops and not to dramatically influence protein folding (Gaboriaud et al., 1987).

RESULTS

1. Human RAP30 is structurally related to bacterial sigma factors

Hydrophobic Cluster Analysis revealed that human RAP30 contains several regions weakly similar to bacterial sigma factors and delta protein. The N-terminus of RAP30, which is the RAP74 binding domain, contributes to RNAP II binding (Chapter III) and masks the DNA-binding C-terminus (Tan et al., 1994), is similar to subregion 1.2 of bacterial sigma factors, which is important for masking DNA-binding domains (Dombroski et al., 1992) (Fig. 1). The central region of RAP30, which is involved in RNAP II binding (McCracken and Greenblatt, 1991; Chapter III), is similar to the core binding domain of sigma factors, subregion 2.1 (Lesley and Burgess, 1989; Shuler et al., 1995) (Fig. 2). The cryptic DNA-binding C-terminus of RAP30 is similar to subregions 3.1 and 4.1 of sigma factors, which are near DNA-binding subregions 2.4 and 4.2 (Dombroski et al., 1992) (Fig. 3). The C-terminal end of RAP30 has some similarity with the N-terminus of delta (Fig. 4).

2. The overall structure of yeast Cdc73p is similar to RAP30

The HCA plots of human RAP30, *Drosophila* RAP30 and yeast Cdc73p (the first 300 residues of 393) were compared (Fig. 5). The comparison of HCA plot of yeast Cdc73p with HCA plots of RAP30s revealed a similar distribution of short hydrophobic clusters over a length of about 300 residues, suggesting that Cdc73 might be structurally related to RAP30s. While the first 300 residues of Cdc73p only share 13% and 11% sequence identities with human RAP30 and *Drosophila* RAP30, respectively, their HCA homology scores are as high as 69% and 66%, respectively (Table 1). If the HCA

homology scores are more than 60%, proteins are found to have the same overall 3D fold, although their sequence identity may be as low as 10% (Gaboriaud et al., 1987). Furthermore, like human RAP30, yeast Cdc73p contains sequences similar to subregions 1.2, 2.1, 3.1 and 4.1 of sigma factors and the N-terminal region of *B. subtilis* delta protein (Fig. 1, 2, 3, 4). Notably, the similarity between the C-terminus of Cdc73p and the N-terminus of delta is more extensive than the similarity between human RAP30 and delta (Fig. 4).

In summary, based on HCA, both RAP30 and Cdc73p were found to contain sequences similar and co-linear with conserved regions of sigma factors. RAP30 and Cdc73p also have sequences similar to *B. subtilis* delta protein (Fig. 6).

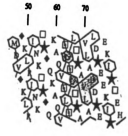
3. Multiple sequence alignment of RAP30 and Cdc73p

A multiple sequence alignment of Cdc73p, a *S. cerevisiae* RAP30 homologue Tfg2p (Henry et al., 1994), and *Drosophila* and human RAP30, based on HCA, is shown in Fig. 7.

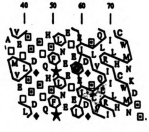
Figure 1. The N-terminal regions of RAP30 and Cdc73p are similar to subregion 1.2 of bacterial sigma factors. HCA plots and deduced sequence alignments of N-terminal regions of human RAP30, *Drosophila* RAP30 and *S. cerevisiae* Cdc73p, and subregion 1.2 of sigma A (*Anabena*). Prosposed similar hydrophobic residues are shaded. Conserved amino acids are grouped as follow: (I, L, M, V, F, Y, W), (H, K, R), (D, E, N, Q), (A, G), (S, T), and are underlined.

```
hRAP30      39  EVGKLRIAKTQG.....RTEVSFTLNEDLANIHDIGGK   71
dRAP30      46  DVGKLRINKTPG....QKAQVSLSLTPAVLALDPEEKI   79

Cdc73p      36  MVETLSSDGSTQHSFPLNEETEIEIDGSLVQLRIILHC   73

Sigma A     88  EIGRIRLLRADE.........EIELARKIADLLELERV   116
(Anabaena)
Tfg2p       91  ELGKIRINKEG.....SKITLLLNENDNDSIPHEYDLE   122

            |<------------------------------>|
                      subregion 1.2
```

hRAP30



dRAP30



Cdc73p

Sigma A

Figure 1

Figure 2. RNA polymerase II binding regions of RAP30 and Cdc73 are similar to the core binding subregion 2.1 of bacterial sigma factors. HCA plots and deduced sequence alignments of the central regions of human RAP30 and *Drosophila* RAP30, *S. cerevisiae* Cdc73p, and subregion 2.1 of sigma A (*Anabaena*) and sigma 70 (*E. coli*). Proposed similar hydrophobic residues are shaded. Bold indicates >50%, underline indicates 30-50%, and italic indicates 10-30% match to a consensus sequence calculated from 31 similar sigma factors. Stars indicate conserved residues between Cdc73p and RAP30. The sequences of RAP30 and Cdc73p required for RNAP II binding are indicated.

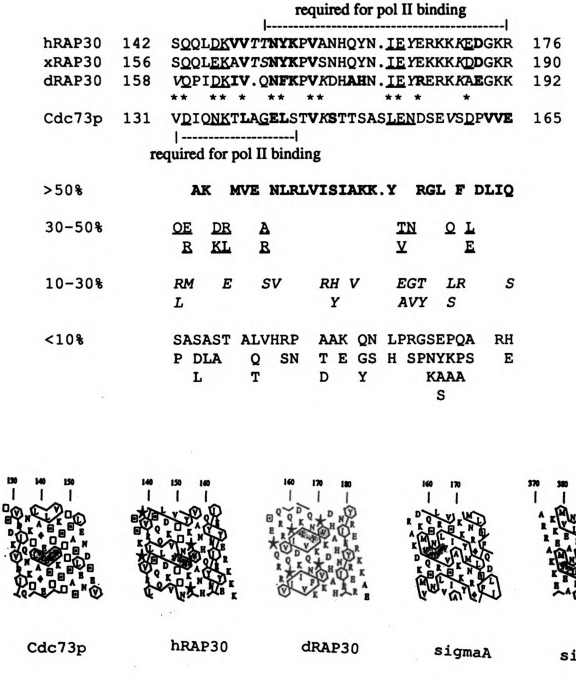
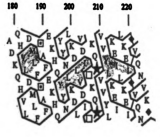

required for pol II binding
hRAP30 142 SQQLDKVVTTNYKPVANHQYN.IEYERKKKEDGKR 176
xRAP30 156 SQQLEKAVTSNYKPVSNHQYN.IEYEKKKKDDGKR 190
dRAP30 158 VQPIDKIV.QNFKPVKDHAHN.IEYRERKKAEGKK 192
Cdc73p 131 VDIQNKTLAGELSTVKSTTSASLENDSEVSDPVVE 165
required for pol II binding
>50% AK MVE NLRLVISIAKK.Y RGL F DLIQ
30-50% QE DR A TN Q L R KL R V E
10-30% RM E SV RH V EGT LR S L Y AVY S
<10% SASAST ALVHRP AAK QN LPRGSEPQA RH P DLA Q SN T E GS H SPNYKPS E L T D Y KAAA S
Cdc73p
hRAP30
dRAP30
sigmaA
sigma70

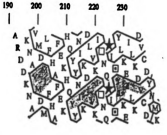
Figure 2

Figure 3. The C-terminal region of RAP30 is similar to subregions 3.1 and 4.1 of bacterial sigma factors. HCA plots and deduced sequence alignments of the C-terminal region of human RAP30 and *Drosophila* RAP30, and subregions 3.1 and 4.1 of sigma K (*B. subtilis*). Prosposed similar hydrophobic residues are shaded. Conserved amino acids are grouped as follows: (I, L, M, V, F, Y, W), (H, K, R), (D, E, N, Q), (A, G), (S, T), and are underlined.

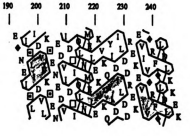
```
hRAP30   178   ADKQHVLDMLFSAFEKH.......  194
dRAP30   195   DDKNAVMDMLFHAFEKH.......  211

Cdc73p   182   GAKPINFGYLIKDAE.........  196

Sigma K  120   GNEISLIDVLKSENEDVIDTIQLN  143

                |<--- subregion 3.1 --->|


hRAP30   195   QYYNLKDLVDIT..KQPVVYLKEILKEIGV  221
dRAP30   212   QYYNIKDLVKIT..NQPISYLKEILKDVCD  238

Cdc73p   197   ....LKLVQSIK*****VANIKQFLLESKS  261

Sigma K  144   ELEKVKQYIDILDDREKEVIVGRFGLDLKK  173

                |<------subregion 4.1-------->|

*****: 46 a.a. deleted.
```



hRAP30



dRAP30

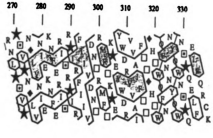


Sigma K

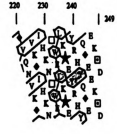
Figure. 3

Figure 4. The C-terminal regions of RAP30 and Cdc73 are similar to the N-terminal region of delta protein. HCA plots and deduced sequence alignments of the C-terminal regions of human RAP30, *Drosophila* RAP30 and yeast Cdc73p, and the N-terminal region of delta protein (*B. subtilis*). Prosposed similar hydrophobic residues are shaded. Conserved amino acids are grouped as follows: (I, L, M, V, F, Y, W), (H, K, R), (D, E, N, Q), (A, G), (S, T), and are underlined.

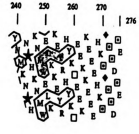
```
hRAP30     223 QNVKGIHK......NTWELKPEY...RHYQGEEKSD      249
dRAP30     239 YNMKNPHK......NMWELKKEY...RHYKTEEKKEEEHK 269
Cdc73p     266 RNLPSVPN......GLVNIFKNF...ERISRPIRFIIVDN 296
Delta      1   MGIKQYSQEELKEMALVEIAHELFEEHKKPVPFQELLNEI 40

Cdc73p     299 TRMF..TKPEYWDRVVAIFTTGH..........TWQFNN 323
Delta      41  ASLLGVKKEELGDRIAQFYTDLNIDGRFLALSDQTWGLRS 80

Cdc73p     324 YQ.WNSPQELFQR                          335
Delta      81  WYPYDQLDEETQP                          93
```



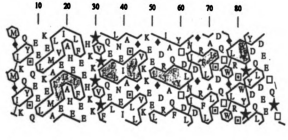
Cdc73p



hRAP30



dRAP30



delta

Figure. 4

Figure 5. HCA plots of human RAP30, *Drosophila* RAP30 and *S. cerevisiae* Cdc73p. Vertical lines indicate the proposed regions of similarity between the segments.

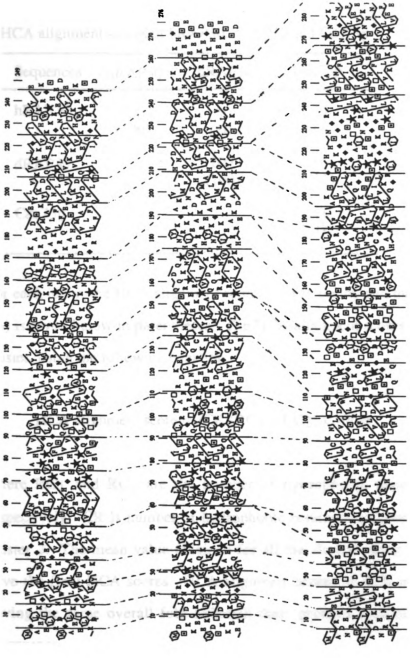


Figure 5

Table 1

HCA alignment scores of hRAP30, dRAP30 and Cdc73p

Sequences	hRAP30	dRAP30	Cdc73p
hRAP30	100 (100)	80 (50)	69 (13)
dRAP30		100 (100)	66 (11)
Cdc73p			100 (100)

For each entry, the HCA alignment score is given on top and the sequence identity score (%) is given below in parentheses. The HCA alignment scores were calculated for each cluster, using the following equation:

$$\text{HCA alignment score} = 2CR/(RC_1 + RC_2) \times 100\%$$

where RC_1 and RC_2 are the number of hydrophobic residues in cluster 1 and 2, respectively. CR is number of hydrophobic residues in cluser 1 that are also found in cluster 2. The mean value obtained for all the clusters along the sequences compared gave the final HCA scores. HCA alignment scores of 60% are found among proteins having the same overall fold although these proteins may have significant structural divergence.

Figure 6. The structure of human RAP30 and Cdc73p. The functional domains and the regions similar to sigma factors and delta protein are indicated.

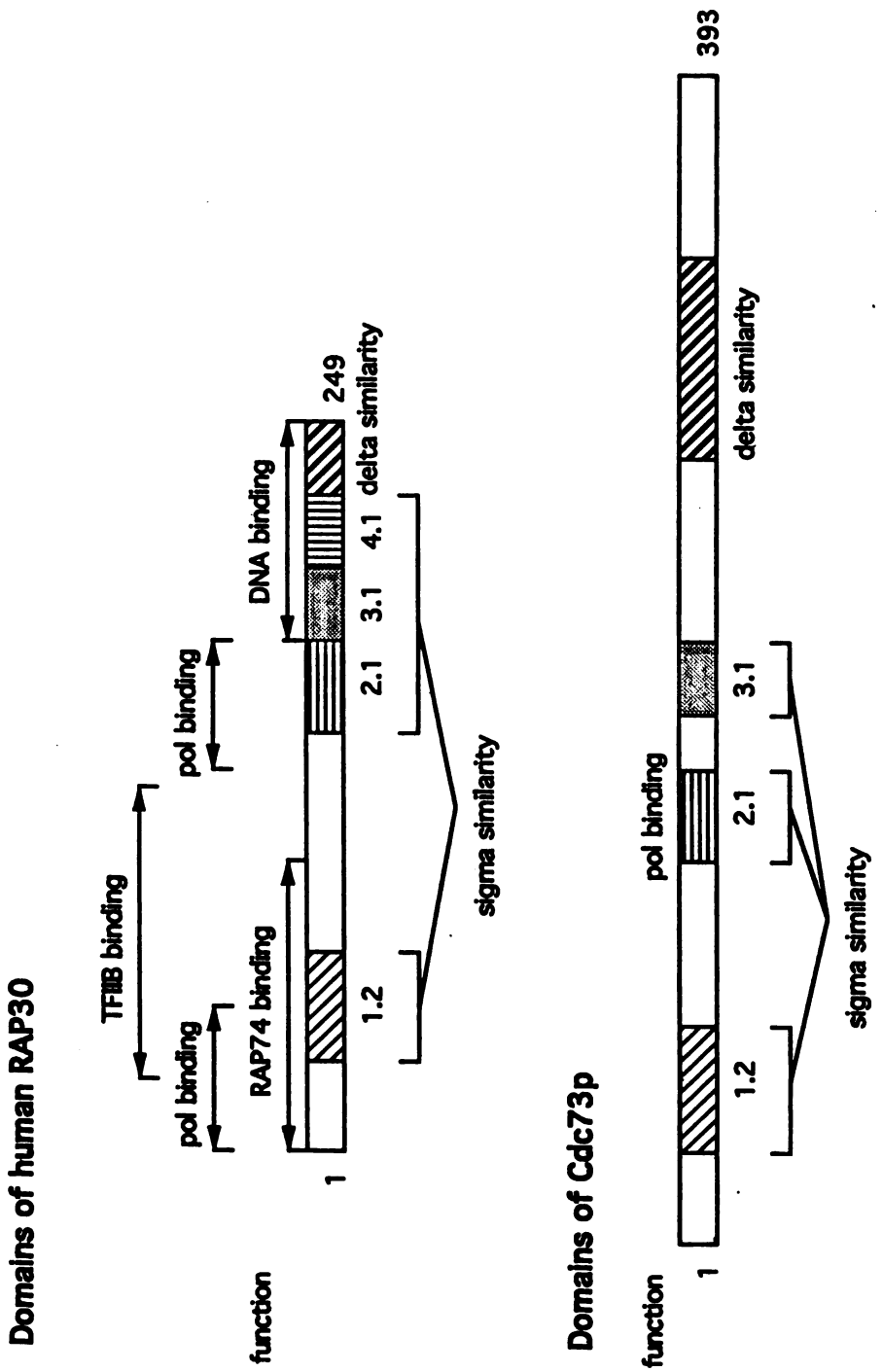

Domains of human RAP30
TFIIB binding
pol binding
RAP74 binding
pol binding
DNA binding
1
249
function
1.2
2.1
3.1
4.1
delta similarity
sigma similarity
Domains of Cdc73p
pol binding
1
393
function
1.2
2.1
3.1
delta similarity
sigma similarity

Figure 6

Figure 7. Multiple sequence alignment of human RAP30, *Drosophila* RAP30, *S. cerevisiae* Tfg2p, and Cdc73p, based on HCA. Gaps are indicated by dots. Conserved amino acids are grouped as follow: (I, L, M, V, F, Y, W), (H, K, R), (D, E, N, Q), (A, G), (S, T), and are underlined.

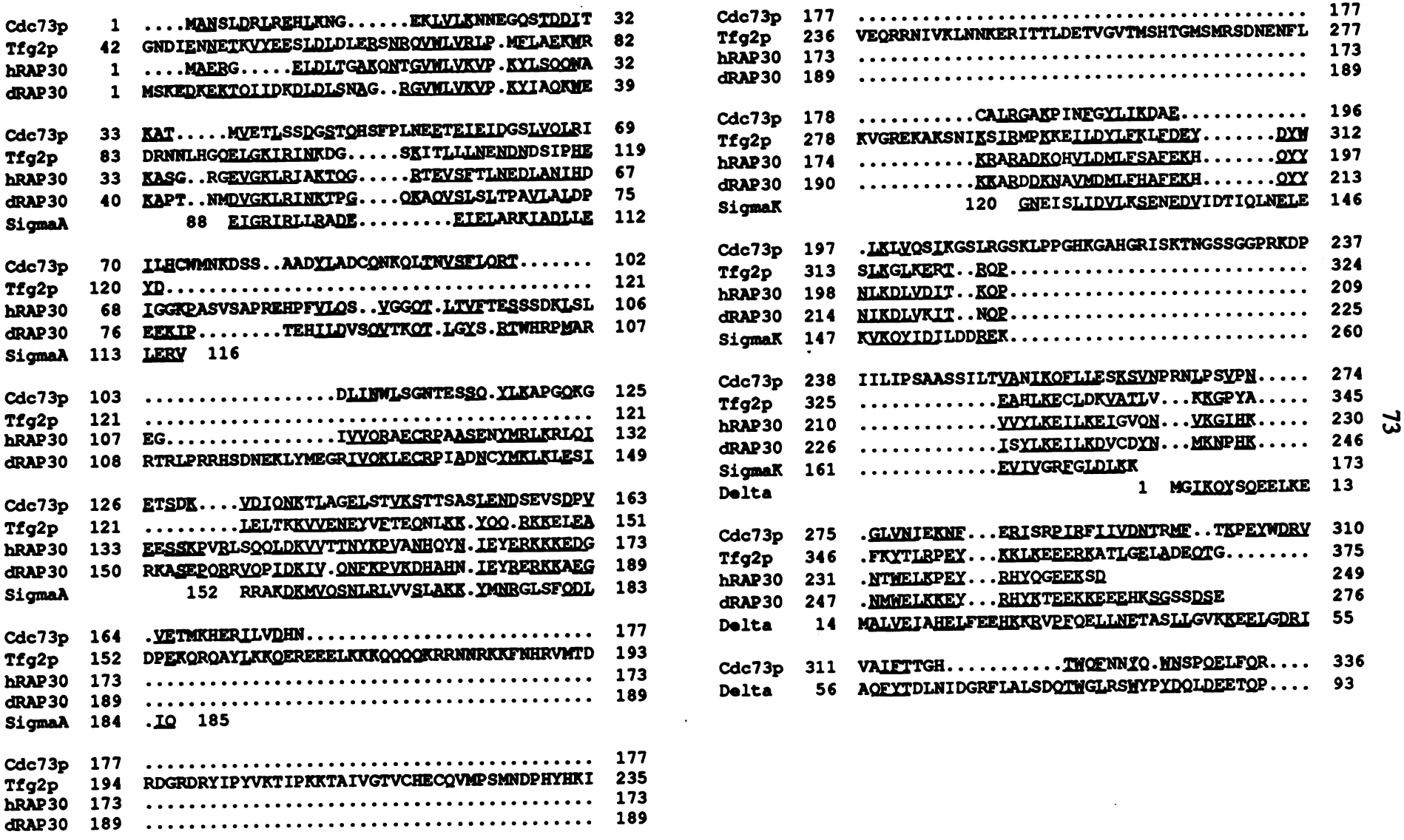


Figure 7

DISCUSSION

RAP30 can be considered a member of the sigma family. RAP30 binds to RNAP II (McCracken and Greenblatt, 1991), recruits RNA polymerase II to the pre-initiation complex (Flores et al., 1991; Killeen et al., 1992), prevents RNAP II from binding nonspecifically to DNA, and greatly inhibits nonspecific transcription by RNAP II (Killeen and Greenblatt, 1992). All these functions are analogous to the functions of bacterial sigma factors.

Several research groups have tried to identify the regions of RAP30 similar to conserved regions of sigma factors. Sopta et al. proposed that human RAP30 contained a region of similarity with sigma 70. The best alignment was between residues 118-165 of RAP30 and residues 373-425 of sigma 70, which corresponds to subregion 2.1 and 2.2. The second-best alignment involved part of region 1b (residues 94-117 of RAP30 and residues 111-139 of sigma 70) (Sopta et al., 1989). Since *Xenopus* RAP30 has an insertion of 12 amino acids within the region of human RAP30 aligned to region 1b, this alignment seemed unlikely to be correct (Gong et al., 1992). Garrett et al. proposed that the C-terminus (residues 150-249) of rat RAP30 was similar to regions 4.1 and 4.2 of SpoIIIC (*B. subtilis* sigma k) (Garrett et al., 1992). The C-terminus of RAP30, however, lacks the helix-turn-helix DNA-binding motif which is the primary feature of subregion 4.2. These previously reported sequence comparisons were based on linear alignment methods, and may not be reliable.

HCA sometimes results in significant alignment in cases where linear methods fail (Gaboriaud et al., 1987). Applying the HCA method, we have identified four sigma-similar regions and one delta-similar region within the RAP30 sequence (Figs. 1, 2, 3, 4). Notably, these regions are arranged in the same linear order in RAP30 as they are in sigma factors, suggesting that RAP30 and sigma factors might have similar structure.

The sequence similar to subregion 1.2 is near the N-terminus of RAP30. This is the region that binds RAP74 and RNAP II (Chapter III) and masks DNA binding by the RAP30 C-terminal region (Tan et al., 1994). This region of RAP30 is immediately distal to a demonstrated RAP74-binding sequence that is highly conserved among RAP30 from different species (Tan et al., 1995). Since subregion 1.2 of sigma factors masks DNA binding domains, this region of RAP30 might be important for masking the C-terminal DNA binding region. This region may also contribute to regulation of RAP74 and RNAP II binding (Chapter III).

The central region of RAP30 was previously proposed to contain a subregion 2.1 similarity (Sopta et al., 1989), but our alignment, which is deduced from HCA, is different from that which was previously proposed. The subregion 2.1 of sigma factors is the core binding domain (Lesley and Burgess, 1989; Shuler et al., 1995). There are several lines of evidence indicating that the region of RAP30 that we have aligned to subregion 2.1 may be involved in RNAP II binding. When RNAP II binds to RAP30, at least one of the serine residues at positions 135, 136, or 142 is protected from phosphorylation (McCracken and Greenblatt, 1991). These serines residues are near the aligned region. Mutagenic analysis suggests that the residues 131-159 and 152-176 of RAP30, which are located in this region, are important for RNAP II binding (Chapter III). A yeast Cdc73p mutant deleted 15 amino acids from this similar region decreases its RNAP II binding affinity dramatically (C.-h. Chang and Z. F. Burton, unpublished data). Thus, among these alignments, although the similarity between RAP30 and sigma in this region is weak, its functional role is the most clearly understood.

The C-terminal region of RAP30 includes a cryptic DNA binding domain (Tan et al., 1994). This region contains sequences similar to subregions 3.1 and 4.1 of sigma factors and the N-terminal region of delta protein. Subregions 3.1 and 4.1 of sigma factors may contribute to DNA binding by subregions 2.4 and 4.2 (Dombroski et al., 1992). Although delta can limit RNAP binding to DNA, there is no direct evidence that

delta itself binds to DNA. Functional domains of delta have not yet been elucidated, so the functional significance of this similarity is unclear. It is possible that N-terminal region of delta has DNA binding ability.

Like RAP30, sigma and delta, Cdc73p can be isolated by RNAP affinity chromatography (Wade et al., 1995). Cdc73p appears to be evolutionarily related to human RAP30, sigma factors and delta protein (Fig. 6). As expected, Cdc73p can bind to Anc1p/Tfg3p, the smallest subunit of yeast TFIIF (L. Lei and Z. F. Burton, unpublished data). Based on these observations, Cdc73p may be a homologue of yeast RAP30, but this function has been ascribed to another yeast protein Tfg2p (Henry et al., 1994). This raises the possibility that yeast, and possibly other eukaryotes, contains an alternative form of RAP30, much as bacteria have alternative sigma factors for transcription of different classes of genes.

REFERENCES

Achberger, E.C., and Whiteley, H.R. (1981). J. Biol. Chem. 256, 7424-7432.

Achberger, E.C., Hilton, M.D., and Whiteley, H.R. (1982). Nucl. Acids. Res. 10, 2893-2910.

Daniels, D., Zuber, R., and Losick, R. (1990). Proc. Natl. Acad. Sci. USA 87, 8075-8079.

Dickel, C.D., Burtis, K.C., and Doi, R.H. (1980). Biochem. Biophys. Res. Commun. 95, 1789-1795.

Dombroski, A.J., Walter, W.A., Record, M.T.Jr., Siegele, D.A., and Gross, C.A. (1992). Cell 70, 501-512.

Flores, O., Lu, H., Killeen, M.T., Greenblatt, J., Burton, Z.F., and Reinberg, D. (1991). Proc. Natl. Acad. Sci. USA 88, 9999-10003.

Gaboriaud, C., Bissery, V., Benchetrit, T., and Mornon, J.P. (1987). FEBS Letters 224, 149-155.

Gardella, T. Moyle, H., and Susskind, M.M. (1989). J. Mol. Biol. 206, 579-590.

Garrett, K.P., Serizawa, H., Hanley, J.P., Bradsher, J.N., Tsuboi, A., Arai, N., Yokota, T., Arai, K., Conaway, R.C., and Conaway, J.W. (1992). J. Biol. Chem. 267, 23942-23949.

Gong, D.-W., Hashimoto, S., Wada, K., Roeder, R.G., Nakatani, Y., and Horikoshi, M. (1992). Nucleic Acids Res. 20, 6414.

Gribskov, M. and Burgess, R.R. (1986). Nucl. Acids Res. 14, 6745-6761.

Helmann, J.D. (1994). Transcription: Mechanisms and Regulation, edited by Conaway, R.C., and Conaway, J.W., Raven Press, New York, 1-17.

Helmann, J.D., and Chamberlin, M.J. (1988). Ann. Rev. Biochem. 57, 839-872.

Henrissat, B., Popineau, Y., and Kader, J.C. (1988). Biochem. J. 255, 901-905.

Henry, N.L., Sayre, M.H., and Kornberg, R.D. (1992). J. Biol. Chem. 267, 23388-23392.

Hilton, M.D., and Whiteley, H.R. (1985). J. Biol. Chem. 260, 8121-8127.

Hinkle, D.C., and Chamberlin, M.J. (1972). J. Mol. Biol. 70, 157-185.

Hyde, E.I., Hilton, M.D., and Whiteley, H.R. (1986). J. Biol. Chem. 261, 16565-16570.

Ishihama, A. (1992). Mol. Microbiol. 6, 3283-3288.

Jones, C.H., and Moran, C.P. (1992). Proc. Natl. Acad. Sci. USA 89, 1958-1962.

Jones, C.H., Tatti, K.M., and Moran, C.P. (1992). J. Bacteriol. 174, 6815-6821.

Juang, Y.-L., and Helmann, J.D. (1994). J. Mol. Biol. 239, 1-14.

Killeen, M., and Greenblatt, J. (1992). Mol. Cell. Biol. 12, 30-37.

Killeen, M., Coulombe, B., and Greenblatt, J. (1992). J. Biol. Chem. 267, 9463-9466.

Lampe, M., Binnie, C., Schmidt, R., and Losick, R. (1988). Gene 67, 13-20.

Lemesle-Varloot, L., Henrissat, B., Gaboriaud, C., Bissery, V., Morgat, A., and Mornon, J.P. (1990). Biochimie 72, 555-574.

Lesley, S.A., and Burgess, R.R. (1989). Biochemistry 28, 7728-7734.

Lonetto, M., Gribskov, M., and Gross, C.A. (1992). J. Bacteriol. 174, 3843-3849.

McCracken, S., and Greenblatt, J. (1991). Science 253, 900-902.

Pero, J. Nelson, J., and Tox, T.D. (1975). Proc. Natl. Acad. Sci. USA 72, 1589-1593.

Reed, S.I., Ferguson, J., and Jahng, K.-Y. (1988). Cold Spring Harb. Symp. Quant. Biol. 53, 621-627.

Shuler, M.F., Tatti, K.M., Wade, K.H., and Moran, C.P.Jr. (1995). J. Bacteriol. 177, 3687-3694.

Siegele, D.A., Hu, J.C., Walter, W.A., and Gross, C. (1989). J. Mol. Biol. 206, 591-603.

Sopta, M., Burton, Z.F., and Greenblatt, J. (1989). Nature 341, 410-414.

Tan, S., Garrett, K.P., Conaway, R.C., and Conaway, J.W. (1994). Proc. Natl. Acad. Sci. USA 91, 9808-9812.

Tan, S., Conaway, R.C., and Conaway, J.W. (1995). Proc. Natl. Acad. Sci. USA 92, 6042-6046.

Tjian, R., Losick, R., Pero, J., and Hinnebush, A. (1977). Eur. J. Biochem. 74, 149-154.

Wade, P.A., Werel, W., Fentzke, R.C., Thompson, N.E., Leykam, J.F., Burgess, R.R., Jaehning, J.A., and Burton, Z.F. (1995). Mol. Cell. Biol., in press.

Waldburger, C., Gardella, T., Wong, R., and Susskind, M.M. (1990). J. Mol. Biol. 215, 267-276.

Zhou, Y.N., Walter, W.A., and Gross, C.A. (1992). J. Bacteriol. 174, 5005-5012.

Zuber, P., Healy, J., Carter, H.L. III, Cutting, S., Moran, C.P.Jr, and Losick R. (1989). J. Mol. Biol. 206, 605-614.

CHAPTER III

FUNCTIONAL DOMAINS OF RAP30

ABSTRACT

A set of deletion mutants of human RAP30, the small subunit of transcription factor IIF (TFIIF; RAP30/74), was constructed to map functional domains. Mutants were tested for accurate transcriptional activity, RAP74 binding, RNA polymerase II binding and TFIIB binding. Transcription assays indicate the importance of both N- and C-terminal sequences for RAP30 function. RAP74 binds to the N-terminal region of RAP30 between amino acids 1-98. RAP30 N-terminal sequence between 1-50 and central sequences between 131-159 and 152-176 are important for RNA polymerase II binding. TFIIB binds to an overlapping region of RAP30, localized to amino acids 1-176 by an affinity bead procedure. A minimal interaction domain was located between amino acids 27-118 by a more sensitive enzyme-linked immunosorbent assay.

INTRODUCTION

Transcription by RNA polymerase II requires auxiliary proteins termed general transcription factors, including TFIIA, TFIIB, TFIID, TFIIE, TFIIF, TFIIH, TFIIJ, TFIIS and SIII/elongin, for accurate initiation and efficient elongation. TFIIA, TFIIB, TFIID, TFIIE, TFIIF and TFIIH are involved in the early steps of initiation and promoter clearance, whereas TFIIF, TFIIS and SIII facilitate elongation of RNA chains. Most studies of the functions of general transcription factors have focused on their interactions with DNA template, transcript, RNAP II, other general transcription factors and regulatory factors. It has been postulated that the order of assembly of the initiation complex begins with the binding of TFIID, TFIIA and TFIIB to the TATA box and start site sequences, followed by the assembly of RNAP II bound to TFIIF to form a minimal transcription machinery capable of accurate initiation on supercoiled templates. The addition of TFIIE, TFIIH and TFIIJ yields a complete basal transcription initiation complex capable of initiation on linear template (reviewed by Zawel and Reinberg, 1993; Conaway and Conaway, 1993).

In this study, we focus on the RAP30 subunit of TFIIF. TFIIF was first identified and purified by affinity chromatography on a column containing covalently immobilized RNAP II (Sopta et al., 1985). It was named RAP30/74, for RNA polymerase II-associating proteins of 30 and 74 kDa apparent molecular mass. Further studies show that TFIIF is a heteromeric complex of subunits RAP30 and RAP74 that stimulates both initiation and elongation of transcription (Burton et al., 1988; Flores et al., 1988, 1990, 1991; Conaway and Conaway, 1989; Conaway et al., 1991; Kitajima et al., 1990). The RAP30 subunit binds directly to RNAP II, reduces the nonspecific binding of RNAP II to DNA, and helps bring RNAP II into the pre-initiation complex (Flores et al., 1991; Conaway et al., 1991; Killeen et al., 1992). RAP30 has also been shown to interact with

TFIIB, and this interaction may be important for pre-initiation complex assembly (Ha et al., 1993). RAP30 is required for accurate initiation using an extract transcription system and linear DNA templates (Chang et al., 1993). With a better defined system and highly supercoiled templates, RNAP II can accurately initiate transcription in the absence of TFIIF (Parvin and Sharp, 1993; Usheva and Shenk, 1994). RAP74 has been found to be dispensable for initiation under some conditions (Chang et al., 1993; Parvin and Sharp, 1993; Tyree et al., 1993; Usheva and Shenk, 1994), and required for RNAP II to escape from the promoter (Chang et al., 1993). TFIIF also stimulates the elongation rate of RNAP II (Bengal et al., 1991; Izban and Luse, 1992).

Several research groups used a mutagenic approach to elucidate the functional domains of RAP30. It has been proposed that RAP30 contains a N-terminal RAP74 binding region (Yonaha et al., 1993; Frank et al., 1995), a central RNAP II binding region (McCracken and Greenblatt, 1991), and a cryptic C-terminal DNA binding region (Tan et al., 1994). Here we report the further characterization of RAP74 binding and RNAP II binding domains of RAP30. The homodimerization and TFIIB binding domains are also identified.

MATERIALS AND METHODS

Construction of RAP30 mutants

Fragments of a human *RAP30* cDNA (Sopta et al., 1989) were subcloned into pET23d (Novagen) and transformed into *E. coli* BL21(DE3) for expression (Wang et al., 1993). To construct RAP30(1-227), an NcoI to BamHI (blunt: end-filled with Klenow DNA polymerase I and deoxynucleoside triphosphates) fragment was subcloned between the NcoI and HincII sites of the vector. To construct RAP30(1-176) an NcoI to XhoI (blunt) fragment was cloned between the NcoI and HincII sites of the vector. These mutants have a DKLAAALEHHHHHH C-terminal extension. To construct RAP30(1-118), an NcoI to PvuII fragment was subcloned between the NcoI and XhoI (blunt) sites of the vector. This mutant has an HHHHHH C-terminal extension. To construct RAP30(1-98) an NcoI to SstI (blunt) fragment was subcloned between the NcoI and HincII sites of the vector. This mutant has a DKLAAALEHHHHHH C-terminal extension. RAP30(1-249), RAP30(27-249), RAP30(51-249), RAP30(110-249), RAP30(160-249), RAP30(1-152) and RAP30(27-152) were constructed by PCR amplification using an upstream primer containing an engineered NcoI site and a downstream primer containing an engineered NotI site. The NcoI to NotI fragments were subcloned between the same sites of the vector. These mutants have a AALEHHHHHH C-terminal extension. The internal deletion mutant RAP30(Δ131-159) was constructed as follows. The 1-130 fragment was constructed by PCR amplification using an upstream primer containing an engineered NcoI site and a downstream primer containing an engineered Pst I site, and the 160-249 fragment was constructed by PCR amplification using an upstream primer containing an engineered PstI site and a downstream primer containing an engineered NotI site. These two fragments were ligated at the engineered

PstI sites and subcloned between NcoI and NotI of the vector. The mutant has a AALEHHHHHH C-terminal extension.

RAP30(1-249), RAP30(1-227), RAP30(1-176), RAP30(1-152), RAP30(1-118), RAP30(1-98), RAP30(27-249), RAP30(27-152), and RAP30(Δ131-159) accumulated within inclusion bodies and were purified by methods previously described for full-length RAP30 (Wang et al., 1993). RAP30(51-249), RAP30(110-249) and RAP30(160-249) were in the soluble fraction and were purified by Ni^{2+}-affinity chromatography (Qiagen). RAP30(110-249) was further purified by Mono S (Pharmacia) chromatography. Protein concentrations were determined using molar extinction coefficients calculated on the basis of aromatic amino acid composition (Luthy and Eisenberg, 1990).

A production vector for human TFIIB was the kind gift of D. Reinberg and R. Tjian and this protein was prepared by a protocol supplied by D. Reinberg (Ha et al., 1991).

In vitro transcription

The transcriptional activity of RAP30 mutants was tested as previously described (Wang et al., 1993). A TFIIF-depleted extract was prepared by immunodepletion using anti-RAP30 antibodies. This extract was combined with Adenovirus major late promoter template DNA (pML digested with SmaI, 50 μg/ml), RAP74, and RAP30 wild type or mutant, and pre-incubated for 1 hour, in buffer containing 12 mM Hepes pH 7.9, 12% glycerol, 60 mM KCl, 12 mM $MgCl_2$, 0.2 mM EDTA, and 3.2 mM EGTA. 600 μM ATP, CTP, UTP and 25 μM α^{32}P-GTP were added to the reaction and transcription was allowed to proceed for 30 min. Transcripts were isolated and analyzed by polyacrylamide gel electrophoresis in 50% (w/v) urea and 1x TBE. The accurately initiated transcript (217 nucleotides) was quantitated using a Molecular Dynamics phosphorimager.

Gel filtration analysis of RAP30 mutants

Storage Buffer (SB) was used for gel filtration and all binding reactions. SB contains 20 mM Hepes pH 7.9, 20 % glycerol w/v, 1 mM EDTA, 1 mM EGTA, and variable KCl concentration (i.e.. SB 0.1 contains 0.1 M KCl). The purified recombinant RAP30 mutants were denatured in SB 0.5 containing 4 M urea for 1 hour at room temperature. The samples were dialyzed with SB 0.5 and subsequently applied onto a Waters Protein Pak 300SW sizing column (8.0 x 300mm) previously equilibrated with SB 0.5. The samples were eluted at 0.5 ml/min, and fractions were collected, precipitated with 125 μg/ml Na-deoxycholate and 6% w/v trichloro-acetic acid for 1 hour on ice. Pellets were collected, resuspended in SDS-PAGE sample buffer and electrophoresed.

Protein-protein interaction assays

RAP74 binding to RAP30 mutants. 200 pmol RAP30 and RAP74 were incubated 1 hr at room temperature in 0.5 ml SB 0.1 containing 0.2% BSA. 10 μl Ni^{2+} resin, pre-incubated with SB 0.1 containing 0.2% bovine serum albumin, was added and incubated with tumbling at 4°C for 1 hour. RAP30 mutants bound the Ni^{2+} resin through their histidine tags. Beads were washed three times with 1 ml SB 0.25. Bound proteins were eluted with 50 μl SDS-PAGE sample buffer. 10 μl of the eluate was electrophoresed by 15% SDS-PAGE and blotted to nitrocellulose. RAP74 was detected in a Western blot developed with anti-RAP74 antiserum.

RNAP II binding to RAP30 mutants. calf thymus RNAP II was immobilized on Affi-gel 10 (Bio-Rad) at a density of about 1 mg RNAP II per ml resin. 50 μl RNAP II beads and 300 pmol RAP30 mutant were incubated at 4°C for 1 hour in 0.5 ml SB 0.1 containing 0.2% BSA. Beads were washed 3 times with 1 ml SB 0.3, and bound proteins were eluted in 50 μl SB 0.5. 30 μl of this eluate was analyzed on a 15% SDS-PAGE gel

developed with silver nitrate. Affi-gel beads without bound protein ligand were used as a negative control.

TFIIB binding to RAP30 mutants. TFIIB was immobilized on Affi-gel 10 (Bio-Rad) at a density of about 1 mg TFIIB per ml resin. 20 µl TFIIB beads and 300 pmol RAP30 mutant were incubated at 4°C for 1 hour in 0.5 ml SB 0.1 containing 0.2% BSA. Beads were washed 3 times with 1 ml SB 0.25, and bound proteins were eluted in 50 µl SB 0.5. 30 µl of this eluate was analyzed on a 15% SDS-PAGE gel developed with silver nitrate. Affi-gel beads without bound protein ligand were used as a negative control.

ELISA assays to observe RAP30-TFIIB interaction were done as described with minor modifications (Marsalek and Kaguni, 1994). Microtiter wells (Bectron Dickinson; Pro-Bind) were coated overnight with the protein to be immobilized (5 µg/ml) in 100 µl 50 mM sodium borate pH 9.0, at 4°C. Wells were washed three times with 200 µl PBS containing 0.2% BSA and 0.05% Tween-20 (PBSBT) and blocked with 200 µl PBSBT for an hour at room temperature. Proteins were added in 50 µl SB 0.1 containing 0.2% BSA and incubated for 15 min at room temperature. Glutaraldehyde was added to a final concentration of 1 % for 30 min to cross-link protein-protein interactions. Wells were washed three times with 200 µl PBSBT. 100 µl of rabbit antiserum (1:1000 diluted) directed against the mobile phase protein was added to each well and incubated two hours at room temperature. Wells were washed three times with PBSBT, and 100 µl HRP-conjugated goat anti-rabbit secondary antibody (Bio Rad; 1:3000 diluted) was added and incubated for one hour at room temperature. Wells were washed 3 times with PBSBT. Color was developed with 100 µl 50 mM sodium citrate (pH 4.0), 0.03% H_2O_2 and 0.4 mg/ml o-phenylenediamine. Color development was stopped by addition of 100 µl 4 N H_2SO_4. The reaction was measured for absorbance at 490 nm using a plate reader (Bio Tek Instruments; EL310). All determinations were done in duplicate and reported as average values.

RESULTS

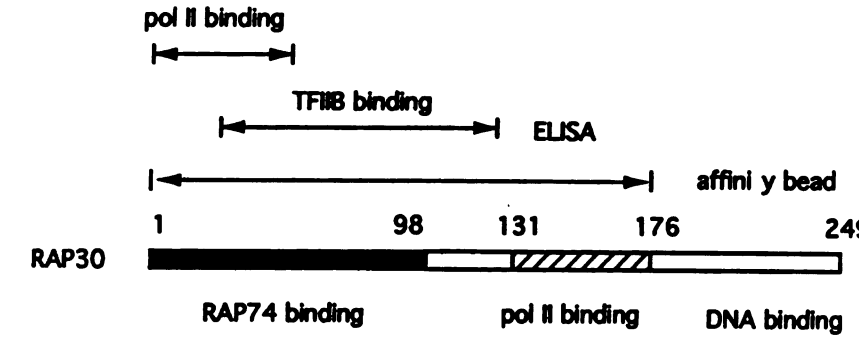
Both N- and C-terminal regions of RAP30 are important for transcription activity

A set of RAP30 deletion mutants was constructed with C-terminal histidine tags to aid in purification and binding reactions (Fig. 1). Each of the RAP30 mutants was tested for the ability to stimulate accurate runoff transcription from the Adenovirus major late promoter (Fig. 2). Full-length RAP30(1-249) was much more active in this assay than any of the deletion mutants. RAP30(1-249) is distinct from normal RAP30 because it has a C-terminal histidine tag. RAP30(1-227) and (1-176) supported a much lower level of activity. RAP30(1-152) had barely detectable activity and RAP30(1-118) was inactive (data not shown). These results are consistent with previously published data from the Conaway laboratory that indicate that the C-terminus of RAP30 is involved in DNA binding and is important for accurate initiation (Garrett et al., 1992; Tan et al., 1994). Activity was also very sensitive to deletion of N-terminal sequences. RAP30(27-249) supported weak activity (Fig. 2), but RAP30(51-249) was inactive (data not shown). Using a set of RAP30 mutants with short internal deletions, Conaway and co-workers have recently shown that amino acid sequences between 16-30 and 136-210 are critical for initiation of transcription (Tan et al., 1995). Our immunodepleted extract system is somewhat more tolerant of C-terminal region mutations than this reconstituted system (see Discussion).

Modification at the C-terminus causes RAP30 to behave as a monomer

Attaching a histidine tag to the C-terminus of RAP30(1-249) caused this protein to stimulate transcription at lower concentrations than RAP30 with no tag, although activity saturated at a similar level at higher concentrations (Fig. 3). The C-terminal histidine tag causes freshly renatured RAP30 to behave as a monomer in gel filtration, whereas unmodified RAP30 is a dimer (Table 1). As a monomer, RAP30 with the C-

Figure 1. RAP30 deletion mutants. RAP30(1-249) is the histidine-tagged version of RAP30. Accurate transcription (tx.) was determined from the Adenovirus major late promoter (Figure 2). RAP74 binding to RAP30 was determined using a Ni^{2+}-affinity bead procedure (Figure 4). RNAP II binding to RAP30 was determined using affinity beads (Figure 5). TFIIB binding to RAP30 was determined by affinity bead and ELISA procedures (Figures 6 and 7). (+) indicates high activity; (+/-) indicates low activity; (?) indicates barely detectable activity; and (-) indicates no detectable activity; (n.d.) indicates that no determination was made for a particular mutant; (*) indicates data is not shown in this report.



	x. ac ivi y	RAP74 binding	TFIIB binding		pol II binding
			ELISA	bead	
RAP30	100%	+	+	n.d.	+
1-249	>100%	+	+	+	+
1-227	30%	+*	+	n.d.	n.d.
1-176	20%	+	+	+	+
1-152	(< 5%)	+	+/-	+/-	+/-
1-118	0	+	+/-	-	+/-
1-98	0	+	-	-	+/-
27-249	10%	-	+	-	?
51-249	0	-	?	-	-
110-249	0	-	?	-	-
160-249	n.d.	n.d.	n.d.	-	-
27-152	0	-	+	-	n.d.
Δ131-159	n.d.	n.d.	n.d.	+*	+/-

Figure 1

Figure 2. Transcriptional activities of RAP30 deletion mutants. The ability of RAP30 and RAP30 deletion mutants to support accurate transcription initiation was determined from the Adenovirus major late promoter. An extract derived from human HeLa cells was depleted of TFIIF by immunoprecipitation with anti-RAP30 antibodies. This extract was supplemented with RAP74 (3.4 pmol) and the indicated amount of RAP30 or RAP30 mutant. The accurately initiated transcript was quantitated using a phosphorimager. Data is shown only for those mutants for which accurate transcription was high enough to be quantitated.

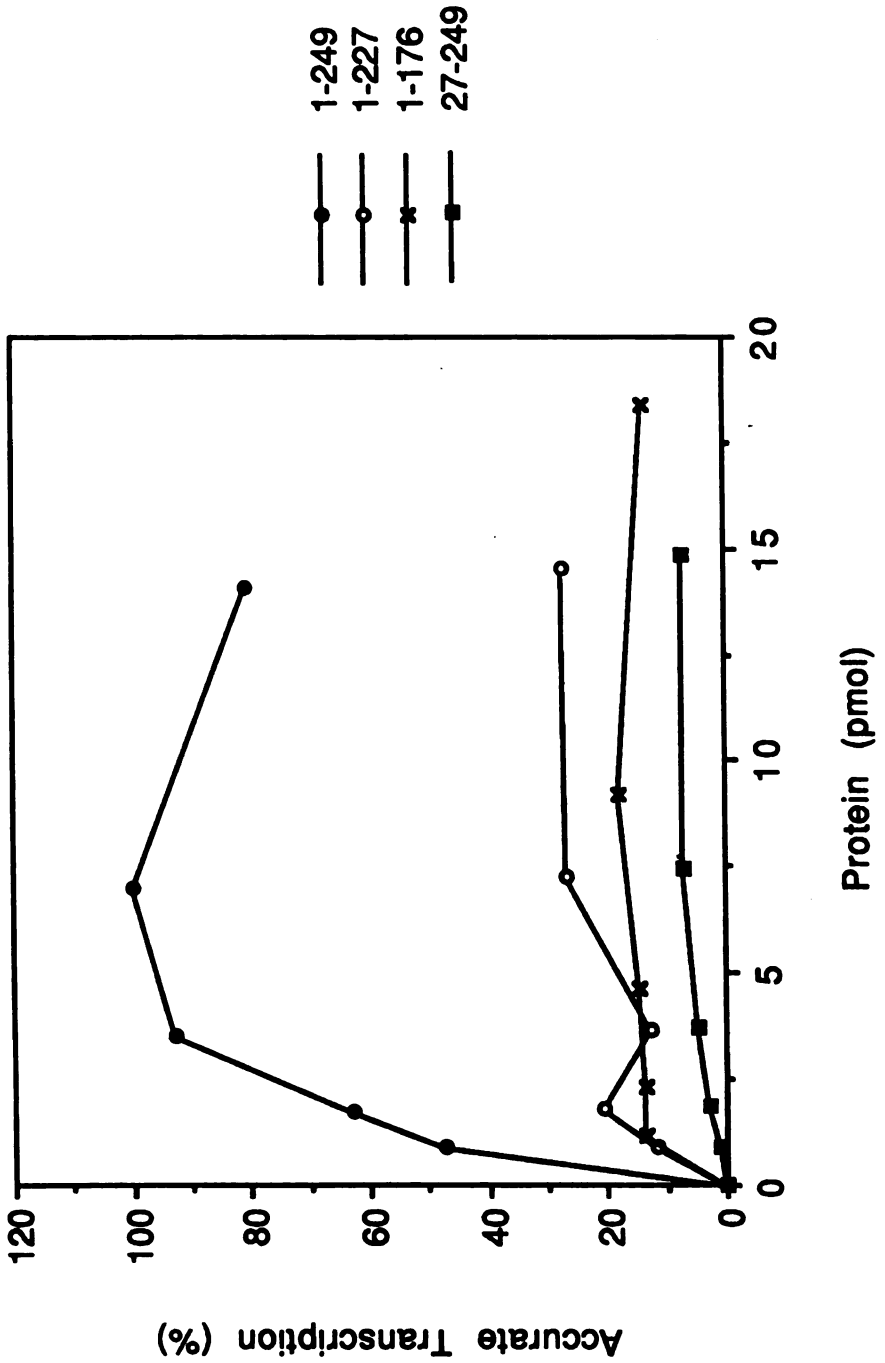

1-249
1-227
1-176
27-249
Accurate Transcription (%)
120
100
80
60
40
20
0
0
5
10
15
20
Protein (pmol)

Figure 2

Figure 3. C-terminal modification affects RAP30 activity. When full length RAP30 with a six histidine extension [RAP30(1-249)] was compared to unmodified RAP30 in run-off transcription assay, a reduced amount of RAP30(1-249) was required to saturate the assay.

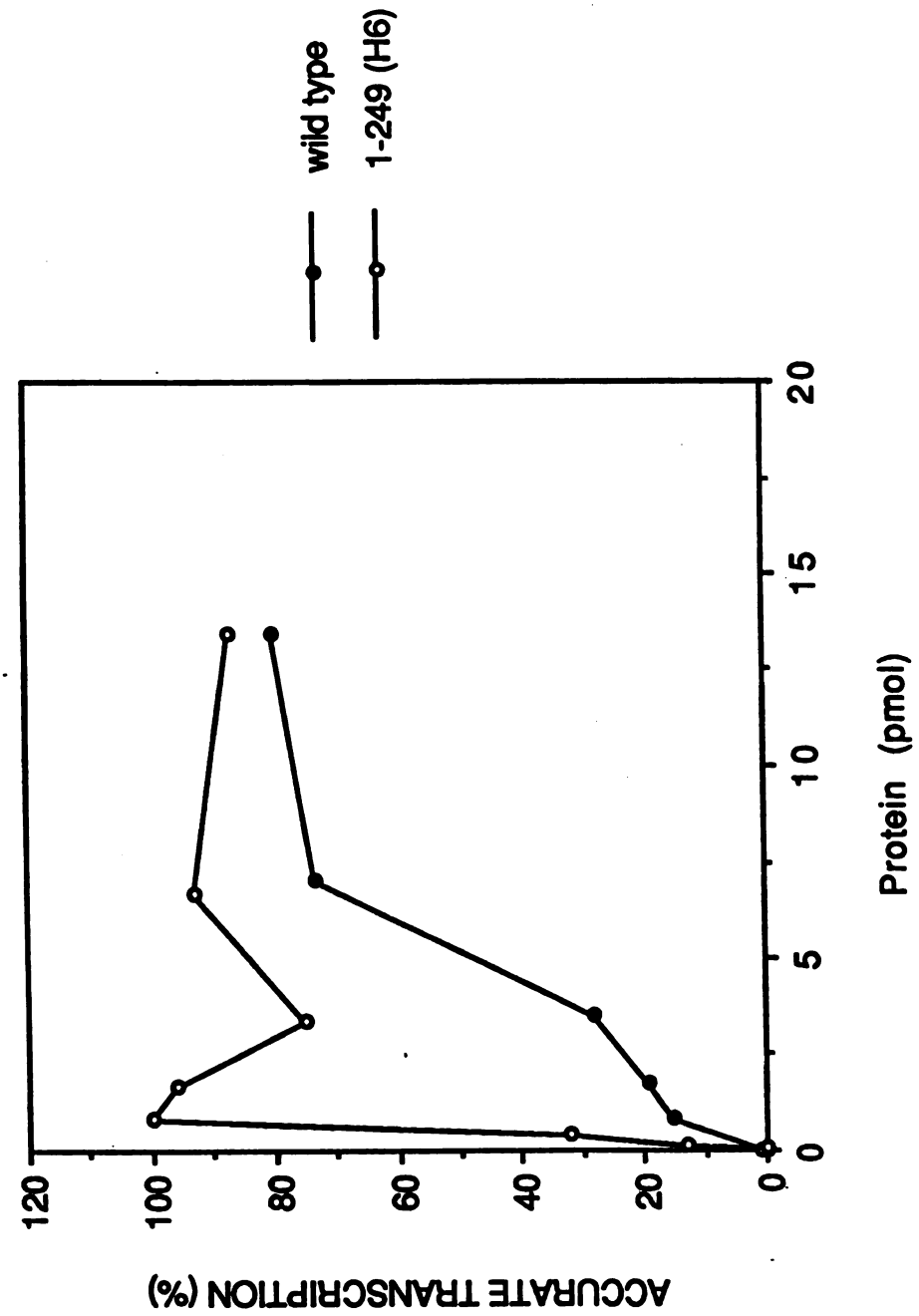

wild type
1-249 (H6)
ACCURATE TRANSCRIPTION (%)
120
100
80
60
40
20
0
0
5
10
15
20
Protein (pmol)

Figure 3

Table I. Molecular weight of RAP30 C-terminal deletion mutants determined by gel filtration

RAP30 mutants	elution time	apparent MW (KD)	deduced MW (KD)	dimer or monomer
WT RAP30	18'32"	55	28.4	dimer
RAP30(1-249)	20'16"	27	29.7	monomer
RAP30(1-227)	21'	21	27.2	monomer
RAP30(1-176)	21'22"	18	21.4	monomer
RAP30(1-118)	22'51"	10	13.8	monomer
RAP30(1-98)	23'18"	8.6	12.2	monomer

The purified recombinant RAP30 mutants were denatured in 4 M urea and dialyzed in 0.5 M KCl transcription buffer. The samples were applied onto a Waters Protein Pak 300SW sizing column (8.0 x 300 mm). The apparent molecular weight of proteins was determined by a standard curve.

terminal tag may more easily make appropriate contacts with other transcription factors and template in assembly of the pre-initiation complex.

RAP74 binds to the N-terminal region of RAP30 between amino acids 1-98

The RAP74 binding site on RAP30 was mapped using an affinity bead procedure (Fig. 4). RAP74 without a histidine tag was combined with RAP30 or a RAP30 mutant that contained a C-terminal tag. The mixture was combined with Ni^{2+}-affinity beads. After washing, bound proteins were eluted with SDS-PAGE sample buffer and analyzed in a Western blot developed with anti-RAP74 antiserum. Amino acid sequences between 1-98 of RAP30 were sufficient to observe binding between RAP30 and RAP74. Attempts to produce RAP30 mutants with more extensive C-terminal deletions were not successful, so no such mutants were tested. Even the shortest N-terminal deletion (27-249) of RAP30 was unable to bind to RAP74. These results are very similar to those previously reported by Yonaha et al. (Yonaha et al., 1993). In this previous report, interaction between RAP30 and RAP74 was determined using a 2-hybrid reporter system in contrast to the direct binding assay shown here.

Both N-terminal and central regions of RAP30 are important for polymerase binding

RAP30 sequences that are important for binding to RNA polymerase II were analyzed by binding RAP30 mutants to RNA polymerase II immobilized on agarose beads. The binding reaction was initially done in SB 0.1 and then beads were washed in SB 0.3 (Figure 5). These experiments indicate that both N-terminal sequences between amino acids 1-50 and central sequences, including those between amino acids 131-159, are important for polymerase binding. Using the 0.3 M KCl wash condition, only RAP30(1-249) and (1-176) bound RNA polymerase II tightly. Much weaker binding was observed with RAP30(1-152), (1-118), and (1-98). This result indicates that part of the

Figure 4. Sequences between amino acids 1-98 of RAP30 are sufficent for RAP74 binding. RAP74 and RAP30 mutant (200 pmol each) were combined with Ni^{2+}-affinity beads to bind the histidine tag on the RAP30 mutant. After washing with SB 0.25, bound proteins were eluted with SDS-PAGE sample buffer and analyzed on a Western blot developed with anti-RAP74 antiserum.

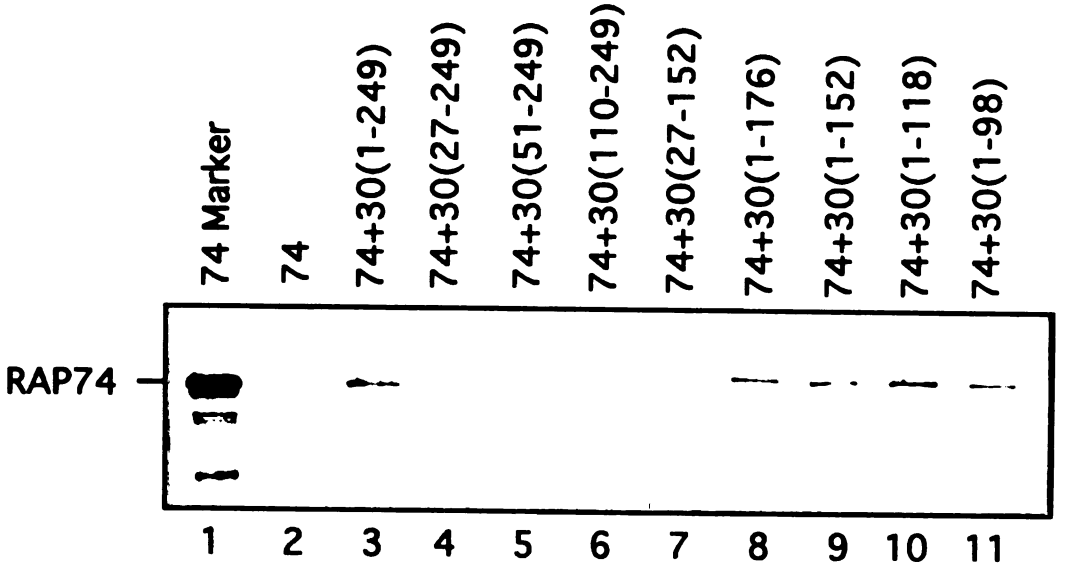

74 Marker
74
74+30(1-249)
74+30(27-249)
74+30(51-249)
74+30(110-249)
74+30(27-152)
74+30(1-176)
74+30(1-152)
74+30(1-118)
74+30(1-98)
RAP74
1
2
3
4
5
6
7
8
9
10
11

Figure 4

Figure 5. RAP30 sequences between 1-50, 131-159 and 152-176 are important for RNAP II binding. 300 pmol RAP30 or mutants was incubated with affinity beads containing covalently immobilized RNAP II. After washing with SB 0.3, bound protein was eluted and analyzed on an SDS-PAGE gel developed with silver nitrate.

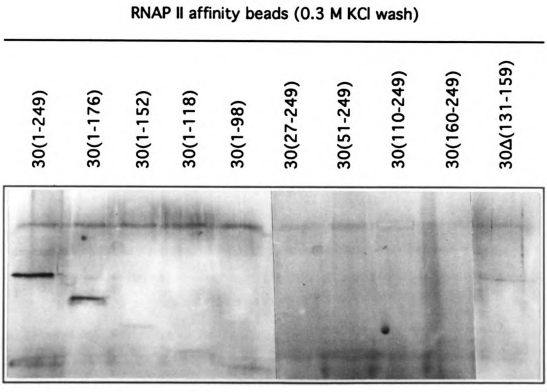

RNAP II affinity beads (0.3 M KCl wash)
30(1-249)
30(1-176)
30(1-152)
30(1-118)
30(1-98)
30(27-249)
30(51-249)
30(110-249)
30(160-249)
30Δ(131-159)

Figure 5

RNA polymerase II binding surface is located between amino acids 152-176. Because Hydrophobic Cluster Analysis (HCA) indicated that RAP30 might have sequence similarity to bacterial sigma factors within RNA polymerase-binding subregions 2.1 (Chapter 2, Figure 2), a small deletion mutant was constructed within this region of RAP30. HCA indicated that this region might include amino acids between positions 144-176. One deletion mutant in this region, RAP30(Δ131-159), was observed to bind RNA polymerase II weakly, with comparable apparent affinity to RAP30(1-152), (1-118), and (1-98).

That RAP30(1-98) binds weakly to RNA polymerase II indicates that N-terminal sequences are important for binding, and analysis of mutants from which these sequences are deleted reveals their importance. RAP30(27-249) was severely impaired for polymerase binding and RAP30(51-249) was not observed to bind. RAP30 sequence between amino acids 1-27, therefore, is important for binding, and sequence between amino acids 27-50 is essential, despite the presence of the binding region within the central portion of RAP30. This observation was of great interest to us, because of our proposed similarity between human RAP30 and homology subregion 1.2 of bacterial sigma factors (Chapter 2, Figure 1). This proposed similarity region lies between RAP30 amino acids 39-71, and this is the same region of RAP30 that interacts with RNA polymerase II. The function of sigma subregion 1.2 is not known, but we predict based on our results with RAP30 and our HCA analysis, that this region of sigma factors may contribute to polymerase binding.

Multiple contacts within the RAP74-binding and polymerase-binding regions of RAP30 contribute to TFIIB binding

RAP30 deletion mutants were tested for binding to TFIIB (Fig. 6 and Fig. 7). Binding was assayed using two methods that may differ in the sensitivity of detection for

Figure 6. RAP30 sequences between amino acids 1-176 are required for tight binding to TFIIB. RAP30(1-152) binds weakly to TFIIB. .300 pmol RAP30 (30) was incubated with affinity beads containing covalently immobilized TFIIB (IIB). After washing with SB 0.25, bound protein was eluted and analyzed on an SDS-PAGE gel developed with silver nitrate.

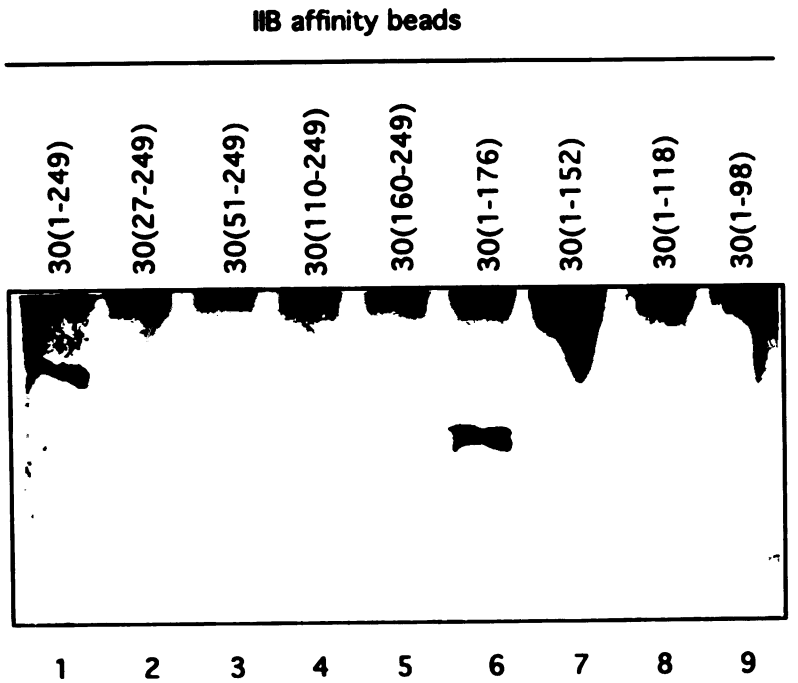
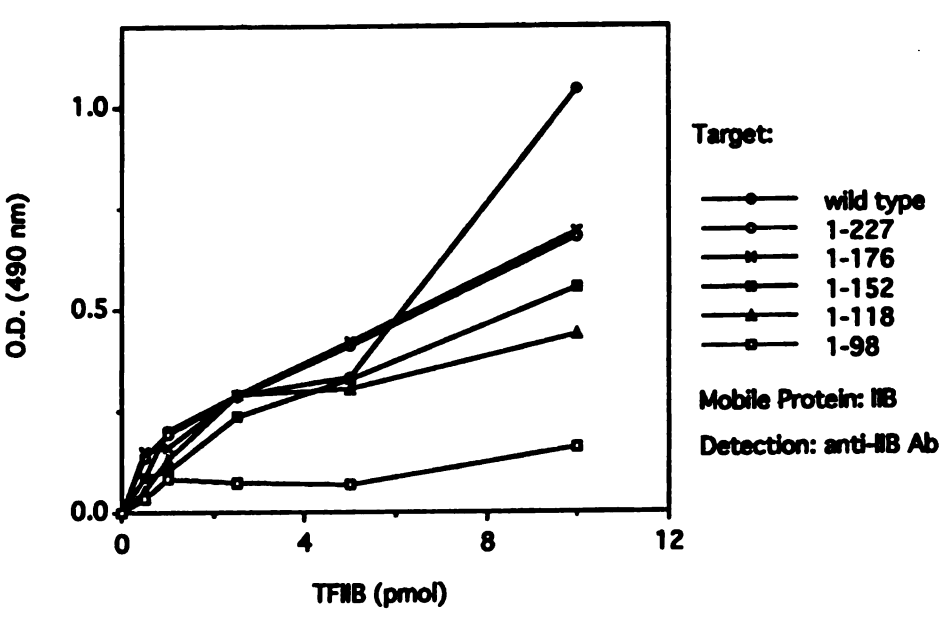

IIB affinity beads
30(1-249)
30(27-249)
30(51-249)
30(110-249)
30(160-249)
30(1-176)
30(1-152)
30(1-118)
30(1-98)
1
2
3
4
5
6
7
8
9

Figure 6

Figure 7. A minimal TFIIB-binding region of RAP30 maps between amino acids 27-118. RAP30 or RAP30 mutants (5 μg/ml) were immobilized in the wells of a microtiter dish. After blocking, increasing amounts of TFIIB were incubated in the wells as indicated. Binding was detected with anti-TFIIB antiserum. The O.D. (490 nm) varies between experiments because of variable time for color development. Wild type is RAP30 with no histidine tag. RAP30(1-249) has a C-terminal histidine tag.


O.D. (490 nm)
1.0
0.5
0.0
0
4
8
12
TFIIB (pmol)
Target:
wild type
1-227
1-176
1-152
1-118
1-98
Mobile Protein: IIB
Detection: anti-IIB Ab

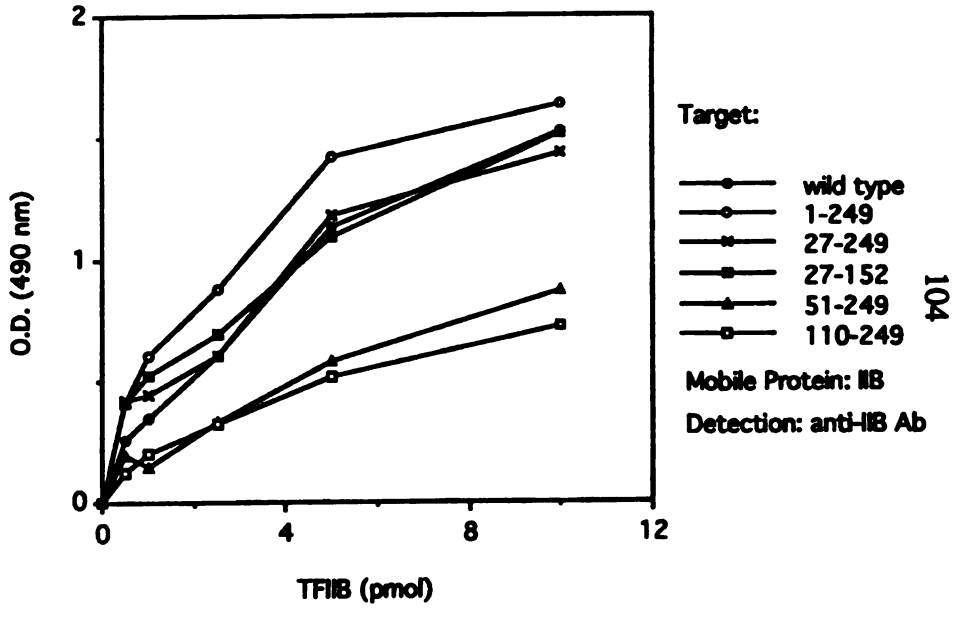

O.D. (490 nm)
2
1
0
0
4
8
12
TFIIB (pmol)
Target:
wild type
1-249
27-249
27-152
51-249
110-249
Mobile Protein: IIB
Detection: anti-IIB Ab

Figure 7

protein-protein interactions. For the experiment shown in Figure 6, TFIIB was covalently immobilized on agarose beads and used as an affinity adsorbant for RAP30 mutants. Binding reactions were done at 0.1 M KCl and washed at 0.25 M KCl. RAP30 and RAP30(1-176) bound most tightly to TFIIB (lanes 1 and 6). RAP30(1-152) also bound but much less efficiently (lane 7). None of the other deletion mutants was observed to bind. Apparently, an extensive surface of RAP30 contributes to TFIIB binding.

In Figure 7, an ELISA plate test was used to measure RAP30-TFIIB binding. Binding, in this case, was determined in buffer containing 0.1 M KCl. Microtiter plate wells were coated with RAP30 or a RAP30 mutant. After blocking, TFIIB was incubated with protein bound to the wells. After glutaraldehyde cross-linking and washing, bound TFIIB was detected with anti-TFIIB antiserum, using a horseradish peroxidase-linked second antibody and a colorimetric enzyme assay. Although interpretation of this experiment requires that binding of mutant RAP30 to microtiter wells be of similar efficiency (The unshown data suggest that RAP30 mutants appear to bind to microtiter wells with similar efficiency), ELISA appears to provide a reliable measure of interaction and a somewhat more sensitive detection of RAP30-TFIIB binding. By this analysis, RAP30(1-249), (1-227), (1-176), (27-249), and (27-152) bound TFIIB most tightly. RAP30(1-152) and (1-118) showed reduced affinity. RAP30(51-249) and (110-249) were further reduced in affinity, and RAP30(1-98) was not observed to bind. Based on ELISA, therefore, the core of the RAP30-TFIIB interaction may lie between amino acids 27-118 of RAP30, and sequences between 1-176 contribute to binding. These results appear to confirm and extend the results of the affinity bead binding procedure.

DISCUSSION

RAP30 appears to consist of three functional regions (Fig. 8). The N-terminal region binds RAP74 (Fig. 4) and contributes to RNA polymerase II binding (Fig. 5), the central region binds RNA polymerase II (Fig. 5, McCracken and Greenblatt, 1991), and the C-terminal region binds DNA (Tan et al., 1994). In the current work, we have mapped the functional domains of human RAP30 that are required for accurate transcription, that bind RAP74, that bind RNAP II and that bind TFIIB (see Figs. 8).

Yonaha et al. previously reported the RAP74 binding region of RAP30 (Yonaha et al., 1993). Our determination is consistent with theirs except that we have located a slightly less extensive surface on RAP30 for binding to RAP74. They mapped the region of 1-110 of RAP30 as the minimum sequence essential for interacting with RAP74 *in vivo* using a 2-hybrid gene reporter system. Our determination was based on direct *in vitro* contacts between purified proteins and we show that the N-terminal region of RAP30 between amino acids 1-98 is sufficient for RAP74 binding (Fig. 4). RAP30(27-249), which does not bind tightly to RAP74 (Fig. 4) or RNAP II (Fig. 5), has weak residual transcriptional activity (Fig. 2), so contacts with other factors and template may allow inefficient assembly of an active complex with this mutant. Since our assay is in an extract system, TFIIF is required both for initiation and elongation of the transcript, so mutants can potentially be defective in one or both of these processes (Chang et al., 1993).

TFIIB binds to an extensive region of RAP30 between amino acids 1-176 (Fig. 6). A much weaker interaction may be maintained with amino acids between 27-118, because RAP30(27-249) and (1-118) appear to interact weakly with TFIIB (Fig. 7). Apparently, multiple contacts within the RAP74-binding and RNA polymerase II-binding regions of RAP30 contribute to TFIIB binding. The central region of RAP30 is weakly

Figure 8. Functional domains of RAP30.

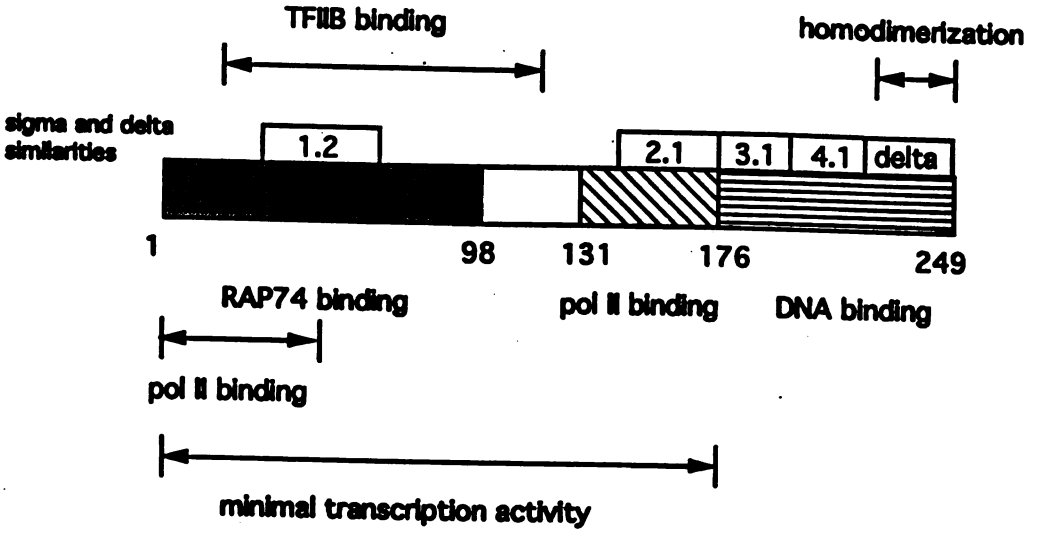

TFIIB binding
homodimerization
sigma and delta similarities
1.2
2.1
3.1
4.1
delta
1
98
131
176
249
RAP74 binding
pol II binding
DNA binding
pol II binding
minimal transcription activity

Figure 8

similar to the core binding region of bacterial sigma factors (Chapter II), suggesting that this region may be involved in RNAP II binding. Consistent with this suggestion, at least one of the serines residues at positions 135, 136, or 142 is protected from phosphorylation when RNAP II binds RAP30 (McCracken and Greenblatt, 1991). We further demonstrated that sequences between 131-159 and 152-176 are important for RNAP II binding (Fig. 5). This is the first direct evidence that the central region of RAP30 is indeed involved in RNAP II binding. Interestingly, the N-terminal region of RAP30 binds to RNAP II weakly, and deletion of the N-terminus abolishes the RNAP II binding activity (Fig. 5). Since deletion of the N-terminal region also eliminates RAP74 (Fig. 4) and TFIIB (Fig. 6) binding, and activates the DNA binding of the C-terminal region (Tan et al., 1994), we propose that in addition to binding RAP74 and RNAP II, the N-terminal region may be a regulatory domain that affects accessibility of central and C-terminal regions. The sigma subregion 1.2 similarity in this region may be important for this regulatory function as this region of sigma factors has been proposed to mask DNA binding by central and C-terminal sigma sequences (Chapter II).

The Conaway laboratory (Tan et al., 1995) has recently published an extensive set of short internal deletion mutants of RAP30 which they have tested for function in accurate initiation and elongation. They mapped a region between amino acids 16-30 that was critical for RAP74 binding, accurate initiation, and stimulation of elongation. Since both RAP30 and RAP74 are required in the reconstituted accurate initiation assay and in stimulation of the dC-tailed template elongation assay (Tan et al., 1994b), blocking RAP30-RAP74 interaction can inhibit both initiation and elongation. An adjacent deletion mutant from amino acids 31-45 fails to stimulate elongation but supports accurate initiation. Based on our results, we expect that these mutations also affect interactions with RAP74 and polymerase. Additional transcription factors (ie. TFIIB) in the reconstituted accurate initiation assay may compensate for suboptimal contacts between TFIIF subunits or between TFIIF and RNAP II. Since the dC tailed template

elongation assay is missing factors other than polymerase, template and TFIIF, this system may be more sensitive to these N-terminal region mutations. Mutants between amino acids 91-120 similarly fail to stimulate elongation on the dC-tailed template but support initiation. Since these mutants are expected to affect RNA polymerase II binding, this contact appears most important for elongation, as might be expected. Tight polymerase binding may be dispensable for initiation because of compensation by other factors in the reconstituted system. Mutations within the central polymerase binding and the C-terminal DNA binding regions, amino acids 136-240, eliminate accurate initiation but do not influence elongation. This result is more difficult to reconcile with our results, since we observe weak accurate transcription with deletion to 1-176 (Fig. 2). Apparently, the immunodepleted extract system is more tolerant of C-terminal deletions than the reconstituted system. Additional factors present in the extract but missing in the more purified system may partially compensate for RAP30-DNA contacts. Taken together, results from the Conaway laboratory indicate that for the elongation assay, RAP30 contacts with RAP74 and polymerase are most essential. For the initiation assay, RAP30 contacts with RAP74, polymerase and DNA are essential.

REFERENCES

Burton, Z.F., Killeen, M., Sopta, M., Ortolan, L.G., and Greenblatt, J. (1988). Mol. Cell. Biol. 8, 1602-1613.

Chang, C., Kostrub, C. F. and Burton, Z. F. (1993). J. Biol. Chem. 268, 20482-20489.

Conaway, J.W. and Conaway, R. C. (1989) J. Biol. Chem. 264, 2357-2362.

Conaway, R.C., and Conaway, J.W. (1993). Annu. Rev. Biochem. 62, 161-190.

Conaway, R.C., Garrett, K.P., Hanley, J.P., and Conaway, J.W. (1991) Proc. Natl. Acad. Sci. USA 88, 6205-6209.

Flores, O., Maldonado, E., Burton, Z.F., Greenblatt, J., and Reinberg, D. (1988) J. Biol. Chem. 263, 10812-10816.

Flores, O., Ha, I., and Reinberg, D. (1990) J. Biol. Chem. 265, 5629-5634.

Flores, O., Lu, H., Killeen, M., Greenblatt, J., Burton, Z.F., and Reinberg, D. (1991) Proc. Natl. Acad. Sci. USA 88, 9999-10003.

Frank, D.J., Tyree, C.M., George, C.P., and Kadonaga, J.T. (1995). J. Biol. Chem. 270, 6292-6297.

Garrett, K. P., Serizawa, H., Hanley, J. P., Bradsher, J. N., Tsuboi, A., Arai N., Yokota, T., Arai, K., Conaway, R. C. and Conaway, J. W. (1992). J. Biol. Chem. 267, 23942-23949.

Ha, I., Lane, W.S., and Reinberg, D. (1991) Nature (Lond.) 352, 689-695.

Ha, I., Roberts, S., Maldonado, E., Sun, X., Kim, L. U., Green, M. and Reinberg, D. (1993). Genes Dev. 7, 1021-1032.

Izban, M.G. and Luse, D. S. (1992) J. Biol. Chem. 267, 13647-13655.

Killeen, M., Coulombe, B., and Greenblatt, J. (1992) J. Biol. Chem. 267, 9463-9466.

Kitajima, S., Tanaka, Y., Kawaguchi, T., Nagaoka, T., Weissman, S. and Yasukochi, Y. (1990) Nucleic Acids Res. 18, 4843-4849.

Marsalek, J., and Kaguni, J.M. (1994) J. Biol. Chem. 269, 4883-4890

McCracken, S. and Greenblatt, J. (1991) Science 253, 900-902.

Parvin, J. D. & Sharp, P. A. (1993) Cell 73, 533-540.

Sopta, M. Carthew, R.W., and Greenblatt, J. (1985). J. Biol. Chem. 260, 10353-10361.

Sopta, M., Burton, Z. F. and Greenblatt, J. (1989) Nature 341, 410-414.

Tan, S., Garrett, K. P., Conaway, R. C. and Conaway, J. W. (1994). Proc. Natl. Acad. Sci. USA 91, 9808-9812.
Tan, S., Aso, T., Conaway, R.C., and Conaway, J.W. (1994b) J. Biol. Chem. 269, 25684-25691.

Tan, S., Conaway, R.C., and Conaway, J.W. (1995) Proc. Natl. Acad. Sci. USA 92, 6042-6046.

Tyree, C.M., George, C.P., Lira-DeVito, L.M., Wampler, S.L., Dahmus, M.E., Zawel, L., and Kadonaga, J.T. (1993). Genes Dev. 7, 1254-1265.

Usheva, A., and Shenk, T. (1994). Cell 76, 1115-1121.

Wang, B. Q., Kostrub, C. F., Finkelstein, A. and Burton, Z. F. (1993) Protein Expr. Purif. 4, 207-214.

Yonaha, M., Aso, T., Kobayashi, Y., Vasavada, H., Yasukochi, Y., Weissman, S. M. and Kitajima, S. (1993). Nucleic Acids Res. 21, 273-279.

Zawel, L. and Reinberg, D. (1993). Prog. Nucl. Acids Res. 44, 67-108.

CHAPTER IV

DYNAMIC INTERACTIONS BETWEEN SUBUNITS OF TRANSCRIPTION FACTOR TFIIF AND TFIIB

ABSTRACT

Both RAP30 and RAP74 bind independently to TFIIB. RAP74 antagonizes the interaction between TFIIB and RAP30, both by binding to RAP30 and by binding to TFIIB. Analysis of deletion mutants of RAP74 shows that a C-terminal region between amino acids 358-517 binds directly to TFIIB. This region of RAP74 also binds to RNA polymerase II. RAP30, therefore, binds TFIIB only in the absence of the RAP74 subunit. When the TFIIF complex is intact, TFIIF-TFIIB contact appears to be maintained through the RAP74 subunit. If these binding relationships are maintained within functional transcription complexes, dynamic interactions between TFIIF subunits and TFIIB may be a mechanism to separate RAP30 and RAP74 functions during various stages of the transcription cycle.

INTRODUCTION

A minimal pathway for assembly of pre-initiation complexes on RNA polymerase II-dependent promoters has been defined in vitro (Buratowski et al., 1989; Maldonado et al., 1990; Zawel and Reinberg, 1993, 1995; Serizawa et al., 1994). On promoters that include a TATA box, TBP (TATA-binding protein) binds first to this recognition sequence. TFIIA facilitates this interaction. TFIIB can enter the complex either by binding to these template-associated factors or by binding to RNA polymerase II. For a tight association with the complex, polymerase must first associate with TFIIF, so TFIIB and TFIIF cooperate to bring polymerase into the pre-initiation complex (Conaway et al., 1991; Flores et al., 1991; Serizawa et al., 1994; Zawel and Reinberg, 1995). TFIIE and TFIIH are additionally required for accurate initiation from linear DNA templates.

TFIIB is a single polypeptide of 33 kDa (Ha et al., 1991; Mailk et al., 1991) that can roughly be divided by sequence and mutational analysis into N- and C-terminal regions. At the N-terminus of TFIIB is a probable Zn^{2+}-binding sequence that is important for pre-initiation complex assembly (Buratowski and Zhou, 1993). Immediately adjacent to the Zn^{2+}-binding motif is a highly conserved sequence that is important for selection of transcriptional start sites (Pinto et al., 1994). This N-terminal region binds the RAP30 subunit of TFIIF and RNA polymerase II (Ha et al., 1993; Hisatake et al., 1993), and mutations in this region inhibit assembly of transcription intermediates, because they fail to interact with either TFIIF or polymerase (Buratowski and Zhou, 1993; Ha et al., 1993; Hisatake et al., 1993; Malik et al., 1993). The C-terminal region of TFIIB contains two imperfect direct repeats. Near the end of the first repeat is an amphipathic alpha helix with a basic charged face. Sequence surrounding this structure is important for binding TBP and the acidic activation domain of herpes simplex virus VP16, although binding sites for these two proteins are separable by

mutation (Roberts et al., 1993; Roberts and Green, 1994). In the absence of other factors, the N- and C-terminal domains of TFIIB are thought to interact, causing masking of protein binding regions. Binding of the acidic transcriptional activator VP16 to the C-terminal region appears to open up the TFIIB structure to expose the C-terminal region, which binds TBP, and the N-terminal region, which binds RNA polymerase II and TFIIF (Roberts and Green, 1994). Thus, TFIIB forms a bridge between TBP and RNA polymerase II/TFIIF, in assembly of the pre-initiation complex, and the assembly steps involving TFIIB are a target of transcriptional regulators. Recent determinations of an x-ray crystal structure of TATA-box DNA with TBP and TFIIB (Nikolov et al., 1995) and a nuclear magnetic resonance structure of TFIIB (Bagby et al., 1995) support this model for TFIIB function. These structures indicate that the N-terminal region of TFIIB is presented as a scaffold for subsequent assembly of RNA polymerase II/TFIIF into the pre-initiation complex (Nikolov et al., 1995; Bagby et al., 1995).

TFIIF is a heteromeric factor of 28 kD (RAP30) and 58 kD (RAP74) subunits (Flores et al., 1988; Conaway and Conaway, 1989; Kitajima et al., 1990; Flores et al., 1990), but there is some indication that in certain contexts these subunits may enter complexes as separate factors (Flores et al., 1991; Killeen et al., 1992; Chang et al., 1993; see Discussion).

Accurate initiation has been demonstrated from highly supercoiled templates using a system consisting of RNA polymerase II, TFIIB, and either TBP (Parvin and Sharp, 1993), or YY1, which is an initiator binding protein (Usheva and Shenk, 1994). One implication of these observations is that RNA polymerase II and TFIIB might minimally suffice to select transcriptional start sites. Consistent with this view, swapping *S. pombe* for *S cerevisiae* TFIIB and RNA polymerase II, in a system otherwise comprised of *S. cerevisiae* factors, shifts the position of the transcriptional start to that characteristic of *S. pombe* (Li et al., 1994). Also, some *sua7* mutants in the *S. cerevisiae* gene encoding TFIIB are altered for selection of initiation sites (Pinto et al., 1992). *sua8*

mutants, in the gene encoding the largest subunit of RNA polymerase II, affect transcriptional starts in a very similar way to these *sua7* mutants (Berroteran et al., 1994). Interestingly, mutations in the *S. cerevisiae SSU71/TFG1* gene, which encodes the homologue of the RAP74 subunit of human TFIIF, suppress abnormal start site selection in a *sua7ssu71* double mutant (Sun and Hampsey, 1995). By itself the *ssu71* mutant does not affect transcriptional starts, so TFIIF may suppress the *sua7* mutant indirectly through another general factor. The RAP30 subunit of human TFIIF has been shown to interact physically with the N-terminal region of TFIIB (Ha et al., 1993), so the large subunit of TFIIF interacts genetically, and the small subunit interacts physically, with TFIIB.

The largest subunit of RNA polymerase II has an interesting carboxy-terminal domain (CTD) which consists of 52 repeats of the consensus sequence YSPTSPS (Corder and Ingles, 1992; Dahmus, 1994). RNA polymerase II enters the pre-initiation complex most efficiently in the dephosphorylated IIa state. Within the complex the CTD is multiply phosphorylated on the SP serines, by a subunit of TFIIH and possibly other CTD kinases. Elongating RNA polymerase II molecules are primarily in the highly phosphorylated IIo state. So the level of phosphorylation of the CTD may regulate elongation, and a CTD phosphatase may be important for re-cycling dephosphorylated polymerase IIa to a promoter after termination.

A CTD phosphatase has recently been identified that binds to RNA polymerase II (Chambers and Dahmus, 1994; Chambers et al., 1995). This phosphatase interacts with a region of polymerase distinct from the CTD, and will not dephosphorylate the CTD unless this domain is covalently linked to polymerase. The RAP74 subunit of TFIIF stimulates CTD phosphatase activity, apparently through its interaction with RNA polymerase II. Consistent with this view, the C-terminal region of RAP74 that binds polymerase is essential for stimulation. The C-terminal region of RAP74 is masked for phosphatase stimulation (Chambers et al., 1995) just as it is for RNA polymerase II binding (Wang and Burton, 1995) and for TFIIB binding (this paper). TFIIB suppresses

stimulation of phosphatase activity by RAP74, and this is further evidence for functional interaction between TFIIB and RAP74 (Chambers et al., 1995).

Based on these observations, there may be significant coupling of TFIIF and TFIIB function in transcription. TFIIF has functions in both initiation and elongation of RNA chains. TFIIB has known functions in initiation but may have unrecognized functions in elongation as well.

We have mapped the region of RAP30 that interacts with TFIIB (Chapter III). In this report, the region of RAP74 that interacts with TFIIB was also mapped. The RAP30 interaction surface for RAP74 overlaps with the RAP30 interaction surface for TFIIB (Chapter III). RAP74 binding to RAP30 competes with the interaction between RAP30 and TFIIB. Additionally, fragments of RAP74 that interact with TFIIB but do not bind RAP30 antagonize the interaction between TFIIB and RAP30. Interaction between TFIIB, RAP30 and RAP74 may allow for distinct pathways of pre-initiation complex assembly on different promoters and may allow for separate timing of RAP30 and RAP74 function in initiation and elongation.

MATERIALS AND METHODS

Protein reagents

RAP30 mutants (Chapter III) and RAP74 mutants (Wang and Burton, 1995) were prepared as described. A production vector for human TFIIB was the kind gift of D. Reinberg and R. Tjian and this protein was prepared by a protocol supplied by D. Reinberg (Ha et al., 1991).

Protein-protein interaction assays

TFIIB binding to RAP74 mutants. Storage Buffer (SB) was used in binding reactions. SB contains 20 mM Hepes pH 7.9, 20 % glycerol w/v, 1 mM EDTA, 1 mM EGTA, and variable KCl concentration (i.e.. SB 0.1 contains 0.1 M KCl). TFIIB was immobilized on Affi-gel 10 (Bio-Rad) at a density of about 1 mg TFIIB per ml resin. 20 µl TFIIB beads and 300 pmol RAP74 mutant were incubated at 4°C for 1 hour in 0.5 ml SB 0.1 containing 0.2% BSA. Beads were washed with 1 ml SB 0.25, and bound proteins were eluted in 50 µl SB 0.5. 30 µl of this eluate was analyzed on a 15% SDS-PAGE gel developed with silver nitrate. Affi-gel beads without bound protein ligand were used as a negative control.

ELISA assays were done as described with minor modifications (Marsalek and Kaguni, 1994). Microtiter wells (Bectron Dickinson; Pro-Bind) were coated overnight with the protein to be immobilized (5 µg/ml) in 100 µl 50 mM sodium borate pH 9.0, at 4°C. Wells were washed three times with 200 µl PBS containing 0.2% BSA and 0.05% Tween-20 (PBSBT) and blocked with 200 µl PBSBT for an hour at room temperature. Proteins in the mobile phase were added in 50 µl SB 0.1 containing 0.2% BSA and incubated for 15 min at room temperature. Glutaraldehyde was added to a final

concentration of 1 % for 30 min to cross-link protein-protein interactions. Wells were washed three times with 200 µl PBSBT. 100 µl of rabbit antiserum (1:1000 diluted) directed against the mobile phase protein was added to each well and incubated two hours at room temperature. Wells were washed three times with PBSBT, and 100 µl HRP-conjugated goat anti-rabbit secondary antibody (Bio Rad; 1:3000 diluted) was added and incubated for one hour at room temperature. Wells were washed 3 times with PBSBT . Color was developed with 100 µl 50 mM sodium citrate (pH 4.0), 0.03% H_2O_2 and 0.4 mg/ml o-phenylenediamine. The development was stopped by addition of 100 µl 4 N H_2SO_4. The reaction was measured for absorbance at 490 nm using a plate reader (Bio Tek Instruments; EL310). All determinations were done in duplicate and reported as average values.

A Ni^{2+}-affinity bead procedure was used to detect the interactions between TFIIF subunits and TFIIB. Histidine-tagged RAP30 or RAP30(1-176) (300 pmol) was incubated with TFIIB (600 pmol) 1 hr at room temperature in 0.5 ml SB 0.1 containing 0.2% BSA, in the presence or absence of RAP74 (600 pmol). 10 µl Ni^{2+} resin, pre-incubated with SB 0.1 containing 0.2% BSA, was added and incubated with tumbling at 4°C for 1 hr. RAP30 bound the Ni^{2+} resin through their histidine tags. Beads were washed three times with 1 ml SB 0.25. Bound proteins were eluted with 50 µl SDS-PAGE sample buffer. 10 µl of the eluate was electrophoresed by 15% SDS-PAGE and blotted to nitrocellulose. TFIIB or RAP74 was detected in a Western blot developed with anti-TFIIB or anti-RAP74 antiserum.

RESULTS

Both RAP30 and RAP74 bind independently to TFIIB

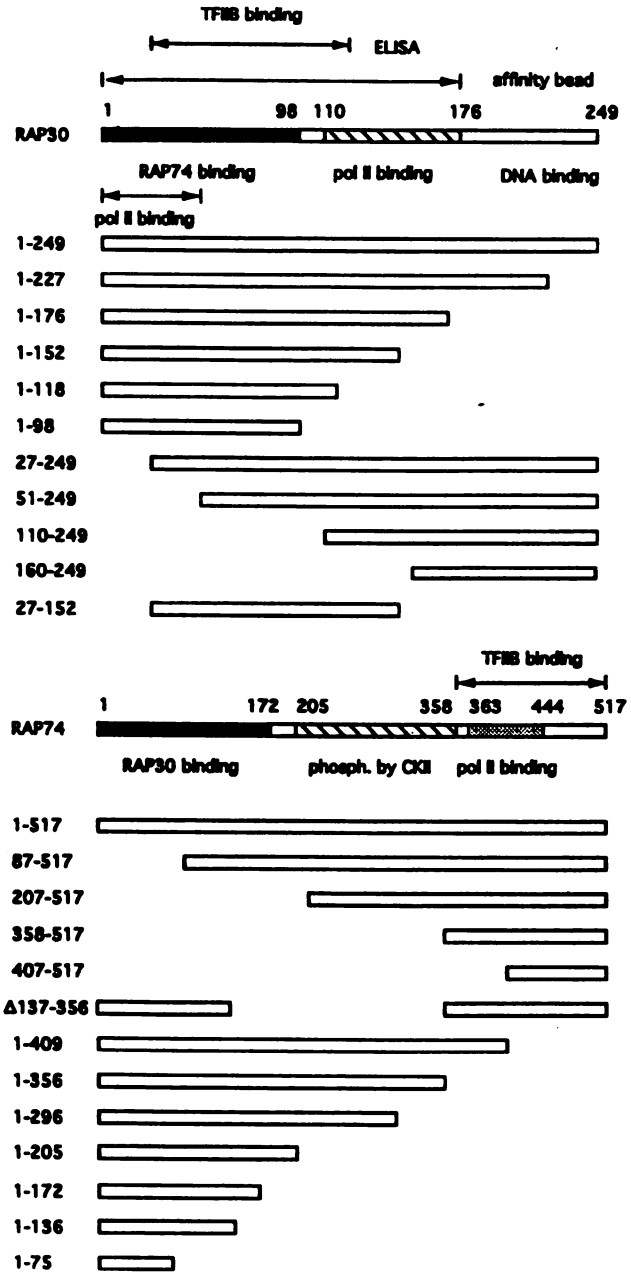
A set of RAP30 and RAP74 deletion mutants was constructed with C-terminal histidine tags to aid in purification and binding reactions (Fig.1; Chapter III; Wang and Burton, 1995). Because of the genetic and physical interaction between TFIIF and TFIIB (Ha et al., 1993; Sun and Hampsey, 1995), RAP30 and RAP74 were tested for direct binding to TFIIB (Fig. 2). TFIIB is shown to bind directly and independently to both RAP30 (lane 2) and RAP74 (lane 4). This result confirms the previous report of a direct interaction between TFIIB and RAP30 (Ha et al., 1993) and demonstrates for the first time that TFIIB also interacts directly with RAP74.

C-terminal region of RAP74 between amino acids 358-517 binds directly to TFIIB

To demonstrate the specificity of the TFIIB-TFIIF interactions and to indicate their functional importance, the sequences of RAP30 and RAP74 that are important for binding TFIIB were mapped. RAP30 deletion mutants have been tested for binding to TFIIB (Chapter III). A set of RAP74 mutants was also tested for binding to immobilized TFIIB (Fig. 3). RAP74(358-517) bound most tightly (lane 4). Full length RAP74, and RAP74(87-517) and (Δ137-356) bound with somewhat lower affinity (lanes 1, 2 and 6). RAP74(207-517) and (407-517) bound much more weakly. RAP74 mutants from which C-terminal sequences were removed bound TFIIB weakly or were not observed to bind TFIIB (Fig. 4). The TFIIB binding site on RAP74, therefore, appears to be located within the C-terminal region between amino acids 358-517.

Sequences within the N-terminal region and central region of RAP74 affect masking of the C-terminal domain for RNA polymerase II binding (Wang and Burton, 1995) and CTD phosphatase stimulation (Chambers et al., 1995). Since RAP74(358-517)

1. RAP30 and RAP74 deletion mutants. RAP30(1-249) is the histidine-tagged version of RAP30. Accurate transcription (tx.) was determined from the Adenovirus major late promoter (Figure 2,Chapter III). RAP74 binding to RAP30 was determined using a Ni^{2+}-affinity bead procedure (Figure 4, Chapter III). TFIIB binding to RAP30 (Figures 6 and 7, Chapter III) and RAP74 (Figures 3 and 4) was determined by affinity bead and ELISA procedures. (+) indicates high activity; (+/-) indicates low activity; (?) indicates barely detectable activity; and (-) indicates no detectable activity; (n.d.) indicates that no determination was made for a particular mutant; (*) indicates data is not shown in this report.



	tx. activity	RAP74 binding	TFIIB binding	
			ELISA	bead
RAP30	100%	+	+	n.d.
1-249	>100%	+	+	+
1-227	30%	+*	+	n.d.
1-176	20%	+	+	+
1-152	(< 5%)	+	+/-	+/-
1-118	0	+	+/-	-
1-98	0	+	-	-
27-249	10%	-	+	-
51-249	0	-	?	-
110-249	0	-	?	-
160-249	n.d.	n.d.	n.d.	-
27-152	0	-	+	n.d.

	TFIIB binding	block IIB-30 binding
RAP74		
1-517	+	+
87-517	+	+
207-517	+/-	+/-
358-517	++	+
407-517	+/-	+/-
Δ137-356	+	n.d.
1-409	?*	+/-
1-356	?*	+/-
1-296	?*	+/-
1-205	+/-*	+/-
1-172	?*	?
1-136	?*	?
1-75	-*	-

Figure 1

Figure 2. Both RAP30 and RAP74 bind independently to TFIIB. 300 pmol RAP30 (30) (lanes 1 and 2) or RAP74 (74) (lanes 3 and 4) was incubated with affinity beads containing covalently immobilized TFIIB (IIB) or no protein ligand (CT for "control"). After washing with SB 0.25, bound protein was eluted and analyzed on an SDS-PAGE gel developed with silver nitrate.

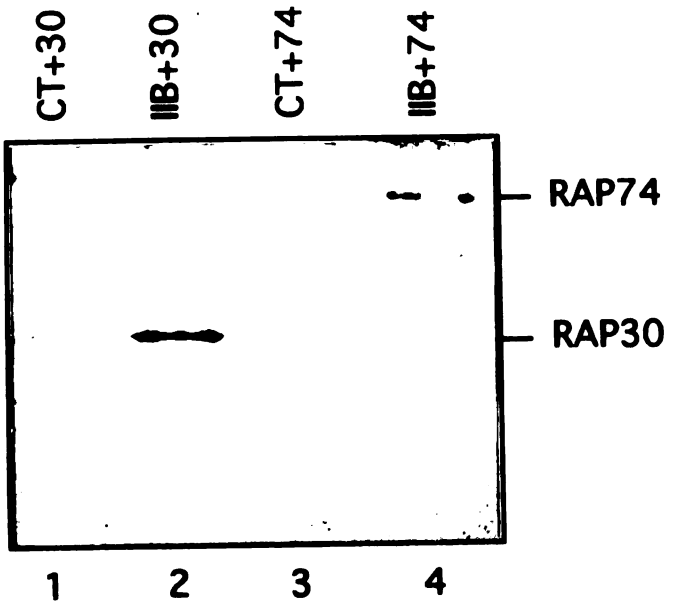


Figure 2

Figure 3. RAP74 has a masked binding site for TFIIB located within a C-terminal region between amino acids 358-517. RAP74 and RAP74 mutants were bound to TFIIB immobilized on agarose beads as described in Figure 2. Bound proteins were analyzed in a polyacrylamide gel developed with silver nitrate. None of these proteins bound to negative control beads (data not shown).

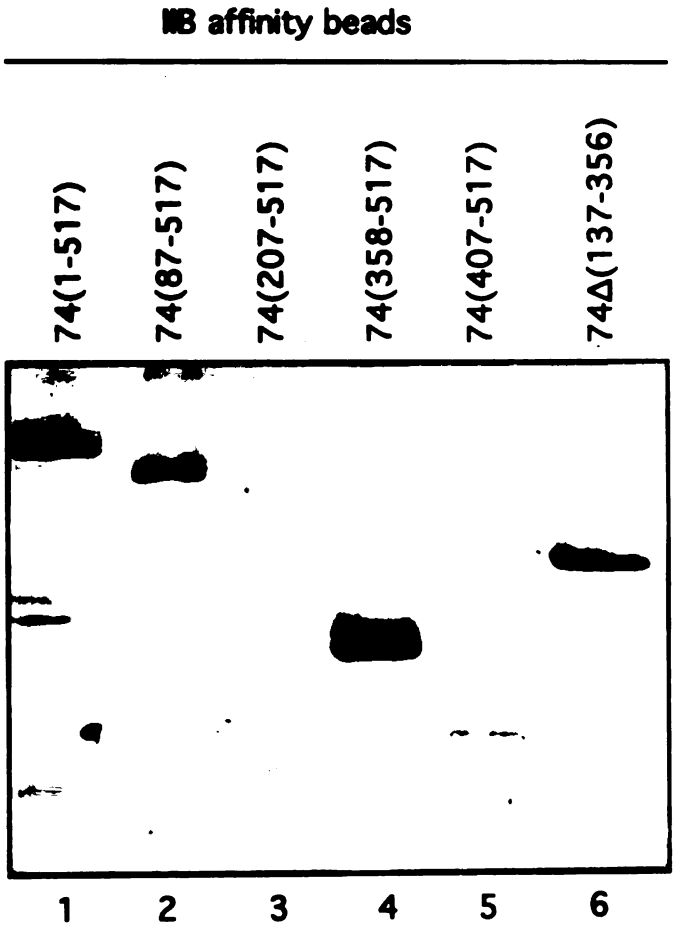

MB affinity beads
74(1-517)
74(87-517)
74(207-517)
74(358-517)
74(407-517)
74Δ(137-356)
1
2
3
4
5
6

Figure 3

Figure 4. C-terminal deletion mutants of RAP74 do not bind TFIIB tightly. RAP74 mutants were bound to TFIIB immobilized on agarose beads as described in Figure 2.

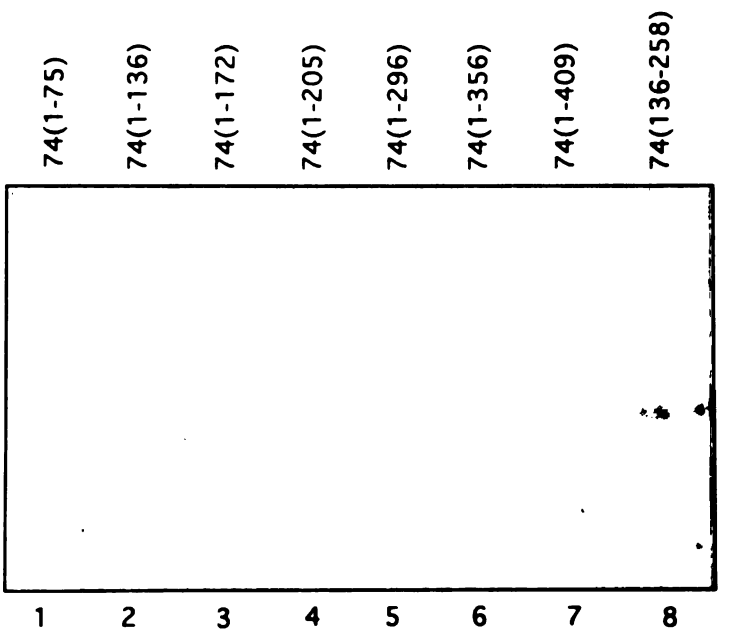

74(1-75)
74(1-136)
74(1-172)
74(1-205)
74(1-296)
74(1-356)
74(1-409)
74(136-258)
1
2
3
4
5
6
7
8

Figure 4

appears to bind more tightly to TFIIB than full-length RAP74, or RAP74(87-517), (207-517), and (Δ137-356), the C-terminal domain also appears to be masked for TFIIB binding. RAP74(207-517) shows the most dramatic masking effect for polymerase binding (Wang and Burton, 1995) and phosphatase stimulation (Chambers et al., 1995) just as it does for TFIIB binding (Fig. 3, lane 3).

RAP74 antagonizes the interaction between TFIIB and RAP30

Since RAP74 and TFIIB interact with overlapping regions of RAP30 (Chapter III), RAP74 might stimulate or antagonize TFIIB binding. In Figure 5, an ELISA test for interactions between TFIIB, RAP30, and RAP74 was done in which one protein was immobilized in the well of the microtiter plate, another was added to the well as a binding partner, to be detected with antibody, and the third was added as a potential competitor or facilitator of the binding interaction. Using this protocol, RAP74 was shown to inhibit TFIIB binding to RAP30 (filled circles). This same conclusion was obtained using a Ni^{2+} affinity bead procedure (Fig. 6A). On the other hand, RAP30 does not dissociate the interaction between RAP74 and TFIIB (x-x). This was perhaps the expected result because the N-terminal region of RAP74 binds to RAP30 (Yonaha et al., 1993; Wang and Burton, 1995) while the C-terminal region of RAP74 binds to TFIIB (Fig. 3-4), so both contacts could be maintained simultaneously. RAP30-RAP74 interactions are also maintained in the presence of TFIIB as a potential competitor (Fig. 5 (open circles) and 6B). Taken together, these data indicate that interactions between RAP30 and RAP74 take precedence over those between RAP30 and TFIIB. On the other hand, the C-terminal domain of RAP74 can bind to TFIIB when RAP74 is complexed with RAP30, in intact TFIIF. So if the RAP30 subunit of TFIIF were to assemble into a transcription complex in the absence of RAP74, RAP30 might be expected to interact with TFIIB. If TFIIF enters as a RAP30/74 complex, however, the C-terminal region of RAP74 is

Figure 5. Dynamic interactions between TFIIF subunits and TFIIB. RAP74 blocks formation of a RAP30-TFIIB complex (filled circles). TFIIB does not block formation of a RAP30/74 (TFIIF) complex (open circles). RAP30 does not block formation of a RAP74-TFIIB complex (x-x). TFIIB or RAP74 (5 μg/ml) was immobilized in the wells of a microtiter plate. A mixture of a mobile protein (3.5 pmol) and increasing quantities of a potential binding competitor (or facilitator) were added. Detection of bound protein was with anti-RAP30 or anti-RAP74 antiserum.

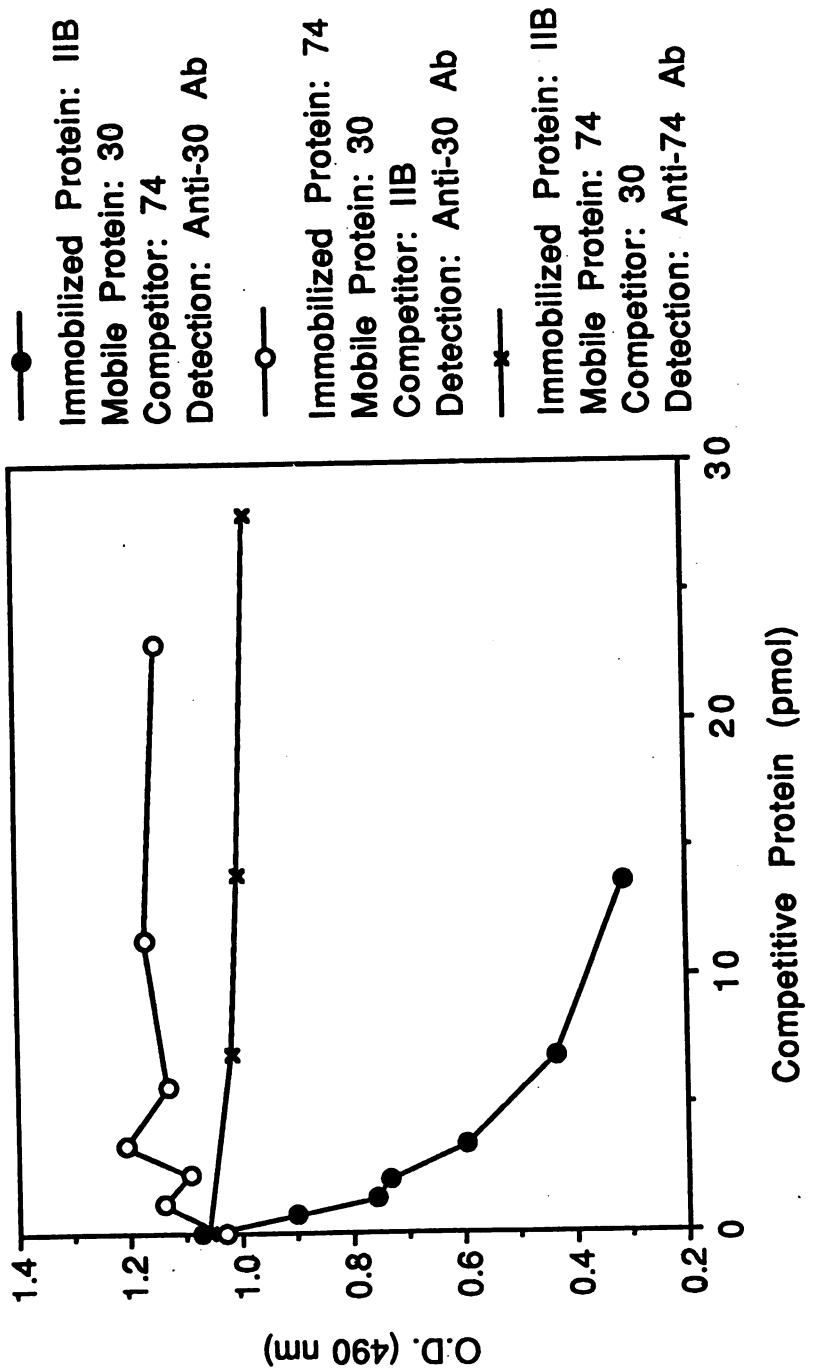
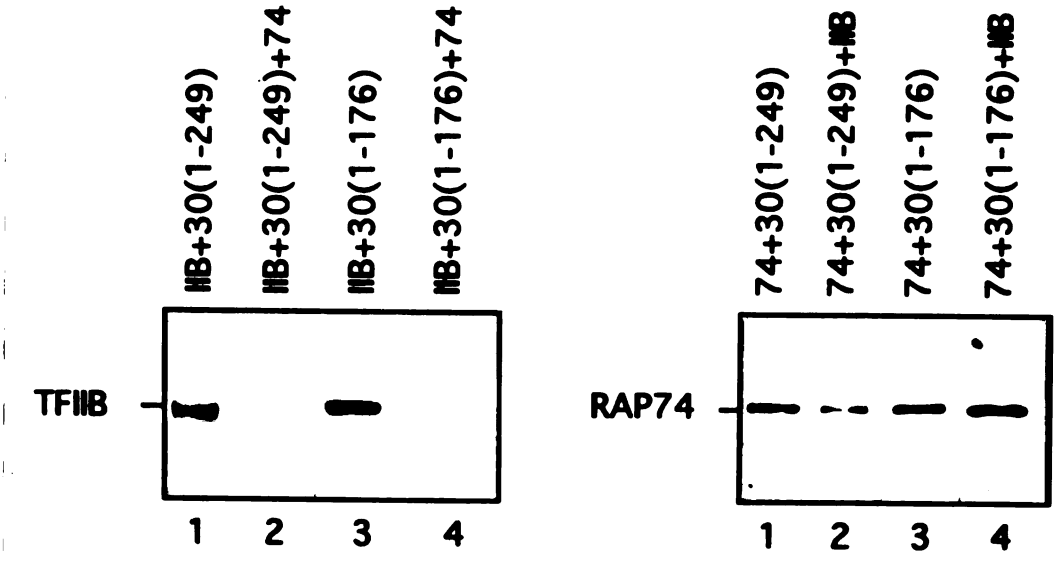

Immobilized Protein: IIB
Mobile Protein: 30
Competitor: 74
Detection: Anti-30 Ab
Immobilized Protein: 74
Mobile Protein: 30
Competitor: IIB
Detection: Anti-30 Ab
Immobilized Protein: IIB
Mobile Protein: 74
Competitor: 30
Detection: Anti-74 Ab
O.D. (490 nm)
Competitive Protein (pmol)
0.2
0.4
0.6
0.8
1.0
1.2
1.4
0
10
20
30

Figure 5

Figure 6. A) RAP74 blocks formation of a RAP30-TFIIB complex. A Ni^{2+}-affinity bead procedure was used, as described in Materials and Methods. Histidine-tagged RAP30 or RAP30(1-176) (300 pmol) was bound to TFIIB (600 pmol) in the presence or absence of RAP74 (600 pmol). In the presence of RAP74, TFIIB did not bind to the Ni^{2+} resin. A Western blot is shown developed with anti-TFIIB antiserum. **B) TFIIB does not block formation of a RAP30/74 (TFIIF) complex.** Histidine-tagged RAP30 or RAP30(1-176) retained RAP74 in both the presence and absence of TFIIB. A Western blot is shown developed with anti-RAP74 antiserum.


IIB+30(1-249)
IIB+30(1-249)+74
IIB+30(1-176)
IIB+30(1-176)+74
TFIIB
1
2
3
4
74+30(1-249)
74+30(1-249)+IIB
74+30(1-176)
74+30(1-176)+IIB
RAP74
1
2
3
4

A. B.

Figure 6

expected to interact with TFIIB, and interaction between RAP30 and TFIIB is expected to be blocked.

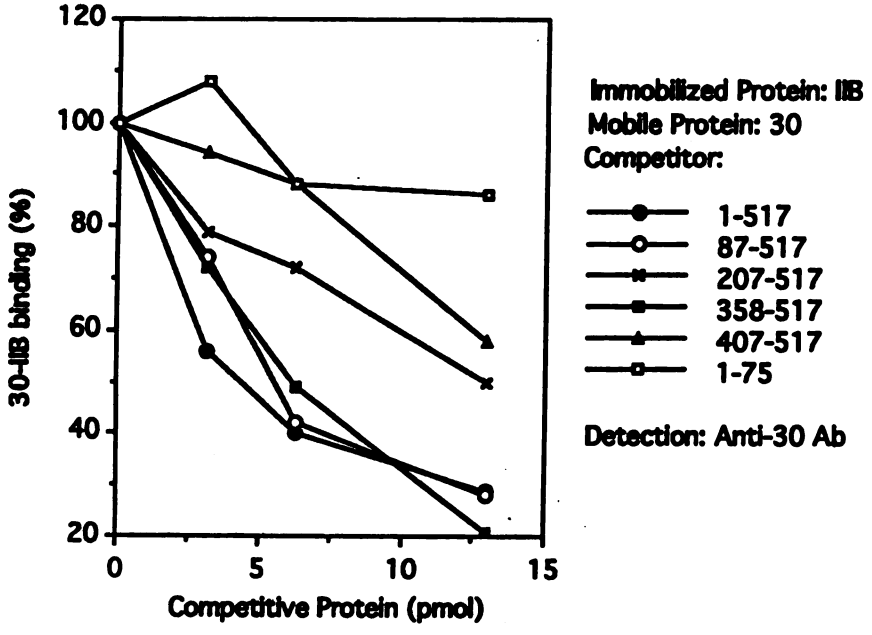
RAP74 disrupts the interaction between TFIIB and RAP30 by two mechanisms

RAP74 could potentially disrupt the interaction between TFIIB and RAP30 by two mechanisms, binding to RAP30 and/or binding to TFIIB. In Figure 7, both mechanisms are shown to contribute to blocking the TFIIB-RAP30 interaction. RAP74 mutants containing a RAP30 binding region disrupt TFIIB-RAP30 binding (left panel), and RAP74 mutants that are missing the RAP30-binding region (amino acids 1-172) but contain the TFIIB-binding region (amino acids 358-517) also disrupt TFIIB-RAP30 binding (right panel). There is a close correlation between the RAP74 mutants that bind TFIIB most strongly and the ability of these mutants to compete the TFIIB-RAP30 interaction (Figs. 3 and 7 (right panel)).

RAP74-TFIIB interaction is maintained in complexes containing polymerase

Since the RAP74 interaction appears to take precedence over the RAP30 interaction with TFIIB, one question that arises is whether the RAP74-TFIIB interaction is maintained in complexes containing RNA polymerase II, for instance within a fully assembled pre-initiation complex. Although this question is difficult to address directly with a simple experiment, we have tested the effect of RNA polymerase II on binding of TFIIB and RAP74 (Fig. 8). Since both TFIIB and RNA polymerase II bind to overlapping regions within the C-terminal domain of RAP74 (amino acids 358-517 vs. 363-444), these proteins might compete or cooperate for binding. RNA polymerase II does not appear to strongly compete or facilitate the TFIIB-RAP74 interaction. When RAP74 was immobilized and TFIIB was added as the binding partner, addition of RNA polymerase II was not observed to affect formation of complexes containing TFIIB and RAP74 (filled circles). When TFIIB was immobilized and RAP74 added as the binding

Figure 7. RAP74 blocks formation of the RAP30-TFIIB complex both by binding to RAP30 and by binding to TFIIB. RAP74 or RAP74 mutants were tested in the ELISA competition assay Figure 5. The region of RAP74 that binds RAP30 maps between amino acids 1-172 (Wang and Burton, 1995). The TFIIB binding region maps between 358-517 (Fig. 3).


120
100
80
60
40
20
0
5
10
15
30-IIB binding (%)
Competitive Protein (pmol)
Immobilized Protein: IIB
Mobile Protein: 30
Competitor:
1-517
87-517
207-517
358-517
407-517
1-75
Detection: Anti-30 Ab

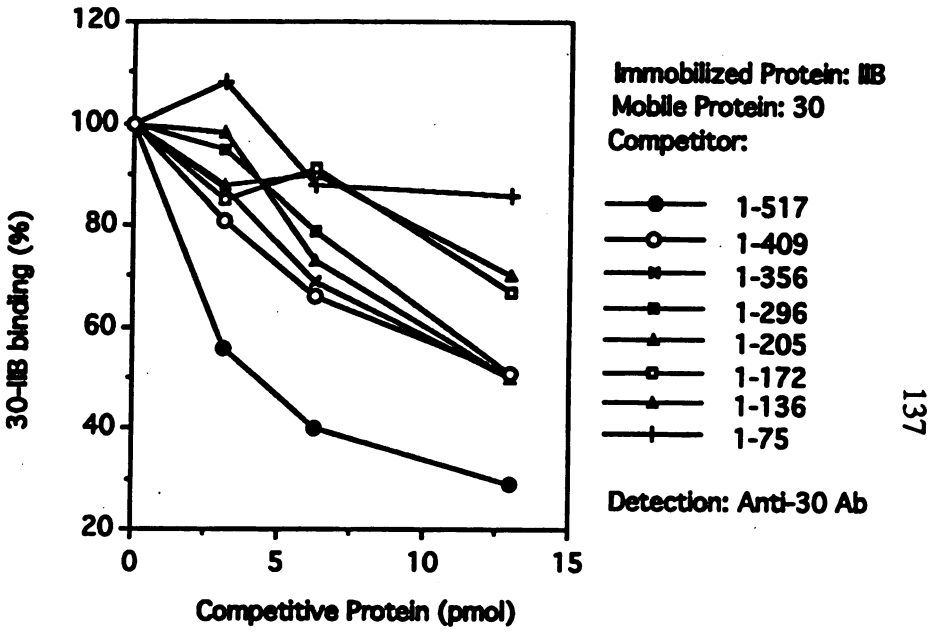

120
100
80
60
40
20
0
5
10
15
30-IIB binding (%)
Competitive Protein (pmol)
Immobilized Protein: IIB
Mobile Protein: 30
Competitor:
1-517
1-409
1-356
1-296
1-205
1-172
1-136
1-75
Detection: Anti-30 Ab

Figure 7

Figure 8. RNA polymerase II does not appear to block TFIIB-RAP74 interactions. This experiment was done essentially as described above for Figure 5 and as indicated in the figure key.

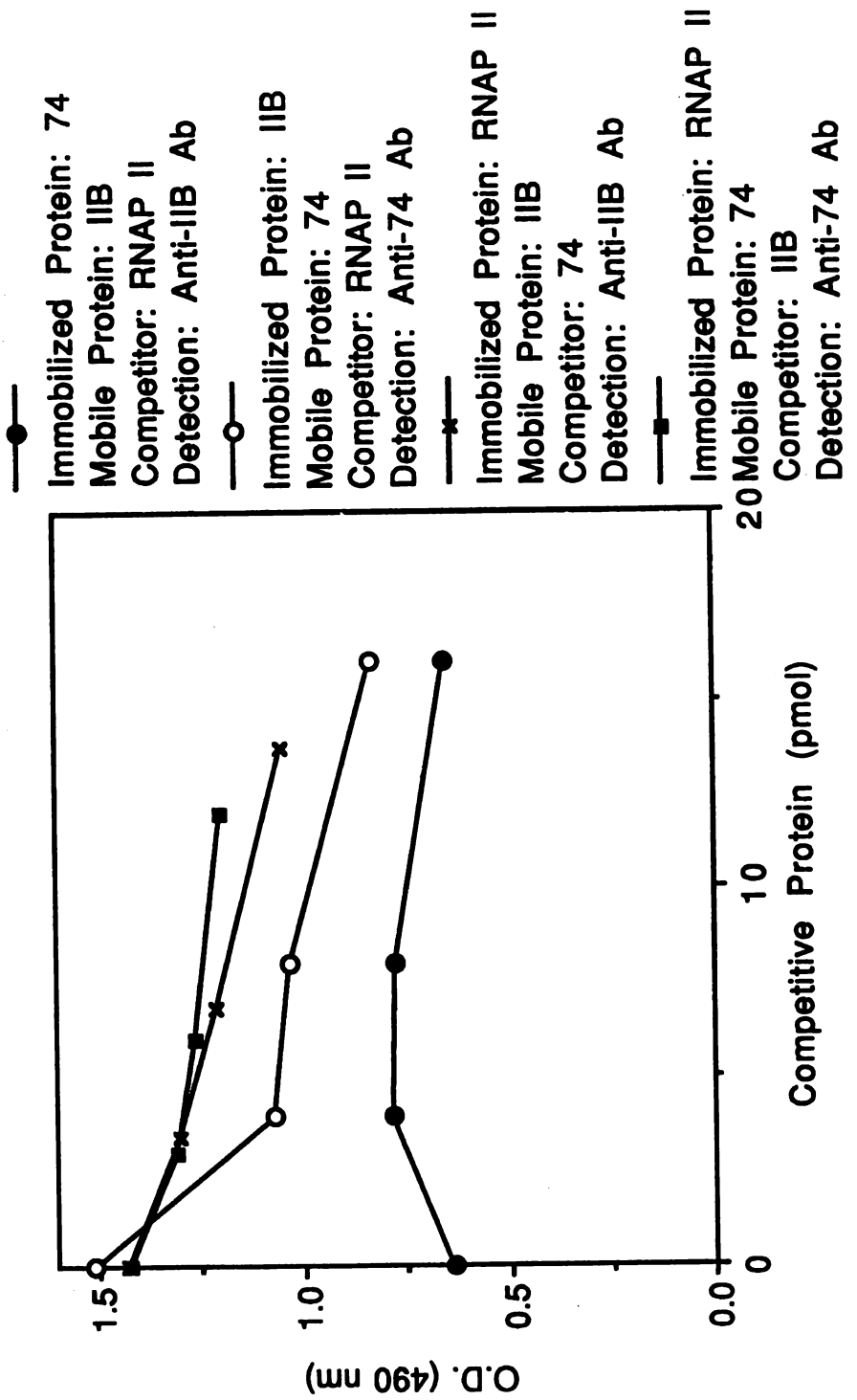

Immobilized Protein: 74
Mobile Protein: IIB
Competitor: RNAP II
Detection: Anti-IIB Ab
Immobilized Protein: IIB
Mobile Protein: 74
Competitor: RNAP II
Detection: Anti-74 Ab
Immobilized Protein: RNAP II
Mobile Protein: IIB
Competitor: 74
Detection: Anti-IIB Ab
Immobilized Protein: RNAP II
Mobile Protein: 74
Competitor: IIB
Detection: Anti-74 Ab
O.D. (490 nm)
0.0
0.5
1.0
1.5
0
10
20
Competitive Protein (pmol)

Figure 8

partner, addition of RNA polymerase II caused a moderate reduction in the retention signal (open circles). Of course, RNA polymerase II is expected to interact with both TFIIB and RAP74, and polymerase may interfere with binding between anti-RAP74 antibody and RAP74. The slight apparent inhibition of binding that is observed, in the experiment in which TFIIB was immobilized (open circles), could be attributable to antibody interference or distribution of TFIIB-polymerase and RAP74-polymerase complexes. It may be difficult to exchange TFIIB and/or RAP74 bound to separate polymerase molecules into a ternary protein complex, since this requires dissociation of one factor from polymerase. Consistent with this idea, when RNA polymerase II was immobilized and competition for binding was between RAP74 and TFIIB, a smaller reduction in the retention signal was observed. This was true whether detection was for TFIIB (x-x) or RAP74 (filled squares). Of course, when RNA polymerase II is immobilized, TFIIB and RAP74 could bind independently without maintaining a TFIIB-RAP74 interaction. Additional experiments will be required to demonstrate TFIIB-RAP74 contacts within the pre-initiation complex.

DISCUSSION

There is mounting evidence for functional interaction between TFIIF and TFIIB. These factors cooperate to bring RNA polymerase II into the pre-initiation complex (Conaway et al., 1991; Zawel and Reinberg, 1993, 1995; Serizawa et al., 1994). TFIIF, TFIIB, and RNA polymerase II cooperate to select transcriptional start sites (Pinto et al., 1992; Li et al., 1994; Berroteran et al., 1994; Sun and Hampsey, 1995). These factors may also cooperate to control dephosphorylation of the CTD (Chambers et al., 1995).

Both the RAP30 and RAP74 subunits of TFIIF contribute to assembly of the pre-initiation complex. The RAP30 subunit is sufficient to bring RNA polymerase II to the complex (Flores et al., 1991), but the RAP74 subunit helps to stabilize this interaction (Tan et al., 1994) and to position RAP30 on template DNA (Coulombe et al., 1994). In some in vitro transcription systems, both the RAP30 and the RAP74 subunit are required for initiation of transcription (Tan et al., 1994), but in an extract system from which TFIIF was removed by immunoprecipitation, only addition of RAP30 was required to reconstitute initiation (Chang et al., 1993). RAP74 was required in this system, however, for early elongation of the transcript. It is possible that a trace of RAP74 remaining after immunodepletion was sufficient to support initiation but not elongation, but no evidence for differential concentration requirements for RAP74 in initiation and early elongation has yet been presented, and such a model requires that RAP74 cycle from the complex after initiation and then reassociate with the complex to support elongation. In any event, our studies showed that in an extract system RAP74 is required for polymerase to escape the promoter. Both RAP30 and RAP74 are required for TFIIF to stimulate elongation strongly using dC-tailed template assays (Tan et al., 1994; Kephart et al., 1994), so we assume both subunits participate in promoter escape. In one report, weak stimulation of elongation was seen with RAP74 alone (Kephart et al., 1994). Both RAP30 and RAP74

can bind to RNA polymerase II (McCracken and Greenblatt, 1991; Wang and Burton, 1995; Chapter III) and TFIIB (Chapter III; Fig. 2) independently, so mechanisms can be imagined in which TFIIF subunits might enter transcription complexes separately or, in some cases, function independently.

RAP30 appears to consist of three functional regions. The N-terminal region binds RAP74 and RNA polymerase II (Yonaha et al., 1993; Tan et al., 1995; Chapter III), the central region binds RNA polymerase II (McCracken and Greenblatt, 1991; Chapter III), and the C-terminal region binds DNA (Tan et al., 1994b). The N-terminus of RAP30 may be involved in regulating the accessibilities of central and C-terminal regions (Chapter III). RAP74 appears to be divided into three functional domains, as indicated by sequence analysis (Finkelstein et al., 1992; Aso et al., 1992) and deletion mutants (Wang and Burton, 1995). The RAP74 primary sequence can be divided into an N-terminal basic region, a highly charged central region with overall negative charge, and a C-terminal basic region. The N-terminal region binds to RAP30 (Yonaha et al., 1993; Wang and Burton, 1995). The central region appears to be a largely unstructured hinge that controls accessibility of the C-terminal region. The C-terminal region binds directly to RNA polymerase II (Wang and Burton, 1995) and is important for stimulation of a CTD phosphatase (Chamber et al., 1995). Deletion from the N-terminus of RAP74 initially increases masking of the polymerase binding activity (Wang and Burton, 1995) and eliminates stimulation of CTD phosphatase activity (Chambers et al., 1995). Further deletion within the central region unmasks both polymerase binding and CTD phosphatase activities. The central region is also the site of phosphorylation by casein kinase II, and phosphorylation by this and/or other kinases appears to stimulate polymerase binding by TFIIF (Kitajima et al., 1994).

Multiple contacts within the RAP74-binding and RNA polymerase II-binding regions of RAP30 contribute to TFIIB binding (Chapter III). Addition of RAP74 blocks interaction between TFIIB and RAP30 through two mechanisms. By binding the N-

terminal region of RAP30, the N-terminal region of RAP74 can block interaction between RAP30 and TFIIB. Also, by binding TFIIB, the C-terminal region of RAP74 can block this interaction. Most likely, the C-terminal region of RAP74 interacts with the N-terminal region of TFIIB to compete with RAP30 binding. The structure of the TATA box/TBP/TFIIB complex strongly indicates that interaction between TFIIF and TFIIB might be through the N-terminal region of TFIIB (Nikolov et al., 1995). Since RAP74 dissociates RAP30-TFIIB interactions, TFIIF/TFIIB contact is most likely maintained through the C-terminal region of RAP74.

The strongest contacts between TFIIB and RAP74 are within the C-terminal region of RAP74 between amino acids 358-517 (Fig. 3). The C-terminal region of RAP74 shows a very similar pattern of masking by N-terminal and central region sequences for TFIIB binding (Fig. 3), RNA polymerase II binding (Wang and Burton, 1995), and CTD phosphatase stimulation (Chambers et al., 1995). RAP74 can bind to RAP30 and TFIIB simultaneously (Fig. 5). This result is consistent with the mapping of RAP74 functional domains, since the N-terminal region binds RAP30 and the C-terminal region binds TFIIB. Although RNA polymerase II and TFIIB bind to overlapping regions within the C-terminal domain of RAP74, RNA polymerase II does not appear to strongly compete for RAP74-TFIIB binding (Fig. 8). That TFIIB and RAP74 might interact in complexes containing RNA polymerase II is also indicated by the genetic interaction between yeast *SUA7* and *SSU71* (Sun and Hampsey, 1995) and by TFIIB-mediated suppression of RAP74 stimulation of CTD phosphatase activity (Chamebers et al., 1995). Interaction between the C-terminal region of human RAP74 and TFIIB, therefore, may be maintained in complexes with RNA polymerase II.

The *SUA7* gene in yeast encodes TFIIB, and the RAP74 homologue in yeast is encoded by *SSU71/TFG1*. *sua7* mutants that confer cold sensitivity and altered transcriptional start site selection, and are suppressed by *ssu71*, map to a particular region of TFIIB near the N-terminus (Pinto et al., 1994). From our work, we would have

predicted that the *ssu71* mutations that suppress *sua7* might map within the C-terminal region of *SSU71* where human TFIIB and RAP74 interact; however, the *ssu71* G363D and G363R alleles (M. Hampsey, personal communication) correspond to the K111 position of human RAP74 (Wang and Burton, 1995), within the RAP30 binding region. So genetics implicate both TFIIF subunits in start site selection. In *S. cerevisiae*, transcriptional start sites can be more than 100 bases downstream from a TATA box. RNA polymerase II appears to "melt in" near the TATA box and scan single-stranded DNA for the +1 position (Giardina and Lis, 1993). Since TFIIB and TFIIF are part of the start site selection mechanism, both factors may traverse a significant length of DNA from the TATA box to the initiation site.

Although the experiments shown here demonstrate specific protein-protein contacts between general transcription factors TFIIB and TFIIF, the importance of RAP74 preventing interaction between RAP30 and TFIIB is not yet clear. Based on our data, TFIIF is expected to interact with TFIIB through the RAP74 subunit, but in the absence of RAP74, a RAP30-TFIIB complex could be maintained. Such alternate structures are most easily understood if TFIIF does not invariably function as an intact unit, but rather, under some circumstances, as two independent factors, RAP30 and RAP74. In Figure 9, models are presented to describe alternate pathways for initiation and productive elongation based on dynamic interaction between TFIIB and TFIIF subunits. These models account for the possibility that RAP30 and RAP74 may have partially separable functions in initiation and early elongation (Flores et al., 1991; Killeen et al., 1992; Chang et al., 1993). Also a model is presented that may explain how dynamic interaction between TFIIB and TFIIF subunits might influence elongation.

According to one scheme (Fig. 9A), TFIIF enters the pre-initiation complex as a unit and interacts with TFIIB through the C-terminal domain of RAP74. This mechanism has been extensively demonstrated using in vitro systems, although alternate pathways to complex assembly are also possible. In this case, RAP74 is present during pre-initiation

Figure 9. models for dynamic TFIIF-TFIIB interactions during initiation, promoter escape, and elongation. A) The most familiar assembly and initiation mechanism. In this scheme, TFIIF enters the pre-initiation complex as an intact factor. Initiation and promoter escape proceed. Contact between TFIIF and TFIIB, in this case, is through the RAP74 subunit. B) on some promoters transcription may be regulated at the level of promoter escape rather than initiation. Initiation can proceed in the absence of RAP74 (Chang et al., 1993). In this case, RAP30 binds TFIIB. Since RAP74 prevents release of newly-initiated transcripts (C.-h. Chang and Z.F. Burton, submitted), regulatory factors must stabilize the initiated complex. Promoter escape requires the RAP74 subunit (Chang et al., 1993), so by timing the entry of RAP74 into the complex, productive initiation is regulated at the level of promoter escape. C) If TFIIB has a role in elongation, dynamic interactions between TFIIB and TFIIF subunits may be important to regulate this stage of the transcription cycle. Since intact TFIIF stimulates elongation rates (Price et al., 1989; Bengal et al., 1991; Izban and Luse, 1992), regulating RAP30-RAP74 subunit contacts will stimulate or inhibit elongation.

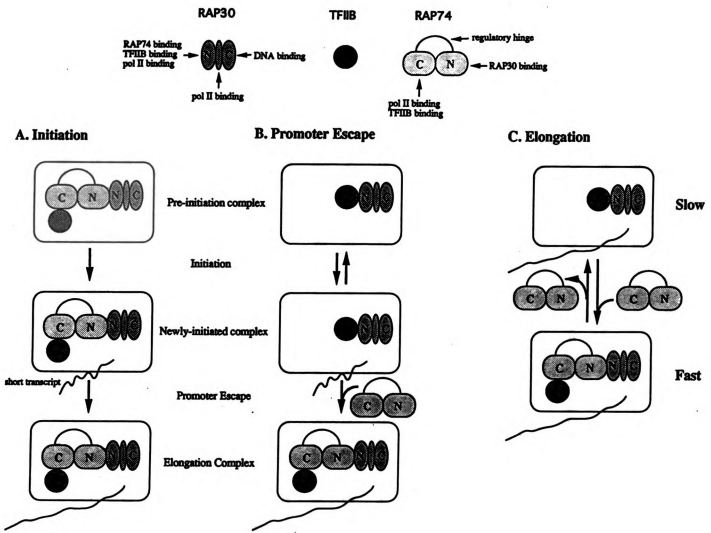

RAP30
TFIIB
RAP74
RAP74 binding
TFIIB binding
pol II binding
DNA binding
pol II binding
regulatory hinge
RAP30 binding
pol II binding
TFIIB binding
A. Initiation
B. Promoter Escape
C. Elongation
Pre-initiation complex
Initiation
Newly-initiated complex
short transcript
Promoter Escape
Elongation Complex
Slow
Fast

Figure 9

complex formation to stabilize the interaction between RAP30 and template DNA (Tan et al., 1994; Coulombe et al., 1994). This function of RAP74 can be complemented by a factor or factors present in extract transcription systems but missing from highly purified systems (Chang et al., 1993). In this scheme, RAP74 is present within the newly-initiated complex to protect the short transcript from release and to drive RNA polymerase II into a productive elongation mode (C.-h. Chang and Z.F. Burton, submitted). Therefore, transcription is regulated at the level of initiation.

Transcription from some promoters, however, may be controlled by regulating promoter escape (Fig. 9B). RNA polymerase II cannot exit the promoter unless RAP74 enters the complex (Chang et al., 1993). The function of RAP74 in pre-initiation complex assembly that has been demonstrated in several defined in vitro systems (Tan et al., 1994; Coulombe et al., 1994) can be replaced by factors present in cell extracts (Chang et al., 1993). In the absence of RAP74, TFIIB binds the N-terminal region of RAP30. Accurate transcription initiation from linear templates has been demonstrated in the absence of added RAP74 in extract systems but not in highly purified systems. The newly-initiated transcript, in this case, requires stabilization. RAP74 can provide this stabilization (C.-h. Chang and Z.F. Burton, submitted) as can other factors present in extract systems (Chang et al., 1993). Productive transcription from such a promoter can be regulated by controlling the stability of the newly-initiated complex and by timing the entry of RAP74 into the complex to allow promoter escape by RNA polymerase II. Other protein factors can affect these mechanisms. For instance, serum response factor binds to the serum response element upstream of the human *c-fos* promoter. This factor regulates transcription through interactions with RAP74 (Zhu et al., 1994; Joliot et al., 1995), and may therefore function by stabilizing the newly-initiated complex and facilitating promoter escape.

TFIIB is known to be an initiation factor, but may have unrecognized functions in elongation as well. Our laboratory became interested in this issue when we isolated RNA

polymerase II-binding transcription factors from yeast by anti-CTD affinity chromatography (Wade et al., 1995). Surprisingly, TFIIB and TFIIF were of similar abundance in complexes with RNA polymerase II. Since anti-CTD chromatography should yield a mixture of initiation and elongation forms, we expected a much lower yield of TFIIB relative to TFIIF, if TFIIB were only involved in initiation. As discussed above, the observation of Chambers et al. that TFIIB suppresses RAP74-mediated stimulation of a CTD-phosphatase also links TFIIB function either to transcription elongation or to RNA polymerase II recycling after termination (Chambers et al., 1995).

Zawel et al. have demonstrated that in a purified transcription system, TFIID remains bound at the promoter, and TFIIB, TFIIF, TFIIE, and TFIIH dissociate after initiation (Zawel et al., 1995). TFIIF then re-associates with elongating polymerase but TFIIB reassociates with TBP at the promoter. These experiments, however, were done in a system that is missing many factors that might contribute to the stability of elongation complexes. In cells and extract systems, therefore, TFIIB and TFIIF may not dissociate and may remain bound to polymerase during elongation. If this is the case, dynamic TFIIB-TFIIF interactions may modulate elongation rates, contribute to pausing, and/or control termination and polymerase re-cycling (Fig. 9C).

Acknowledgements

We thank M. Hampsey for helpful comments on this manuscript and for permission to mention *ssu71* G363D and G363R alleles. We thank D. Reinberg and R. Tjian for a TFIIB production clone and for instructions for TFIIB purification.

This work was initially supported by a grant from the National Institutes of Health. This work was also supported by the Michigan State University Agricultural Experiment Station, by Michigan State University, and by Ms. Verna C. Finkelstein.

REFERENCES

Aso, T., Vasavada, H. A., Kawaguchi, T., Germino, F. J., Gangelly, S., Kitajima, S., Weissman, S. M. and Yasukochi, Y. (1992). Nature 355, 461-464.

Bagby, S., Kim, S., Maldonado, E., Tong, K.I., Reinberg, D., and Ikura, M. (1995) Cell 82, 857-867.

Bengal, E., Flores, O., Krauskopf, A., Reinberg, D. and Aloni, Y. (1991) Mol. Cell. Biol. 11, 1195-1206.

Berroteran, R.W., Ware, D.E., and Hampsey, M. (1994) Molec. and Cell. Biol. 14, 226-237

Buratowski, S., and Zhou, H. (1993) Proc. Natl. Acad. Sci. USA 90, 5633-5637.

Buratowski, S., Hahn, S., Guarente, L., and Sharp, P.A. (1989) Cell 56, 549-561.

Chambers, R.S., and Dahmus, M.E. (1994) J. Biol. Chem. 269, 26243-26248.

Chambers, R. S., Wang, B. Q., Burton, Z. F. and Dahmus, M. E. (1995) J.Biol.Chem. 270, 14962-14969.

Chang, C., Kostrub, C. F. and Burton, Z. F. (1993). J. Biol. Chem. 268, 20482-20489.

Coulombe, B., Li, J., and Greenblatt, J. (1994) J. Biol. Chem. 269, 19962-19967.

Conaway, J.W. and Conaway, R. C. (1989) J. Biol. Chem. 264, 2357-2362.

Conaway, R.C., Garrett, K.P., Hanley, J.P., and Conaway, J.W. (1991) Proc. Natl. Acad. Sci. USA 88, 6205-6209.

Corden, J.L., and Ingles, C.J. (1992) in Transcriptional Regulation, McKnight, S.L., and Yamamoto, K.R., eds., Cold Spring Harbor Laboratory Press, N.Y., pp. 81-107.

Dahmus, M.E. (1994) in Transcription: Mechanism and Regulation, Conaway, R.C., and Conaway, J.W., eds., Raven Press, N.Y., pp. 243-262.

Finkelstein, A., Kostrub, C. F., Li, J., Chavez, D. P., Wang, B. Q., Fang, S. M., Greenblatt, J. and Burton, Z. F. (1992) Nature 355, 464-467.

Flores, O., Maldonado, E., Burton, Z.F., Greenblatt, J., and Reinberg, D. (1988) J. Biol. Chem. 263, 10812-10816.

Flores, O., Ha, I., and Reinberg, D. (1990) J. Biol. Chem. 265, 5629-5634.

Flores, O., Lu, H., Killeen, M., Greenblatt, J., Burton, Z.F., and Reinberg, D. (1991) Proc. Natl. Acad. Sci. USA 88, 9999-10003.

Garrett, K. P., Serizawa, H., Hanley, J. P., Bradsher, J. N., Tsuboi, A., Arai N., Yokota, T., Arai, K., Conaway, R. C. and Conaway, J. W. (1992). J. Biol. Chem. 267, 23942-23949.

Giardina, C., and Lis, J.T. (1993) Science 261, 759-762.

Ha, I., Lane, W.S., and Reinberg, D. (1991) Nature (Lond.) 352, 689-695.

Ha, I., Roberts, S., Maldonado, E., Sun, X., Kim, L. U., Green, M. and Reinberg, D. (1993). Genes Dev. 7, 1021-1032.

Hisatake, K., Roeder, R.G., and Horikoshi, M. (1993) Nature (Lond.) 363, 744-747.

Izban, M.G. and Luse, D. S. (1992) J. Biol. Chem. 267, 13647-13655.

Joliot, V., Demma, M. and Preywes, R. (1995) Nature 373, 632-635.

Kephart, D.D., Wang, B.Q., Burton, Z.F., and Price, D.H. (1994) J. Biol. Chem. 269, 13536-13543.

Killeen, M., Coulombe, B., and Greenblatt, J. (1992) J. Biol. Chem. 267, 9463-9466.

Kitajima, S., Tanaka, Y., Kawaguchi, T., Nagaoka, T., Weissman, S. and Yasukochi, Y. (1990) Nucleic Acids Res. 18, 4843-4849.

Kitajima, S., Chibazakura, T., Yonaha, M., and Yasukochi, Y. (1994) J. Biol. Chem. 269, 29970-29977.

Li, Y., Flanagan, P.M., Tschochner, H., and Kornberg, R.D. (1994) Science 263, 805-807.

Luthy, R., and Eisenberg, D. (1990) in Sequence Analysis Primer, Gribskov, M., and Devereux, J., eds., Stockton Press, N.Y., pp. 61-87.

Maldonado, E., Ha, I., Cortes, P., Weis, L., and Reinberg, D. (1990) Mol. Cell. Biol. 10, 6335-6347.

Malik, S., Hisatake, K., Sumimoto, H., Horikoshi, M., and Roeder, R.G. (1991) Proc. Natl. Acad. Sci. USA 88, 9553-9557.

Malik, S., Lee, D.K., and Roeder, R.G. (1993) Mol. Cell. Biol. 13, 6253-6259.

Marsalek, J., and Kaguni, J.M. (1994) J. Biol. Chem. 269, 4883-4890

McCracken, S. and Greenblatt, J. (1991) Science 253, 900-902.

Nikolov, D.B., Chen, H., Halay, E.D., Usheva, A.A., Hisatake, K., Lee, D.K., Roeder, R.G., and Burley, S.K. (1995) Nature (Lond.) 377, 119-128

Parvin, J. D. & Sharp, P. A. (1993) Cell 73, 533-540.

Pinto, I., Ware, D. E., and Hampsey, M. (1992) Cell 68, 977-988.

Pinto, I., Wu, W.-H., Na, J.G., and Hampsey, M. (1994) J. Biol. Chem. 269, 30569-30573.

Price, D.H., Sluder, A.E., and Greenleaf, A.L. (1989) Mol. Cell. Biol. 9, 1465-1475.

Roberts, S.G.E., and Green, M.R. (1994) Nature (Lond.) 371, 717-720.

Roberts, S.G.E., Ha, I., Maldonado, E., Reinberg, D., and Green, M.R. (1993) Nature (Lond.) 363, 741-744.

Serizawa, H., Conaway, J.W., and Conaway, R.C. (1994) Transcription: Mechanisms and Regulation, Conaway, R.C., and Conaway, J.W. eds., Raven Press, N.Y., pp. 27-43.

Sopta, M., Burton, Z. F. and Greenblatt, J. (1989) Nature 341, 410-414.

Sun, Z. and Hampsey, M. (1995) Proc. Natl. Acad. Sci. USA 92, 3127-3131.

Tan, S., Aso, T., Conaway, R.C., and Conaway, J.W. (1994) J. Biol. Chem. 269, 25684-25691.

Tan, S., Garrett, K. P., Conaway, R. C. and Conaway, J. W. (1994b). Proc. Natl. Acad. Sci. USA 91, 9808-9812.

Tan, S., Conaway, R.C., and Conaway, J.W. (1995) Proc. Natl. Acad. Sci. USA 92, 6042-6046.

Usheva, A. & Shenk, T. (1994) Cell 76, 1115-1121.

Wade, P.A., Werel, W., Chang, C.-h., Shi, X., Wolf, A.J., Sherman, J., Fentzke, R.C., Thompson, N.E., Leykam, J.F., Reed, S.I., Burgess, R.R., Jaehning, J.A., and Burton, Z.F. (1995) Molec. Cell. Biol., in press.

Wang, B.Q. and Burton, Z.F. (1995) J. Biol. Chem. 270, 27035-27044.

Wang, B. Q., Kostrub, C. F., Finkelstein, A. and Burton, Z. F. (1993) Protein Expr. Purif. 4, 207-214.

Yonaha, M., Aso, T., Kobayashi, Y., Vasavada, H., Yasukochi, Y., Weissman, S. M. and Kitajima, S. (1993). Nucleic Acids Res. 21, 273-279.

Zawel, L. and Reinberg, D. (1993) Prog. Nucl. Acids Res. 44, 67-108.

Zawel, L. and Reinberg, D. (1995) Annu. Rev. Biochem. 64, 533-561.

Zawel, L., Kumar, K.P., and Reinberg, D. (1995) Genes Develop. 9, 1479-1490.

Zhu, H., Joliot, V. and Prywes, R. (1994) J. Biol. Chem. 269, 3489-3497.