STRUCTURAL AND FUNCTIONAL ANALYSIS OF THE YEAST GENERAL TRANSCRIPT ELONGATION FACTOR, TFIIS

By

DONALD E. AWREY, B.Sc., M.Sc.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Doctor of Philosophy

McMaster University

YEAST GENERAL TRANSCRIPT ELONGATION FACTOR, TFIIS

Doctor of Philosophy (1997) McMaster University

(Biochemistry) Hamilton, Ontario

TITLE: Structural and functional analysis of the yeast general transcript elongation factor, TFIIS.

AUTHOR: Donald E. Awrey, B.Sc. (McMaster University), M.Sc. (McMaster University)

SUPERVISOR: Professor A.M. Edwards

NUMBER OF PAGES: x, 162

ABSTRACT

Transcription by RNA polymerases is regulated at several stages, including initiation, elongation, and termination. The rate of transcript elongation can be reduced by a multitude of factors, two of which include the template DNA sequence and DNA binding proteins. DNA sequences can cause the elongating polymerase to become stalled. Stalled ternary complexes containing eukaryotic RNA polymerase II require the transcription factor TFIIS to stimulate the resumption of transcription. This stimulation involves the endonucleolytic cleavage and release of oligonucleotides from the 3' terminus of the nascent transcript, with subsequent progression through these modulatory sequences.

I have determined the domain structure of yeast TFIIS by limited proteolytic digestion and assigned functions to the domains required for TFIIS activities *in vitro*. Using a novel binding assay, the surfaces of both yeast RNA polymerase II and yeast TFIIS responsible for their association were determined. The region of yeast RNA polymerase II mediating the interaction with TFIIS was determined to include amino acid residues between 1208 and 1264, which is conserved among eukaryotic RNA polymerase II homologues. The region of TFIIS that interacts with RNA polymerase II was localized to the region 143 to 240 and subsequently to a cluster of amino acid residues contained within this domain. The carboxyl terminal domain of TFIIS is required for the stimulation of endonucleolytic activity by elongating RNA polymerase complexes. Based on the three dimensional structure of TFIIS, I have identified

amino acid residues that are important for the stimulation of the nascent transcript cleavage activity. Several site-directed mutants of TFIIS were unable to stimulate the cleavage activity, yet maintained wild type binding characteristics. I propose a model that describes the mechanism of TFIIS stimulated release of RNA polymerase II from stalled ternary complexes.

ACKNOWLEDGEMENTS

I would like to express my gratitude to Dr. Caroline Kane for gifts in the form of wisdom, plasmids, and reagents. I greatly appreciate the tutelage provided by Dr. Aled Edwards.

I thank all my colleagues for making the time spent at McMaster an enjoyable and memorable experience. Steve, Sally, Jean, Rich, Alexei, Lena, Chris, Kirsten, Mabrouk, and all the rest of the people who could put up with me - best of luck.

Exceptional thanks are given to Christina, Terance, Brianne, and Jessica for their guidance in experimental design, technical assistance, and the preparation of this thesis.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

Å	angstrom
ATP	adenosine 5'-triphosphate
CAPS	3-[cyclohexylamino]-1-propanesulfonic acid
CTD	carboxyl terminal domain
CTP	cytosine 5'-triphosphate
DNA	deoxyribonucleic acid
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
GTP	guanosine 5'-triphosphate
HEPES	N-(2-hydroxyethyl)piperazine-N'-(2-ethanesulfonic acid)
HMK	heart myosin kinase
HSQC	heteronuclear single quantum coherence
IPTG	isopropyl β-D-thiogalactoside
kDa	kilodalton
NMR	nuclear magnetic resonance
NOE	nuclear Overhauser effect
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PMSF	phenylmethylsulfonylfluoride

r.m.s.	root mean squared
RNA	ribonucleic acid
SDS	sodium dodecyl sulfate
TAF	TBP associated factor
TBP	TATA binding protein
Tris	tris(hydroxymethyl)aminomethane
UTP	uridine 5'-triphosphate

CHAPTER 1

Introduction

I. General overview

DNA dependant RNA polymerases synthesize ribonucleic acid (RNA) using deoxyribonucleic acid (DNA) as a template. RNA is synthesized by the 5' to 3' esterification of monomeric ribonucleotide triphosphate substrates in a processive process, directed by Watson-Crick base pairing with the template DNA. The synthesized RNA can direct protein synthesis or provide the structural component of complexes that require RNA as a catalytic or structural component.

Eukaryote RNA polymerases can contain anywhere from 12 to 15 subunits. There are three classes of eukaryotic RNA polymerases called I, II, and III, based upon the class of RNA that each synthesizes. The structure determination of several multisubunit RNA polymerases revealed the conservation of structural features, and provide insight into the catalytic mechanisms of transcription (Darst et al., 1989; Darst et al., 1991; Schultz et al., 1993).

The regulation of transcription by RNA polymerase II plays a major role in the control of cellular function. RNA polymerase is regulated by general transcription factors throughout the transcription cycle; which includes promoter recognition, transcript initiation (Conaway et al., 1990), elongation (Reines, 1994) and termination (Kane, 1994). In this thesis, I will be considering the regulation of transcript elongation by eukaryotic RNA polymerase II. The expression level of many cellular genes (including human c-myc, human L-myc, murine c-fos, and human adenosine deaminase) is controlled by transcription factors during transcript elongation by RNA

polymerase II (Kerppola and Kane, 1990; Spencer and Groudine, 1990). General transcript elongation factors act on the transcribing eukaryotic polymerase to increase both the rate of nucleotide addition and productivity of transcription, where productive transcription is defined as the ability of elongating polymerases to produce full length transcripts. TFIIS is the only known eukaryotic transcript elongation factor that can stimulate the progression of transcribing RNA polymerases through regions of DNA that have induced the complex to halt transcription (Reines, 1994).

II. RNA polymerase structure and architecture

Three families of RNA polymerases have been established based upon the complexity of RNA polymerase subunit composition. Single subunit polymerases are encoded by the genomes of bacteriophage and are also contained within the mitochondria of eukaryotic cells (Masters et al., 1987). Eubacterial polymerases and several eukaryotic viral polymerases have a multi-subunit structure that is typified by the *E. coli* RNA polymerase, which contains four core subunits (Burgess, 1976). A third class of RNA polymerases containing 12 to 15 subunits, includes the archaebacterial and eukaryotic enzymes. The eukaryotic enzymes are similar in architecture and many subunits share considerable amino acid sequence homology.

Single subunit RNA polymerases

The genomes of bacteriophage encode for monomeric RNA polymerases that perform the functions of promoter recognition, promoter clearance, transcript elongation, and termination without a requirement for trans-acting factors. T7 RNA polymerase is 99 kDa in size and its three dimensional structure has been solved (Sousa et al., 1993). The structure, in conjunction with site-directed mutagenic data, revealed the conformation of the active site and nucleic acid binding domains. T7 polymerase is composed of several conserved structural features, for which functions have now been identified. These domains include a 'thumb' and 'finger' region, which are implicated in DNA template binding, and a 'palm' region containing the residues important for catalytic activity. These three regions are spatially arranged with the 'thumb' and 'finger' domains located on either side of the 'palm' region forming a deep cleft that can accommodate double stranded DNA.

E. coli RNA polymerase

The *E. coli* RNA polymerase is composed of four subunits, β, β', and two α polypeptides. The largest subunit, β', contains 1407 amino acid residues and binds to the template DNA (Nudler et al., 1996). Crosslinking and mutagenesis studies have localized the catalytic site to three aspartic acid residues, contained within one of several conserved regions of amino acid sequence shared between eukaryotic and prokaryotic polymerases (Zaychikov et al., 1996). The second largest subunit, β, is

1342 amino acids in length and interacts with the nascent RNA chain (Markovtsov et al., 1996), ribonucleotides (Grachev et al., 1987), and the DNA template (Nudler et al., 1996). The carboxyl terminal half has been shown to bind to both the α and β' subunits. The third largest subunit, α, interacts with the RNA polymerase as a dimer and is composed of 329 amino acids (Ishihama, 1981). The 85 carboxyl terminal amino acid residues of the α subunit are involved in the dimerization of α subunits and in sequence specific DNA binding to sequences immediately upstream of a promoter (Blatter et al., 1994). *E. coli* RNA polymerase can be reconstituted from individually expressed and purified subunits, allowing for extensive biochemical analysis of subunit functions (Borukhov and Goldfarb, 1993). The forthcoming identification of functions for each domain within *E. coli* RNA polymerase will contribute to our understanding of transcription by multisubunit polymerases because the catalytic subunits are conserved from bacteria to man.

The three dimensional low resolution structure (23Å resolution) of *E. coli* RNA polymerase revealed a large cleft that can accommodate double stranded DNA (Darst et al., 1989). One side of this cleft is composed of a 'thumb' projection, as described for the structure of T7 polymerase. Electron microscopy of both core and holoenzyme RNA polymerase preparations showed two distinct conformations. The structure of the holoenzyme complex indicates the presence of the thumb domain in an 'open' position, exposing the cleft; whereas analysis of the core RNA polymerase indicates that, in the absence of a sigma factor, this cleft becomes a channel (Polyakov et al.,

1995). The 'thumb' domain is postulated to be flexible and capable of moving across the cleft to surround the double stranded DNA. This 'closed' conformation may be responsible for yielding the processive elongation complex.

Eukaryotic and Archaebacterial RNA Polymerase Composition

The subunit composition of eukaryotic RNA polymerases is conserved from yeast to man. Amino acid sequence comparisons of available eukaryotic RNA polymerase subunits indicate that RNA polymerase isoforms I, II, and III are similar in architecture across a wide range of species and contain 12 to 15 polypeptides (Sentenac et al., 1992). The subunit structure of yeast RNA polymerases will be described here since yeast and higher eukaryotes typically share 50% or higher amino acid sequence identity between corresponding subunits. Archaebacterial RNA polymerases have a similar subunit architecture to yeast RNA polymerases; these enzymes share approximately 25% amino acid sequence identity (Puhler et al., 1989).

The nomenclature system for identification of eukaryotic RNA polymerases and their subunits define the three eukaryotic RNA polymerases I, II, and III as A, B, and C, respectively (Roeder and Rutter, 1969). The subunits are identified by their molecular weight, as estimated from denaturing SDS-PAGE. For example, the second largest subunit of RNA polymerase I is identified as A135, and the third largest subunit from RNA polymerase II is identified as B45.

The three eukaryotic RNA polymerase isoforms each contain high molecular

weight subunits that share regions of homology with the two largest subunits of *E. coli* RNA polymerase, β' and β (Allison et al., 1985). There are eight conserved regions among β' and the largest subunits of the eukaryotic polymerases, designated A to H, starting from the amino terminus. These conserved regions are thought to harbour the nucleic acid binding and catalytic domains; the intervening divergent regions may be involved in polymerase specific regulatory events. The carboxyl terminal domain (CTD) of the largest subunit of eukaryotic RNA polymerase II is composed of a repetitive sequence that undergoes phosphorylation (Cadena and Dahmus, 1987) and glycosylation (Kelly et al., 1993). The relevance of these post-translational modifications to the CTD is unclear; the CTD is required for cell viability and has been implicated in the activation of transcription initiation by transcriptional activators (Nonet et al., 1987). The role of the phosphorylation state of the CTD in transcript elongation will be discussed later. The CTD also acts to recruit RNA processing activities to the transcribing polymerase which may play a role in transcript termination as well as promote transcript processing (McCracken et al., 1997). The second largest subunit of eukaryotic RNA polymerases contains nine regions of sequence similarity to the *E. coli* RNA polymerase subunit β, designated A to I (starting from the amino terminus) (Nonet et al., 1987) and is implicated in binding to nucleotide substrates (Riva et al., 1987). The third largest subunit of RNA polymerase II, B45, contains a region of homology with the alpha subunit from *E. coli* (Kolodziej and Young, 1989), RNA polymerase II subunit B13 (also known as RPB11), and the

subunits AC40 and AC19 (Woychik and Young, 1993), which are common subunits shared between RNA polymerases I and III. B32 and B17 are two subunits that can dissociate from RNA polymerase II under mildly denaturing conditions and are required for transcription initiation from promoter elements (Edwards et al., 1991). Disruption of the gene encoding B32 in yeast results in viable cells (Woychik and Young, 1989); whereas B17 is required for cell viability (McKune et al., 1993). Subunit B17 is homologous to C25, a subunit of RNA polymerase III (Sadhale and Woychik, 1994). The ninth largest subunit of yeast RNA polymerase II is not required for cell viability and is homologous to subunit A12.2 (Woychik et al., 1991). The eleventh largest subunit of RNA polymerase II has been cloned from yeast and contains regions of homology to subunit AC19 (Woychik and Young, 1993). There are five subunits, ABC 27, ABC 23, ABC 14, ABC 10α, and ABC 10β, encoded by single copy genes, that are shared between the three isoforms of RNA polymerase (Woychik et al., 1990).

Although there are similarities between the three classes of RNA polymerases, there are several subunits that are unique to each isoform. Subunits that do not appear to have homologues within the other polymerase isoforms include: three subunits of RNA polymerase I (A49, A43, and A34.5), RNA polymerase II subunit B32, and 4 subunits of RNA polymerase III (C82, C53, C34, and C31) which are required for cell viability (reviewed in Sentenac et al., 1992). Subunits unique to each polymerase are likely involved in regulatory events particular to each isoform.

The three dimensional structures have been solved at low resolution for both yeast RNA polymerases I and II (Schultz et al., 1993; Darst et al., 1991). The low resolution structures (30 Å and 16 Å resolution for RNA polymerase I and RNA polymerase II, respectively) obtained by electron microscopy showed the conservation of structural features with the bacterial RNA polymerase. These features include the presence of a cleft that is large enough to accommodate double stranded DNA, a 'thumb' projection that forms one wall of the cleft, and a domain forming the 'fingers' of the polymerase which protrude from the cleft wall opposite the 'thumb'. The simplest model of RNA polymerase structure describes the polymerase as a hand which grasps the DNA double helix by closing the thumb projection across the DNA within the cleft.

III. Multisubunit RNA polymerase activities are regulated by protein factors.

RNA polymerase transcription factors regulate RNA synthesis at all stages of the transcription process; including start site selection, promoter clearance, productive transcription of the template, and efficient termination. The actions of regulatory factors are not necessarily restricted to one stage of transcription.

Initiation

General transcription factors are required for the recruitment of RNA polymerases to the 5' end of a coding region, however the exact transcription start site

is determined by the polymerase. In contrast, bacteriophage RNA polymerases do not require transcription factors since the polymerase can independently recognize promoter elements (Oakley and Coleman, 1977). The distance from the position of the promoter element to the start site in bacteriophage SP6 and T7 systems is four to five nucleotides (Chapman and Burgess, 1987; Nam and Kang, 1988), 35 nucleotides in *E. coli* (Kobayashi et al., 1990), 40 to 120 nucleotides in yeast (Maicas and Friesen, 1990), and in higher eukaryotes approximately 40 nucleotides (Kollmar and Farnham, 1993). The factors responsible for the assembly of RNA polymerases at promoters are described for *E. coli* and eukaryotic RNA polymerases.

Promoter dependent transcription initiation by *E. coli* RNA polymerase is performed by a holoenzyme that contains a core polymerase and a sigma factor (Burgess et al., 1969). Sigma factors are small, dissociable factors that bind to promoter sequence elements and therefore direct transcription from promoters containing the consensus binding sequence for each particular sigma factor (Gardella et al., 1989). For example, specific sigma factors are responsible for the regulated expression of genes involved in nitrogen metabolism (Kutsu et al., 1989), heat shock response, and stationary phase metabolism (Hengge-Aronis, 1993). The various sigma factors compete for binding to the polymerase, and therefore slight changes in the relative concentrations can rapidly alter the profile of gene expression. The binding of a sigma factor to the core polymerase is disrupted by the interaction of an 8 to 12 nucleotide long nascent RNA with the holoenzyme and dissociation of the sigma

factor is postulated to result in the polymerase conformational change that defines the stage of promoter clearance (Krummel and Chamberlin, 1989).

Initiation by eukaryotic RNA polymerases is also regulated by transcription factors. Transcription initiation by RNA polymerase I is regulated by two general transcription factors, TIF-IB and UBF, which bind to promoter elements of the 18S, 5.8S, and 28S ribosomal genes (Iida et al., 1985). The transcription factor TIF-IB is a complex of at least 5 polypeptides of which one is TATA binding protein (TBP) (Comai et al., 1992). Assembly of these two factors in addition to RNA polymerase I at a promoter is sufficient for the initiation of transcription.

RNA polymerase II requires five general transcription initiation factors for accurate, basal transcription from templates *in vitro*; TFIID, TFIIB, TFIIF, TFIIE, and TFIIH (reviewed in Serizawa et al., 1994). The formation of an initiation complex requires the assembly of these factors at a promoter element. TFIID contains TBP as well as at least seven TBP associated factors (TAF's) in yeast and higher eukaryotes (Zhou et al., 1992). TBP binds to the TATA element of a promoter and, in conjunction with the remaining four general transcription factors, directs basal transcription in the absence of TAF's (Kambadur et al., 1990). The presence of a consensus TATA sequence at the -25 to -30 position of the promoter is not necessary for the assembly of a pre-initiation complex since several class II promoters have been identified that do not encode a TATA element. However, even at these promoters, TBP containing factors are still required for basal transcription (Pugh and Tjian, 1991).

TFIIB is a monomeric protein that can bind to the TBP-promoter complex and interacts with RNA polymerase II (Buratowski et al., 1989). TFIIF, also known as RAP 30/74, is a heterodimer that binds to the RNA polymerase II-TFIID-TFIIB complex (Burton et al., 1988). TFIIH purified from *Drosophila* (Austin and Biggin, 1996), yeast (Svejstrup et al., 1994), and human (Gerard et al., 1992; Flores et al., 1992) contains 8 to 9 polypeptides and displays ATPase and helicase activities, as well as a kinase activity that phosphorylates the carboxyl terminal domain of RNA polymerase II (Schaefer et al., 1993). TFIIE is a heterodimer that interacts with the RNA polymerase II-TFIID-TFIIB-TFIIF-TFIIH complex, forming a fully assembled pre-initiation complex (reviewed in Serizawa et al., 1994). The association of the pre-initiation complex can also occur via the interaction of subassemblies of these factors. Free holoenzyme complexes that contain all of the general transcription factors defined above as well as many TAF's and RNA polymerase associated factors have been identified (Koleske and Young, 1994).

RNA polymerase III transcribes the genes that encode the 5S RNA, tRNA, 7SL RNA, and U6 snRNA as well as several other small RNA products (reviewed in Geiduschek and Tocchini-Valentini, 1988 and Kassavetis et al., 1994). These genes typically contain an internal promoter element that makes base specific contacts with the transcription initiation factors, TFIIIA (monomeric, 51 kDa) and TFIIIC (multimeric, 6 subunits, 132, 120, 74, 90, 55, and 62 kDa) (Churchill et al., 1990; Parsons and Weil, 1990). TFIIIB (multimeric, 3 subunits of 27, 67, and 90 kDa)

contains TBP as a subunit and interacts upstream of the transcription start site forming a DNA-TFIII(A, B, C) complex (Kassavetis et al., 1990). RNA polymerase III associates with this complex resulting in transcript initiation.

Elongation

Following promoter clearance, the RNA polymerase is a component of a nucleic acid-protein complex, defined as a ternary complex, which contains the template DNA, nascent RNA transcript, and RNA polymerase. The nucleic acid components of the ternary complex can influence the RNA polymerase to reduce the rate of transcription (Kerppola and Kane, 1990; Spencer and Groudine, 1990; Kane, 1994). The nucleic acid sequences that affect the transcriptional ability of RNA polymerases are defined as attenuating sequences. Ternary complexes that have a reduced or abrogated rate of transcription are defined as non-productive complexes. The avoidance of attenuation and the release from non-productive complexes, is facilitated by two classes of general transcript elongation factors resulting in an overall increased rate of transcription (reviewed in Kane, 1994).

The first class of eukaryotic transcription elongation factors includes TFIIF (Bengal et al., 1991), elongin (Bradsher et al., 1993), and ELL (Shilatifard et al., 1996). These factors enhance the rate of transcription by binding to the ternary complex and decreasing the time spent by the ternary complex in a non-productive state. TFIIF stimulates RNA polymerase II during both the initiation (Flores et al.,

1988) and elongation stages (Bengal et al., 1991) to promote an increased rate of transcription. The interaction of elongin and ELL with the ternary complex is required prior to encountering an attenuating region of the template DNA (Bradsher et al., 1993; Shilatifard et al., 1996). Homologues of elongin have been isolated from rat (Bradsher et al., 1993), and humans (Duan et al., 1995), and their activities reconstituted *in vitro*. This elongation factor is composed of three subunits, A, B, and C. The elongin A subunit demonstrates weak activity that is stimulated by the presence of subunits B and C. Human elongin subunits B and C have been shown to interact with the von Hippel-Lindau (VHL) tumour suppressor protein (Kibel et al., 1995). Removal or inactivation of the VHL protein results in the inability of VHL to sequester the stimulatory subunits away from elongin A. This, in turn, may result in a constitutively active transcription factor, but its role in oncogenesis, if any, remains to be established.

A second group of transcription factors, GreA and GreB in bacteria and TFIIS in archae and eukaryotes affect elongation by stimulating non-productive complexes to resume transcript synthesis (reviewed in Kassavetis and Geiduschek, 1993). The bacterial transcription factors, GreA and GreB, have a dual role in transcription, stimulating promoter clearance and increasing productive transcription by RNA polymerase (Hsu et al., 1995). TFIIS, GreA, and GreB all stimulate endonucleolytic cleavage near the 3' end of the nascent RNA chain, producing oligonucleotide products that range in length from 2 to 14 nucleotides (Izban and Luse, 1992; Reines, 1992;

Borukhov et al., 1993). The cleavage event is thought to be a prerequisite for stimulating the reactivation of non-productive complexes. Following cleavage of the nascent chain, RNA polymerase resumes synthesis of RNA from the new 3' hydroxyl which is in the correct base pairing register with the template DNA. A more detailed mechanism of stimulation by these factors will be described below.

An increased production of full length transcripts *in vivo* is stimulated by strong transcriptional activators. Transcriptional activators were previously thought to stimulate an increase in the frequency of transcript initiation, however, recent results indicate that the elongating polymerase is also more productive in the synthesis of full length transcripts (Yankulov et al., 1994). Enhancement of RNA polymerase productivity may be due to the recruitment of transcription elongation factors to the pre-initiation complex or the stimulation of the complex after promoter clearance. For example, in the absence of transcriptional activators an increased density of RNA polymerases can be seen at the 5' end of the coding region, indicating that many ternary complexes do not produce full length transcripts. The increased productivity of ternary complexes is postulated to involve the stimulation of CTD phosphorylation by TFIIH which could serve to recruit the general transcript elongation factors to the elongating polymerase (Yankulov et al., 1996).

Termination

The final requirement for production of biologically active transcripts is the

termination of transcription, which involves the release of the transcript and dissociation of the polymerase from the DNA template. In *E. coli*, a well characterized RNA binding protein, rho, stimulates transcription termination by binding to the nascent RNA chain and removing the RNA from the ternary complex through an ATP dependant process (Das et al., 1978; Brennan et al., 1987). TTF-I, an RNA polymerase I transcription termination factor (Bartsch et al., 1988), enhances termination by inducing the ternary complex to pause, stimulating the release of the RNA transcript, and promoting the 3' end processing of the transcript (Kuhn and Grummt, 1989). A termination factor, La, has been described which promotes the accurate termination by RNA polymerase III (Maraia et al., 1994). Termination factors, although they may exist for RNA polymerase II, have not been characterized.

IV. Regulatory Mechanisms of Transcript Elongation

Nucleic acid sequences promote attenuation of ternary complexes

Elongation by RNA polymerase *in vitro* can be modulated by sequence specific interactions with the DNA template (Kerppola and Kane, 1990). Consequently, the rate of RNA synthesis and the stability of the ternary complex are determined by the position of the ternary complex on the template. *In vivo*, RNA polymerases encounter a variety of additional impediments. For example, DNA within the nucleus of a cell is complexed with proteins such as histones (Felsenfeld, 1992), DNA sequence specific

binding proteins (Kuhn et al., 1990), as well as components of the repair and replication machineries (French, 1992; Drapkin et al., 1994). The transcribing polymerase may also become blocked by DNA and RNA conformation, intra- and inter-strand DNA crosslinks (Donahue et al., 1994), and the presence of DNA binding drugs (Mote et al., 1994).

Attenuating sequences also can induce the formation of non-productive complexes, of which there are two types. The first class of non-productive complex is defined as a paused ternary complex. These complexes are fully capable of resuming transcript elongation in the absence of transcription factors. The formation of paused complexes can be induced by conditions that slow the rate of nucleotide addition, such as limiting nucleotide concentrations (Izban and Luse, 1993), non-optimal buffer conditions, nucleic acid secondary structure (Arndt and Chamberlin, 1990), and the presence of DNA binding proteins (Donahue et al., 1994). The ability to induce the formation of a paused complex at any nucleotide position indicates that there is an equilibrium between two competing conformations of RNA polymerase; transcription can either progress by the addition of the next nucleotide, or revert to paused state. The release from a paused complex is a time dependent event with the half-life of the paused complex being governed by the nature of the transcriptional impediment (Gu and Reines, 1995). Transcription factors TFIIF, elongin, and ELL act to either prevent the formation of paused ternary complexes or to facilitate the conversion of a paused complex to an elongation competent conformation.

A second type of non-productive ternary complex cannot resume transcription without stimulation by transcription factors. These complexes, defined as 'arrested complexes', are unable to effect nucleotide addition (Reines et al., 1989), yet maintain contacts with both the DNA and RNA. The formation of an arrested complex is proposed to involve an intermediary paused complex. The rate of conversion from a paused complex to an arrested conformation is dependant upon the DNA and RNA secondary structure (Kerppola and Kane, 1990), as well as nucleic acid interactions with the ternary complex (Markovtsov et al., 1996). The arrested complexes are extremely stable and are likely regulatory control points of transcript elongation.

There are at least two types of intrinsic pause and arrest sites. The first type involves the presence of RNA secondary structure near the 3' terminus of transcripts contained within stalled ternary complexes. This secondary structure induces a conformational change within the ternary complex, which results in the formation of non-productive complexes. For example, the human c-myc and N-myc genes have predicted RNA stem loop structures approximately 10 nucleotides upstream of one of the pause sites (Kerppola and Kane, 1990), and in prokaryotic systems RNA secondary structure can also result in transcript pausing, arrest, and termination events (reviewed in Landick and Yanofsky, 1987).

The regulation of transcript elongation by pause and arrest sites is the best characterized in bacteria. The *E. coli trp* and *his* genes contain strong pause sites approximately 10 nucleotides from the base of RNA stem loop structures (Lee et al.,

1990), and one of the most well characterized *E. coli* attenuation sites is contained within the *trp* operon (reviewed in Chan and Landick, 1994). Transcription of the *trp* operon is regulated at the level of elongation, coupled with translation, in which the nascent transcript is aborted if tryptophan is present in abundance. A decreased tryptophan concentration results in the productive transcription of the *trp* operon. This regulation is controlled by the formation of RNA secondary structures as well as downstream DNA sequences that signal, through an unknown mechanism, for either transcript attenuation or the formation of the full length transcript. Mutagenesis or substitutions that affect either the stem loop structure or the downstream DNA sequence alter the propensity of the ternary complex to pause at these sites.

The second type of intrinsic transcript attenuation site does not contain any predicted RNA secondary structure. Mutagenesis or deletion of DNA sequences downstream of the arrest site reduces the propensity for ternary complexes to become non-productive (Kash et al., 1989). Sites at which non-productive complexes arise can not be predicted by the presence of secondary structure or nucleotide sequence due to the lack of a correlation between known attenuating sites and the DNA sequence surrounding them. These observations indicate that there may be additional proteins or nucleic acid interactions with the ternary complex that signal for attenuation of transcript elongation.

Structural topology of ternary complexes

Knowledge of the structural topology of the elongating complex is required prior to the development of an understanding of the mechanism by which transcript elongation is regulated. Although the atomic resolution structure of a ternary complex is not known, the proximity of nucleic acids to each other and to several domains of polymerase has been determined by photo cross-linking, site-directed mutagenesis, and digestion of nucleic acids by nucleases. The DNA duplex contains a region in which the DNA strands are separated (melted), and the 3' terminal 3 to 12 nucleotides of the nascent RNA form an RNA:DNA hybrid within the melted region (Hanna and Meares, 1983). This nucleic acid structure is called a transcription bubble and the thymidine bases within this bubble are $KMnO_4$ sensitive, indicating a distortion of nucleotide base pairing (Gamper and Hearst, 1982). The size of the transcription bubble is proposed to be approximately 18 bases long, however transcription bubbles can range in size from 12 to 23 bases in length depending upon the template DNA sequence. Artificially created transcription bubbles, formed by annealing DNA strands containing a 12 base pair region of non-complementarity and an RNA containing 12 nucleotides of complementarity at the 3' terminus can be used to promote transcription from the 3' hydroxyl of the annealed RNA strand (Daube and von Hippel, 1992).

Nucleolytic digestion of the DNA template contained within both paused and arrested complexes indicates that the RNA polymerase footprint varies in size (Levin et al., 1987; Krummel and Chamberlin, 1992). This change in the size of the protected template is dependent upon the DNA sequence (Nudler et al., 1994).

Ternary complexes were analyzed by nucleolytic digestion of ternary complexes halted by nucleotide depletion at sequential nucleotides along the DNA template. These studies indicated that the forward progression of the leading edge of the polymerase was not monotonous (one base at a time), but rather discontinuous, advancing by up to 10 nucleotides in one increment upon the addition of the next nucleotide. Analysis of successive nucleotide additions to the halted ternary complex demonstrated that the trailing edge of the polymerase was moving in concert with the growth of the nascent RNA chain. The discontinuous movement of the leading edge of the polymerase was described as 'inchworming' (Chamberlin, 1995), and first showed the conformational heterogeneity of the elongating ternary complex. Further analysis demonstrated that transcription by RNA polymerase consists of both monotonic as well as discontinuous movement, which method is utilized by the polymerase is dependent upon the template DNA sequence (Nudler et al., 1994). Inchworming by the polymerase is associated with the encounter of both pause and termination sites.

Protein contacts with the transcription bubble have been determined by mutant analysis, salt sensitivity, and the use of photo cross-linkable nucleotides. Residues that are important for stabilizing interactions with the nucleic acid have been localized to the amino terminal conserved region A of the largest subunit, and both the amino terminal conserved region B and the carboxyl terminal conserved region I of the second largest subunit of *E. coli* RNA polymerase (Nudler et al., 1996). Mutagenesis of cysteine residues that are proposed to form a zinc binding motif in conserved region

A of the largest subunit reduces the processivity of the ternary complex and confers salt sensitivity to ternary complexes (Nudler et al., 1996). The conserved region B of the second largest subunit can be cross-linked to a region of the DNA duplex located downstream of the point where strand separation occurs. These two regions of the largest and second largest subunits are likely involved in stabilization of the ternary complex through interactions with nucleic acid near the downstream edge of the polymerase. Conserved region I of the second largest subunit is identified by photo cross-linkable nucleotides incorporated into the template strand near the catalytic site. This region may play a role in the stabilization of the RNA:DNA hybrid immediately upstream of the catalytic site.

Amino acid residues of RNA polymerases found to be in close proximity to the catalytic site include the conserved regions D and G of the largest subunit as well as the conserved regions D and H of the second largest subunit (Markovtsov et al., 1996). Substitution of Mg^{2+} with Fe^{2+} following the formation of halted ternary complexes, and subsequent treatment with dithiothreitol, enabled hydroxyl radical cleavage of the largest subunit within conserved region D (Zaychikov et al., 1996). Mutagenesis of three aspartic acid residues within conserved region D of the largest subunit abolished Fe^{2+} mediated hydroxyl radical cleavage. The mutant RNA polymerase was deficient in nucleotide polymerization, yet nucleic acid binding activity was maintained.

Photo cross-linkable nucleotides incorporated at the 3' end of the nascent RNA chain were used to show that there is a change in the location of the 3' terminus of the

RNA upon conversion from an elongation competent conformation to an arrested complex (Markovtsov et al., 1996). Two conserved regions, D and H, of the second largest subunit of RNA polymerase are preferentially reactive to photo cross-linkable nucleotides when the ternary complex is in an elongation competent conformation. Analysis of arrested ternary complexes, which have a reduced elongation competence, demonstrate a reduced reactivity of photo cross-linking at these same sites. Two conserved regions, D and G, of the largest subunit demonstrated a preferential reactivity to photo cross-linkable nucleotides incorporated into the 3' terminal position of the nascent transcript when the ternary complex is an arrested conformation. Again this reactivity is reduced when the halted ternary complexes remain in an elongation competent conformation. These results identify regions of the RNA polymerase which selectively interact with the nascent RNA depending upon the elongation competence of the ternary complex and establish a structural difference between productive and arrested complexes.

A model of transcript elongation

The mechanism by which polymerases recognize and progress through DNA sequences that attenuate transcription is not fully understood. The inchworm model of transcript elongation (Chamberlin, 1995) incorporates the facts stated above and describes two DNA binding sites, an RNA binding site and a catalytic site. The main tenet of this model describes a leading DNA binding site that can move independently

of the remainder of the polymerase. Upon encountering an attenuating DNA sequence the forward progress of the polymerase is impeded and the polymerase converts from monotonic transcription to a discontinuous mode. The catalytic site and the trailing edge of polymerase then continue to advance synchronously in a 5' to 3' direction even after the leading edge halts forward progression along the template. The discontinuous movement of the leading edge of the polymerase may be due to two conformations of the leading DNA binding domain, locked and unlocked. Upon encountering the attenuating region of DNA the leading DNA binding site locks in position until an RNA binding site is filled. The leading DNA binding domain then unlocks, resulting in a large incremental downstream shift and subsequently locks on the template starting a new round of inchworm movement. By necessity, all three postulated nucleic acid binding sites interact with nucleic acid in an essentially non-sequence dependant manner. However, the relative stability of these binding domains is determined by the nucleic acid sequence and presence of nucleic acid binding proteins. In this model, RNA secondary structure and downstream untranscribed sequences contact the leading DNA binding domain and induce a conformational change within the ternary complex, resulting in the formation of a paused complex. Three events are then possible; progression through the inhibitory region, conversion to an arrested complex, or premature termination.

An extension of the inchworm model describes the location of the RNA transcript within the ternary complex (Wang et al., 1995). Upon formation of an

arrested complex there is a conformational change resulting in the relocation of the 3' terminal nucleotide of the nascent RNA. Although the spatial location of the 3' terminal region of the nascent RNA in arrested complexes has not been determined, Fe^{2+} hydroxyl radical cleavage of the nascent RNA indicates that the 3' terminal nucleotide is not contained within the catalytic site (Nudler et al., 1997). The 3' terminus may be annealed with the DNA to form an RNA:DNA hybrid downstream of the catalytic site or RNA:protein contacts could be made with the polymerase. The conversion of a paused complex to an arrested complex likely involves the stable binding of the 3' terminus of the RNA to these downstream components of the ternary complex. This stable binding inhibits transcript elongation and requires the action of transcription factors to promote release from this conformation.

This model of transcript elongation incorporates observations from the *E. coli* transcription apparatus, but can also be applied to the eukaryotic RNA polymerase II (Rice et al., 1991) and RNA polymerase III enzymes.

V. Transcript cleavage activity

Cleavage of the nascent transcript is associated with the re-activation of arrested ternary complexes (Reines, 1992; Borukhov et al., 1993) and the maintenance of the fidelity of transcription (Erie et al., 1993; Jeon and Agarwal, 1996). The transcript cleavage activity has been observed in many transcription systems including vaccinia (Hagler and Shuman, 1992), *E. coli* (Borukhov et al., 1993), and all three nuclear

eukaryotic polymerases (Reines, 1992; Whitehall et al., 1994; Labhart, 1997). The products generated by the transcript cleavage activity include a short oligonucleotide containing a 5' phosphate (Izban and Luse, 1993), which is released from the ternary complex, and the remainder of the transcript, containing a 3' hydroxyl, which is retained within the ternary complex. The transcript can be re-extended from the new 3' hydroxyl in a template dependent manner.

The endonucleolytic transcript cleavage activity of RNA polymerases can be stimulated by several transcription factors which include GreA (Borukhov et al., 1992), GreB (Borukhov et al., 1993), TFIIS homologues (Reines, 1992; Christie et al., 1994), and a column fraction that stimulates the cleavage of transcripts by eukaryotic RNA polymerase I (Labhart, 1997). These stimulatory transcription factors are known to bind to their respective RNA polymerases (Sawadogo et al., 1980; Orlova, 1995) and either contact or are in close proximity to the 3' terminus of the nascent RNA of halted ternary complexes (Powell et al., 1996).

The size of the cleavage product is dependent upon the nucleotide sequence of the RNA transcript (Izban and Luse, 1993). Binary complexes constructed with RNA oligonucleotides and *E. coli* RNA polymerase demonstrated that GreA and GreB can stimulate the cleavage of RNA in the absence of a DNA template and, upon addition of nucleotides, the binary complex can add several residues in a non-template dependent manner (Altmann et al., 1994). Unfortunately, the same transcripts were not analyzed for cleavage site specificity in the context of a ternary complex.

Analysis of RNA polymerase footprints at positions immediately upstream of a pause site indicates that the length of the cleavage product is associated with the discontinuous movement of the leading edge of the polymerase (Wang et al., 1995; Nudler et al., 1997). Footprint analysis of successive halted complexes along a template indicate that the leading edge of the polymerase moves closer to the position of the 3' hydroxyl at an attenuating DNA sequence. Ternary complexes that are 'walked' along the template DNA by addition of a limited set of nucleotides demonstrate an increased sensitivity to GreB, as the ternary complex approaches a pause site. The length of the cleavage product generated by stimulation with GreB correlates with the distance by which the leading edge of the polymerase has moved closer to the 3' hydroxyl of the RNA.

TFIIS dependent transcript shortening is not dependent upon translocation of the ternary complex on the template DNA. Analysis of the DNA footprint of RNA polymerase II before and after transcript cleavage showed that the position of the polymerase on the template does not change. Although the position of the polymerase does not change, the size of the protected DNA increases to include additional bases at the leading edge of the polymerase. Footprint analysis of polymerase, paused at successive nucleotides along the DNA template, showed a reduction in the size of the footprint occurs prior to the conversion of an elongation competent complex to an arrested complex (Gu et al., 1993). The lack of translocation following the cleavage event indicates that the polymerase is positioned to reform the proper register between

the RNA at the cleavage site with the DNA template prior to transcript cleavage.

Transcript cleavage by the ternary complex has been demonstrated to be an intrinsic property of both RNA polymerase II and *E. coli* RNA polymerase (Orlova et al., 1995). The intrinsic cleavage activity of RNA polymerase II results in the release of several oligonucleotides, some of which are the same length as those produced by TFIIS stimulated ternary complexes. Ternary complexes formed with human RNA polymerase II also demonstrate endopyrophosphorolysis, the internal cleaving of the nascent RNA chain upon addition of pyrophosphate, which produces several cleavage products of a similar size as those produced by TFIIS stimulation (Rudd et al., 1994). GreA and GreB stimulated cleavage products are also a similar size as those produced by the intrinsic cleavage activity of *E. coli* RNA polymerase. Furthermore, the shortened transcripts produced by the intrinsic transcript cleavage activity of *E. coli* RNA polymerase are extended upon addition of nucleotides and restoration of the optimal pH for transcript elongation (Orlova et al., 1995).

VI. Transcript elongation factor TFIIS

TFIIS homologues have been identified, and the DNA sequences cloned, from human, rat, mouse, *Drosophila*, *Xenopus*, yeast, archaebacteria, and as a component of the vaccinia virus RNA polymerase (reviewed in Chen et al., 1992). Several tissue specific isoforms of TFIIS exist in rat, mouse, *Xenopus*, and humans; the major divergent amino acid sequences among isoforms are localized to the amino terminal

region. The yeast TFIIS homologue is 309 amino acid residues in length (Nakanishi et al., 1992). The amino terminal 100 amino acid residues contain a site(s) at which TFIIS is phosphorylated *in vivo* (Horikoshi et al., 1985), and also contain a region of homology to human elongin A (Aso et al., 1995). This region of TFIIS may be involved in the regulation of these transcription factors. The central 100 amino acids of TFIIS homologues share considerable similarity, are required for TFIIS activity, and are postulated to contain the binding surface which interacts with RNA polymerase II (Agarwal et al., 1991). The carboxyl terminal 60 amino acids of TFIIS homologues have a high degree of sequence conservation and contain a tetrad of cysteine residues which chelate a zinc ion (Agarwal et al., 1991). The three dimensional structure of this domain was solved by NMR (Qian et al., 1993a; Qian et al., 1993b) and found to contain a novel zinc binding motif, a 'zinc ribbon', whose structure is similar to a domain of TFIIB (Zhu et al., 1996). The function of the zinc ribbon is unknown, however it has been suggested to be involved in binding to both single stranded and duplex nucleic acid (Agarwal et al., 1991). Several other proteins have been predicted to contain the zinc ribbon motif, eg. B13 (the ninth largest subunit of RNA polymerase II) (Woychik et al., 1991), TFIIE (Petersen et al., 1991), DNA polymerase α (Pizzagalli et al., 1988), and T4 DNA primase (GenBank accession #118732).

TFIIS was initially identified as a column fraction that stimulated transcription in promoter independent assays (Sekimizu et al., 1976). This transcription factor was subsequently determined to act during transcript elongation and to stimulate

transcription by the ternary complex through pause and arrest sites *in vitro* (Sawadogo et al., 1981; Reinberg and Roeder, 1987; Reines et al., 1989). Ternary complexes paused by nucleotide depletion (Izban and Luse, 1992), DNA binding drugs (Mote et al., 1994), and DNA binding proteins (Donahue et al., 1994) as well as arrested complexes formed by sequence specific interactions with the template are responsive to TFIIS (Reines, 1992). Paused and arrested ternary complexes are stimulated by TFIIS to cleave near the 3' terminus of the nascent RNA chain resulting in the release of two to fourteen nucleotide oligomers (Izban and Luse, 1993). The length of the released 3' terminal fragment is dependent upon the nucleotide sequence of the template DNA. A comparison of the length of the cleavage products generated by TFIIS stimulation of either paused or arrested complexes reveals that paused ternary complexes produce cleavage products of a shorter length than products generated by arrested ternary complexes (Izban and Luse, 1993). This difference is likely related to the structural difference between these two types of complexes, as defined above. The cleavage event leaves the nascent RNA chain in register with the template DNA allowing nucleotide addition to occur in a template dependent manner. Purified ternary complexes containing shortened RNA, due to stimulation by TFIIS, are transcriptionally active; nucleotide supplementation results in transcript elongation (Reines, 1992). Following factor stimulated cleavage of the transcript, the polymerase resumes transcription and attempts to progress through the arrest site; repeated cleavage and re-extension of the transcript may be required to progress through an

attenuating DNA sequence (Gu et al., 1993). The cleavage activity is not solely responsible for converting an arrested complex to an elongation competent form; additional activities associated with TFIIS may influence the ternary complex (Ciprés-Palacín and Kane, 1994, Awrey et al., 1997). TFIIS has also been attributed the activity of stimulating the transfer of one strand of duplex DNA to a complementary single stranded DNA (Clark et al., 1991). The significance of the strand transfer activity is unknown and is not reproducible with TFIIS produced by over-expression in *E. coli*.

Disruption of the TFIIS gene in yeast results in an *in vivo* phenotype involving sensitivity to 6 azauracil (Hubert et al., 1983), a compound that is thought to reduce the intracellular concentrations of GTP (Exinger and Lacroute, 1992). A genetic screen was used to identify mutants of RNA polymerase II by the selection of yeast strains that are both 6 azauracil sensitive and demonstrate suppression of 6 azauracil sensitivity upon over-expression of TFIIS. The identification of seven RNA polymerase II mutants localized between the conserved G and H regions of the largest subunit of RNA polymerase II indicated a possible site of interaction between RNA polymerase II and TFIIS (Archambault et al., 1992) that was confirmed by biochemical studies (Wu et al, 1996).

VII. Objectives of this research

The role of TFIIS is therefore defined as a stimulatory activity that induces an

RNA cleavage event by the RNA polymerase allowing the resumption of transcription from the newly formed 3' terminal nucleotide. The mechanism by which TFIIS stimulates an arrested ternary complex is unknown. I have undertaken a structural approach, coupled with site-directed mutagenesis and activity assays, to determine the nature of the interactions between TFIIS and the ternary complex. The objective of this research is to carry out a structure based assignment of TFIIS activities, which includes identification of the surfaces of TFIIS and RNA polymerase II responsible for the protein-protein interaction. The amino acid residues responsible for stimulation of nascent transcript cleavage are also characterized. Based on the results of this investigation a model is presented which provides the framework for future investigation of transcript elongation through regulatory sequences that promote the arrest of ternary complexes.

CHAPTER 2

Materials and Methods

Purification of Yeast RNA Polymerase II

RNA polymerase II was purified from baker's yeast as described previously (Edwards et al., 1990), except that 10 μM $ZnCl_2$ was included in all buffers. The polymerase preparations were stored at -70 °C in 20 mM Tris-HCl pH 7.9, 40 mM ammonium sulphate, 10% glycerol, 10 mM DTT, 1 mM EDTA and 10 μM $ZnCl_2$.

Purification of Yeast TFIIS

Yeast TFIIS was cloned and expressed as a fusion protein containing an amino terminal hexahistidine tag in BL21(DE3) cells. The cells were grown in either Luria-Bertani broth or M9 minimal medium containing ^{13}C-labelled glucose (Cambridge) and / or $^{15}NH_4Cl$ for NMR studies. Bacterial cultures were grown at 37 °C to an OD_{600} between 0.4 and 0.8, and TFIIS expresssion induced with IPTG at either 30 °C or 37 °C (expression temperature for mutant and wild-type TFIIS, respectively). All subsequent procedures were performed at 4 °C. Bacterial cells expressing the yeast TFIIS protein were lysed using a French Pressure cell (American Instrument Company, MD) in Buffer A (20 mM Hepes pH 7.5, 10% glycerol, 10 μM $ZnCl_2$, 1 mM PMSF, and 1 mM benzamidine) containing 100 mM NaCl, 10 mM DTT, and 1 mM EDTA. The lysate was clarified by centrifugation at 55,000 x g for 30 minutes in a Beckman SW28 rotor, and the supernatant loaded onto a DE52 column (2.5 x 2.5 cm; Whatman, Maidstone, England) equilibrated with Buffer A. The flow through was loaded onto an SP Fast Flow column (1 x 8 cm; Pharmacia, Baie D'Urfe, Quebec), washed with

Buffer A containing 100 mM NaCl and eluted with Buffer A containing 500 mM NaCl. The eluate was loaded directly onto a Poros MC column (0.75 x 10 cm; PerSeptive Biosystems, Cambridge, MA) equilibrated with $NiSO_4$ and washed with Buffer A. The Poros MC column was washed with Buffer A containing 500 mM NaCl and 5 mM imidazole, yeast TFIIS was eluted with Buffer A containing 500 mM NaCl and 500 mM imidazole, and immediately dialysed against Buffer A containing 150 mM NaCl and 10 mM DTT. The dialysed sample was incubated with a 500:1 mass ratio of sample to bovine thrombin for one hour at room temperature to remove the amino terminal hexahistidine tag. Digestion with thrombin was terminated by addition of 1 mM PMSF and stored on ice. Following thrombin digestion the sample was adjusted to 50 mM NaCl and loaded at 1 ml / min onto a MonoS cation exchange column (Pharmacia, Baie D'Urfe, Quebec) and eluted at 1 ml / min with a 10 ml gradient containing a maximum of 500 mM NaCl. The fractions containing TFIIS with the hexahistidine tag removed were dialysed in Buffer A containing 50 mM NaCl and 10 mM DTT and stored at -70 °C. For NMR studies, isotopically labelled TFIIS was extensively dialysed against 5 mM potassium phosphate, 10 mM DTT and 10 μM $ZnSO_4$ at pH 6.5. The protein was then concentrated via ultra-filtration to a final protein concentration of 2 mM, as determined by the spectroscopic absorbance at 280 nm.

A second strategy for the purification of TFIIS and derivatives, expressed with an amino terminal hexahistidine fusion in *E. coli* BL21(DE3) cells, was used in the

preparation of TFIIS for analysis in biochemical assays. Cells containing TFIIS or derivatives were sonicated in Buffer A containing 500 mM NaCl and 5 mM imidazole, clarified by centrifugation for 15 minutes in a microcentrifuge at 4 °C and purified using one millilitre of His Bind resin according to the manufacturer's directions (Novagen, WI). The eluate from the His Bind resin was dialysed in Buffer A containing 50 mM NaCl and 10 mM DTT and stored at -70 °C.

Domain Mapping of Yeast TFIIS

The domain structure of TFIIS was investigated by using limited proteolytic digestion to identify stably folded, protease resistant domains. For analytical purposes, 5 μg of purified yeast TFIIS poly-histidine fusion protein was incubated with 10 ng of Endoproteinase Glu-C from Staphylococcus V8 (V8 protease; Boeringer Manheim), chymotrypsin (Sigma), pronase (Boeringer Manheim) or trypsin (Sigma) for up to 60 minutes in protease buffer (20 mM Hepes, pH 7.5, 150 mM NaCl, 5 mM dithiothreitol, 10 μM $ZnSO_4$ and 10% glycerol). The proteolytic products were resolved by SDS-PAGE electrophoresis and visualized by staining with Coomassie blue. For preparative purposes, 5 mg of purified yeast TFIIS fusion protein was incubated with 10 μg of V8 protease for 30 minutes in protease buffer. The proteolytic products were resolved by metal chelate chromatography and purified by reverse phase chromatography on a C4 column (Vydac, Mississauga, ON) developed with an acetonitrile gradient. The molecular weights of the purified products were determined

by mass spectroscopy.

Transcript elongation assays: read through and nucleolytic cleavage by arrested ternary complexes

The ability of TFIIS to stimulate arrested ternary complexes to transcribe through intrinsic arrest sites was analyzed by an *in vitro* transcript elongation assay. These procedures were carried out as previously described (Christie et al., 1994). Briefly, transcription by purified RNA polymerase II was initiated from a 3' deoxycytidine tailed template containing a human histone H3.3 gene fragment, which contains well characterized blocks to elongation. The RNA transcript was pulse-labelled at the 5' end by incubating in the presence of α^{32}P-CTP (3000 Ci/mmol, ICN), 0.8 mM GTP, ATP, and UTP for 75 seconds, followed by transcription to arrest sites in the presence of 0.1 mM unlabelled CTP for 75 seconds. Ternary complexes stalled at these sites were treated in two different ways. For the read through assay, yeast TFIIS was added and transcription was allowed to proceed for the designated intervals. The resulting transcripts were collected by ethanol precipitation and resolved by electrophoresis on a 6% polyacrylamide (19:1 acrylamide:bisacrylamide), 8.3 M urea, TBE gel. The transcripts were quantified by phosphorimaging (Molecular Dynamics, Sunnyvale, CA).

For the analysis of TFIIS stimulated cleavage of the nascent RNA transcript, the stalled ternary complexes were purified from unincorporated nucleotides by two

sequential Bio-gel 30 spin columns (BioRad, Mississauga, ON). TFIIS was then added to the purified complexes and incubation was continued at 30 °C. At designated time points, aliquots were removed and the reaction stopped by addition of stopping buffer. The transcripts were resolved and quantified as described above.

To analyze the intrinsic cleavage activity of RNA polymerase II, stalled ternary complexes were purified as described for cleavage analysis and adjusted to pH 9.5 by addition to buffer containing 50 mM CAPS, 100 mM ammonium sulphate, 5 mM $MgSO_4$, 6 mM spermidine, 0.1 mg/ml heparin, 5 mM DTT, and 5% glycerol. The reaction was allowed to proceed at 30 °C for the designated intervals, aliquots were removed, and the products resolved as described for the cleavage assay.

Binding assay between Yeast TFIIS and RNA polymerase II

The interaction between TFIIS and RNA polymerase II was investigated by a native gel mobility shift assay. TFIIS and derivatives were expressed as a fusion protein contained an hexameric histidine repeat followed by a thrombin site, the heart myosin kinase (HMK) consensus phosphorylation sequence, and the coding region of TFIIS. The fusion protein was constructed with an oligonucleotide encoding the HMK consensus phosphorylation sequence (RRASVDF) inserted into the Nde1 site of the plasmid pET 15b, generating pET 15bHMK. TFIIS was inserted into the Nde1 and BamH1 sites (5' and 3', respectively) and the resulting TFIIS fusion protein was expressed and purified as described above. The protein was phosphorylated at the

HMK consensus site using HMK (Sigma, St. Louis, USA) and γ ^{32}P-ATP according to the manufacturer's instructions. One nanomole of TFIIS was incubated with heart myosin kinase, five nanomoles of ATP, and 5 picomoles of γ ^{32}P-ATP at 30 °C for one hour. The labelled protein was dialysed against Buffer A containing 10 mM DTT and 50 mM NaCl to remove the free radiolabel. The ^{32}P-labelled TFIIS and unphosphorylated TFIIS were equally active in assays of read through and transcript cleavage. For the binding assay, 0.85 picomoles of 32P-HMK-TFIIS was incubated in a 5 µl volume for a minimum of 20 minutes with 0.1-4.5 pmol of RNA polymerase II in buffer containing 20 mM HEPES pH 7.5, 10% glycerol, 10 µM $ZnSO_4$, and 10 mM DTT. The binary TFIIS-RNA polymerase II complex was resolved from free yeast TFIIS by electrophoresis for 2 to 3 hours at 100 volts at 4 °C on a 5% polyacrylamide (30:0.8 acrylamide:bisacrylamide) gel containing 50 mM Tris, 50 mM borate, 10 µM $ZnSO_4$, and 1% glycerol at pH 8.3. The amount of bound ^{32}P-HMK-TFIIS was quantified by phosphorimaging.

Site-directed mutagenesis of TFIIS

TFIIS mutants were constructed by PCR amplification of yeast TFIIS using oligonucleotides encoding the desired mutation. For each mutant constructed, two PCR products were produced, one containing a portion of the wild type sequence and the other containing the remainder of the TFIIS sequence harbouring mutations introduced with the mutagenic oligonucleotide. These PCR fragments were

phosphorylated to introduce a 5' phosphate, and subsequently ligated using T4 ligase. The ligated products were used as the template in a second PCR reaction which produced the coding region for amino acid residues 131-309, containing the site-directed mutations, as well as NdeI and BamHI restriction sites at the 5' and 3' termini, respectively. The PCR products were digested with these restriction enzymes and ligated into pET 15bHMK, containing the heart myosin kinase consensus sequence. All derived plasmids were sequenced to verify the DNA sequence and ensure the introduction of mutations.

CHAPTER 3

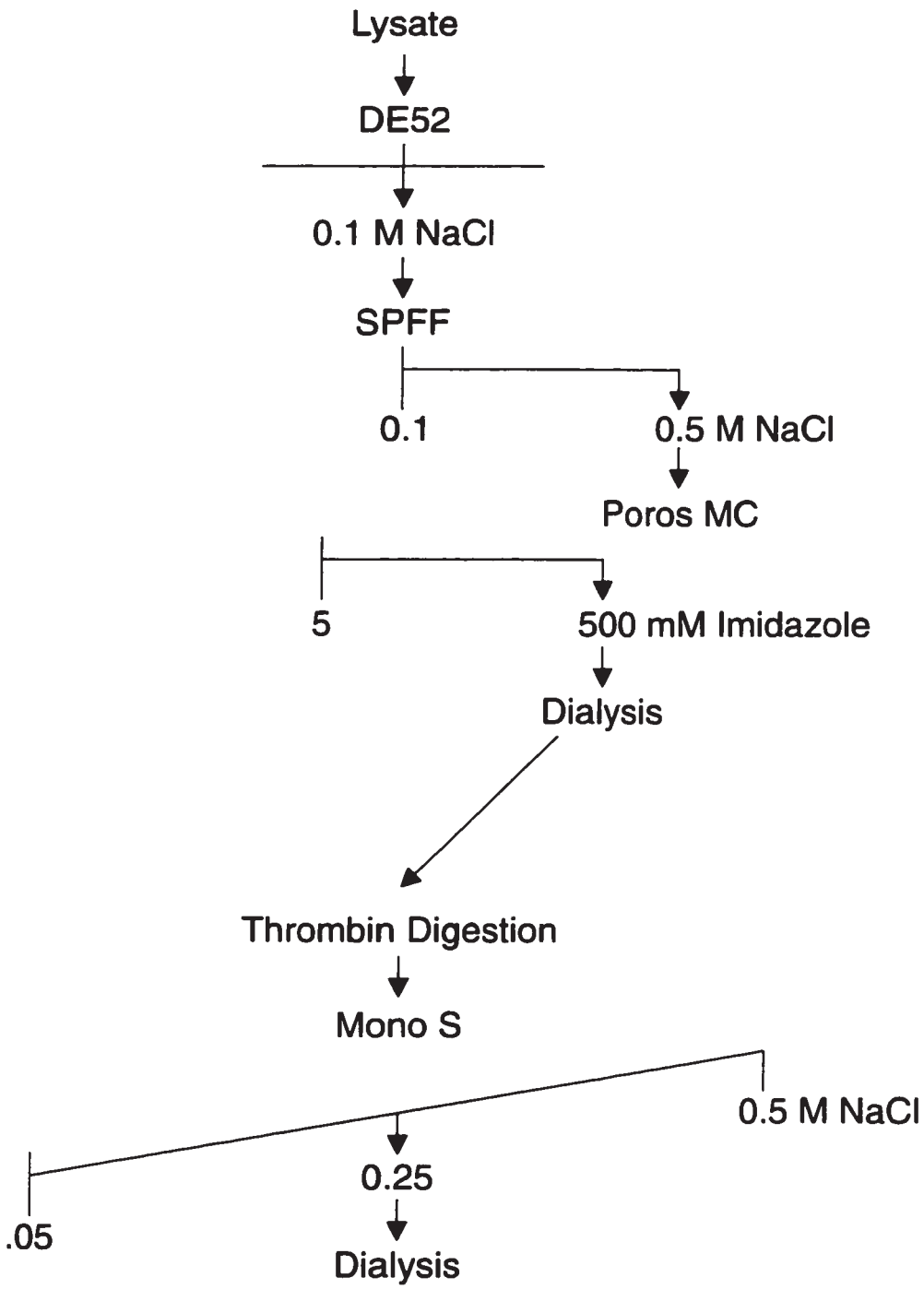
Results

I. Purification of TFIIS

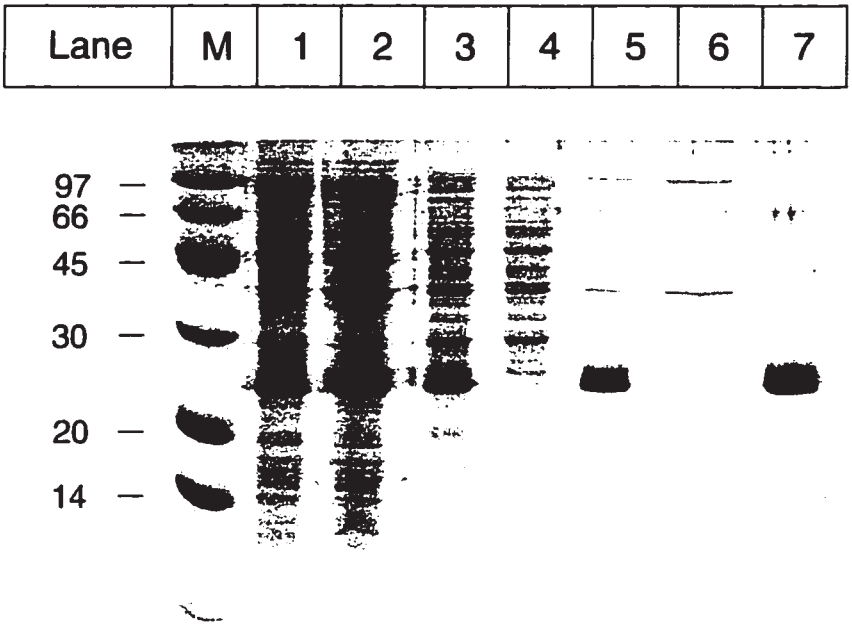
TFIIS was expressed in *E. coli* cells and purified as described in Figure 1A. The typical yield of purified wild type TFIIS containing an amino terminal hexahistidine fusion was 35 to 40 milligrams of protein per litre of cell culture. This purification strategy, when applied to TFIIS derivatives containing the amino terminal hexahistidine fusion (his tagged), resulted in TFIIS greater than 95% pure as determined by SDS-PAGE gels visualized with either Coomassie Brilliant Blue (Sigma,) (Figure 1B, lane 12) or silver stain (data not shown). TFIIS derivatives purified using His Bind resin were not as homogenous, containing approximately 85% TFIIS by mass (data not shown).

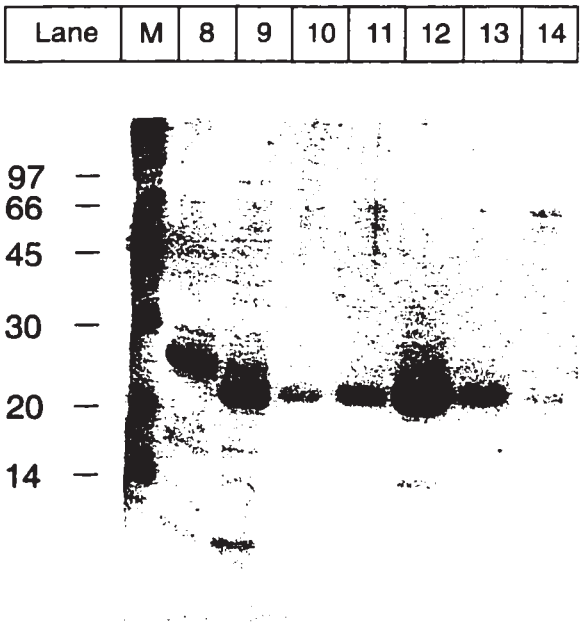
Figure 1: Purification of TFIIS$_{131-309}$. A) A clarified *E. coli* extract was purified by chromatographic techniques as described in Materials and Methods. Abbreviations used to describe column resins are defined in the text. B) Coomassie Brilliant Blue stained SDS-PAGE gel of column fractions during purification. The molecular weight of protein markers contained in Lane M are to the left. Contents of each lane correspond to column fractions obtained from the purification strategy outlined in panel A. Lane 1, clarified *E. coli* extract; lane 2, cell debris; lane 3, DE52 flow through; lane 4, SPFF flow through; lane 5, SPFF eluate; lane 6, metal chelate flow through; lanes 7 and 8, eluate from metal chelate column; lane 9, bovine thrombin digestion of dialysed metal chelate eluate; lanes 10-14, successive column fractions collected from MonoS chromatography.

A


Lysate
DE52
0.1 M NaCl
SPFF
0.1
0.5 M NaCl
Poros MC
5
500 mM Imidazole
Dialysis
Thrombin Digestion
Mono S
0.5 M NaCl
0.25
.05
Dialysis

B


Lane
M
1
2
3
4
5
6
7
97
66
45
30
20
14


Lane
M
8
9
10
11
12
13
14
97
66
45
30
20
14

II. Identification of the smallest active fragment and domain organization

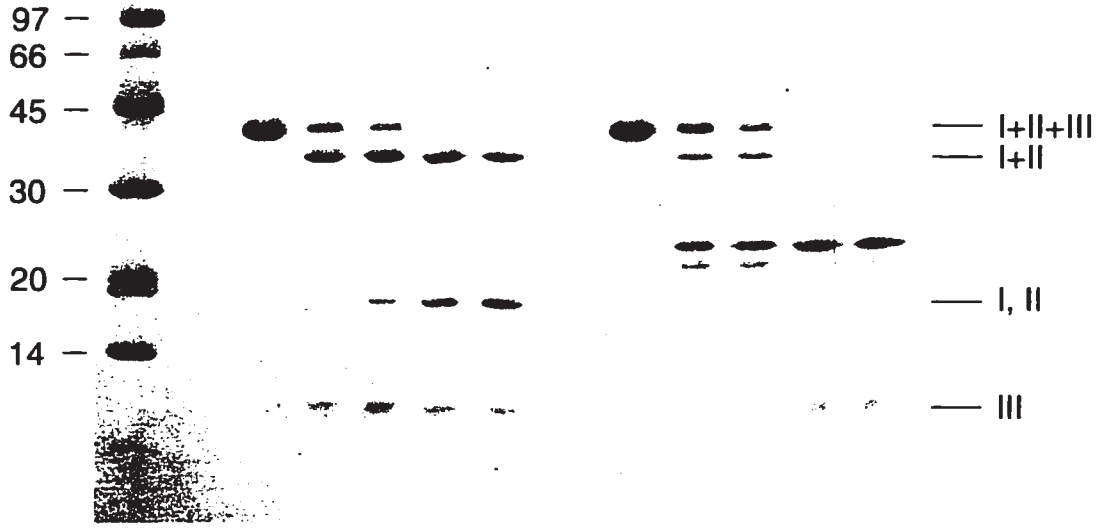
The initial structural characterization of TFIIS involved the identification of the smallest transcriptionally active fragment of TFIIS. Toward this goal, histidine tagged TFIIS was partially digested with either trypsin, chymotrypsin, pronase, or endoproteinase Glu-C from *Staphylococcus aureus* (V8 protease). The limited digestion pattern for V8 protease and chymotrypsin identified several proteolytically sensitive residues (Figure 2A). These residues are postulated to be sensitive to protease digestion due to increased accessibility, and are unlikely to be folded within a domain. A protein domain is an independently folded region of a protein and individual domains can be responsible for one or more functional attributes. The characterization by mass spectroscopy of the partial proteolytic digest indicated that TFIIS is composed of three domains. The V8 protease sensitive amide bonds were identified at the carboxyl side of residues E105, D124, and E246. The proteolytic digestion at residue D124 was incomplete compared to digestion at residues E105 and E246 indicating that this residue may be less solvent accessible. The three domains identified by this approach were domain I (amino acids 1-105), domain II (amino acids 106-246), and domain III (amino acids 247-309). This domain structure correlates with regions of TFIIS containing sequence homology with TFIIS homologues and other known proteins (Figure 2B). Domain I of yeast TFIIS, which extends from the amino terminus to the region 105-124, corresponds to a homologous region of human TFIIS which in turn is 29% identical and 53% similar to the amino terminus of rat

elongin A (Aso et al., 1995). Domain II extends from the 105-124 region to residue 246. The amino acid sequence of both domains II and III is conserved among TFIIS homologues and both domains are essential for activity.

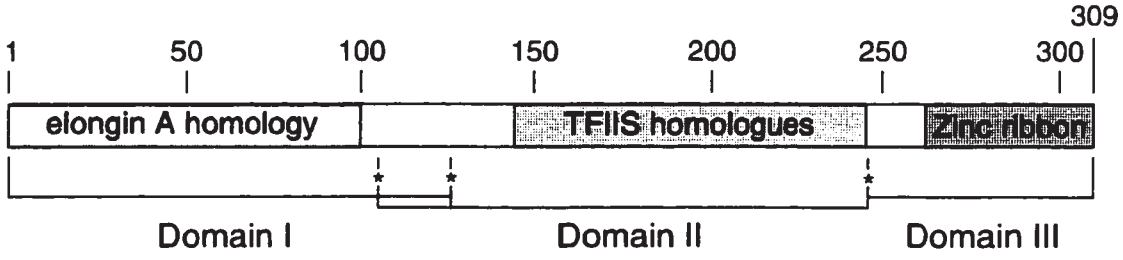
Figure 2: Limited proteolytic digestion of TFIIS. A) $TFIIS_{1-309}$ containing the amino terminal histidine fusion was digested with V8 protease from *Staphylococcus aureus* and chymotrypsin for the indicated times. Proteolytic products were resolved on a 15% SDS-PAGE gel and stained with Coomassie Brilliant Blue, and identified by mass spectroscopy. The proteolytic products containing domains of TFIIS are indicated to the right. B) Schematic representation of the domain structure of TFIIS. Regions of homology to TFIIS homologues and other transcription factors are indicated as shaded boxes. Asterisks indicate the identified proteolytically sensitive amide bonds.

A

	Staph V8						Chymotrypsin					Protease
M		0	5	15	30	60	0	5	15	30	60	Time (min)



B



Deletion mapping of contiguous residues was performed to further define the minimal amino acid sequence required for TFIIS activities *in vitro* (Table 1). Deletion mutants lacking the amino terminal 143 amino acids are fully active in biochemical assays; however, removal of an additional 5 amino acids resulted in functionally inactive protein. The ratio of soluble : insoluble TFIIS derivatives in *E. coli* lysates, an indicator of proper protein folding, was not affected by removal of the first 131 amino acids and reduced upon removal of the first 143 amino terminal residues, indicating that residues between 130 and 143 are required for proper protein folding. An amino terminal deletion of the first 263 amino acids resulted in the expression of soluble protein containing the zinc ribbon motif. The construction of deletion mutants lacking both the amino terminal 130 amino acids as well as carboxyl terminal residues resulted in the identification of two soluble mutants contained within domain II, $TFIIS_{131\text{-}240}$ and $TFIIS_{131\text{-}263}$. The proteolytic mapping coupled with deletion analysis identified the smallest active fragment of TFIIS as 144-309, however NMR spectroscopic analysis revealed that this deletion mutant contained regions of conformational heterogeneity, indicating possible misfolding. NMR analysis of the deletion mutants $TFIIS_{131\text{-}309}$ and $TFIIS_{131\text{-}240}$ indicated that these mutants were monomeric and conformationally homogeneous.

Table 1: Characterization of deletion mutants

Deletion mutant	Soluble Expression	Read through activity
113-309	+++	+++
131-309	+++	+++
138-309	++	+++
144-309	+	+++
149-309	+	-
182-309	+	-
186-309	+	-
222-309	++	-
264-309	+++	-
114-224	+	-
114-240	+	-
131-224	+	-
131-240	+++	-
131-263	+++	-
144-224	-	-
144-240	-	-
131-309 (ins. @ 240:LKQKI)	+++	-
131-309 (del. @ 240-245:LKQKI)	+++	-

Legend: The solubility of TFIIS expression is a ratio of soluble : insoluble TFIIS in *E. coli* lysates: -, insoluble; +, 10-30% solublility; ++, 30-60% solublility; +++, 60-100% solublility. The read through activity of mutants is relative to wild type TFIIS activity: -, inactive; +, 10-30% activity; ++, 30-60% activity; +++, 60-100% activity.

III. TFIIS activity requires domain II and III

Yeast TFIIS activates transcription through the histone H3.3 TIa intrinsic arrest site (Figure 3, and (Christie et al., 1994)). I discovered that the deletion mutant 131-309 also stimulates transcription through arrest sites (Figure 4) at a 5:1 molar ratio of TFIIS:RNA polymerase II. Thus stimulation of arrested ternary complexes through intrinsic arrest sites within the histone H3.3 intron requires both domain II and domain III of TFIIS. Addition of increasing quantities of TFIIS slightly increases the rate of transcription through arrest sites but does not alter the amount of transcript remaining at the TIa arrest site after 30 minutes (compare Figure 4, lanes containing 5:1 with 100:1 and 500:1 ratios of TFIIS:RNA polymerase). Transcripts remaining at TIa after stimulation by TFIIS may be the product of premature termination events.

Both wild type TFIIS (Figure 5) and the 131-309 deletion mutant stimulate transcript cleavage by the ternary complex at a 5:1 molar ratio of TFIIS : RNA polymerase II (Figure 6). The production of two distinct cleavage products (C1 and C2) was observed at the TIa arrest site as was shown previously for mammalian ternary complexes (Reines, 1992). The 5' fragments of the TFIIS-shortened transcripts are capable of being elongated upon addition of nucleotides. Addition of increasing amounts of TFIIS stimulates the production of shortened transcripts at a higher rate and also resulted in the production of many shorter transcripts (Figure 6, 100:1 and 500:1 molar ratio of TFIIS : RNA polymerase II). Stimulation of extensive transcript shortening by high molar ratios of TFIIS : RNA polymerase also resulted in the

formation of transcripts that could not be extended upon addition of nucleotides (compare Figure 6, lanes containing 5:1 with 100:1 and 500:1 ratio of TFIIS:RNA polymerase II). The production of the C1 shortened transcript at 5:1 molar ratios of TFIIS:RNA polymerase II occurred at a more rapid rate than the production of full length transcripts in the read through assay, however the rate of production of C2 did correlate with the rate of progression through the TIa arrest site (compare TFIIS: RNA polymerase II at 5:1 molar ratios in Figures 4A and 6).

IV. Binding of TFIIS to RNA polymerase II

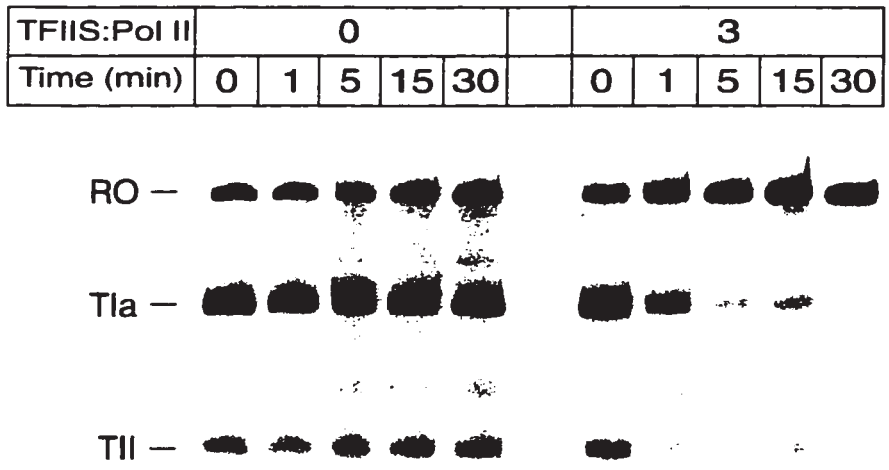
TFIIS fusion proteins containing the amino terminal heart myosin kinase consensus phosphorylation sequence were radiolabelled to an average specific activity of 4.47 ± 0.87 x 10^6 cpm/nmol with phosphorylation of 67.1 ± 13.0 % of TFIIS molecules. The deletion mutant, $TFIIS_{131-309}$, binds to RNA polymerase II with an affinity of 58 nM as determined by native gel mobility shift assays. Radiolabelled deletion mutant $TFIIS_{131-309}$ interacts with RNA polymerase II in a saturatable and stoichiometric manner (Figure 7) and the interaction is both reversible and salt sensitive (data not shown). Although both domain II and domain III are required for TFIIS stimulatory activity, the region responsible for interaction with RNA polymerase II was previously unknown. A series of deletion mutants were analyzed in the binding assay to determine the region of TFIIS responsible for interacting with RNA polymerase II. The minimal region required for interaction with RNA polymerase II is

contained within $TFIIS_{131-240}$ (Figure 8A), which interacted with a similar affinity (80 nM) as did $TFIIS_{131-309}$ (Figure 8B). Analysis of ^{32}P labelled $TFIIS_{1-309}$ was not undertaken because the full length TFIIS was retained in the wells of the native gel. Removal of the amino terminal 130 amino acids improved the electrophoretic properties and enabled the quantitative analysis of the interaction with RNA polymerase II.

Using limited proteolytic digestion and deletion mutants coupled with transcript elongation assays and protein-protein interaction analysis, I have determined that the smallest active fragment contains amino acids 143-309, and residues 131-240 are required for the TFIIS-RNA polymerase II interaction. The available data suggests that $TFIIS_{143-240}$ contains the RNA polymerase II binding site, however the analysis of this fragment was not performed because this particular fragment was insoluble (Table 1).

Figure 3: Read through assay with wild type TFIIS. A) Arrested ternary complexes were incubated for the designated times in either the absence or presence of a 3 fold molar ratio of TFIIS : RNA polymerase II and resolved as described in the Materials and Methods. Transcript products identified to the right of the autoradiogram are; RO, run off product; TIa, major arrest site; and TII, minor arrest site. B) Graphical representation of phosphorimager quantified transcript products at TIa and run off. % at TIa is as described in the text (TIa/(TIa+RO) x 100%). Filled squares represent no TFIIS added and filled diamonds represent products of 3:1 molar ratio of TFIIS:RNA polymerase II.

A


TFIIS:Pol II
0
3
Time (min)
0 1 5 15 30
0 1 5 15 30
RO
Tla
TII

B

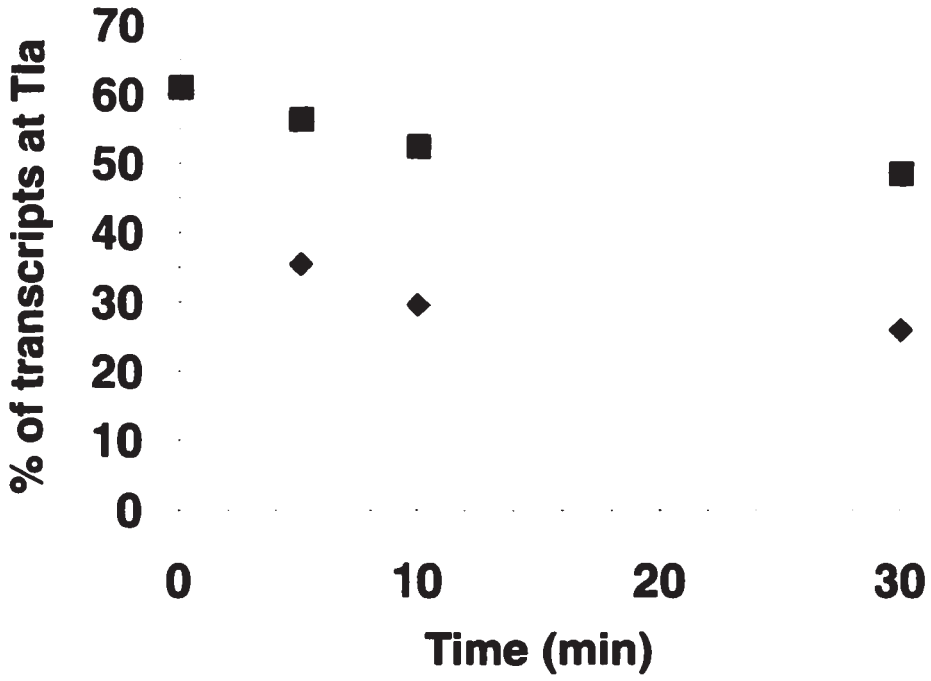
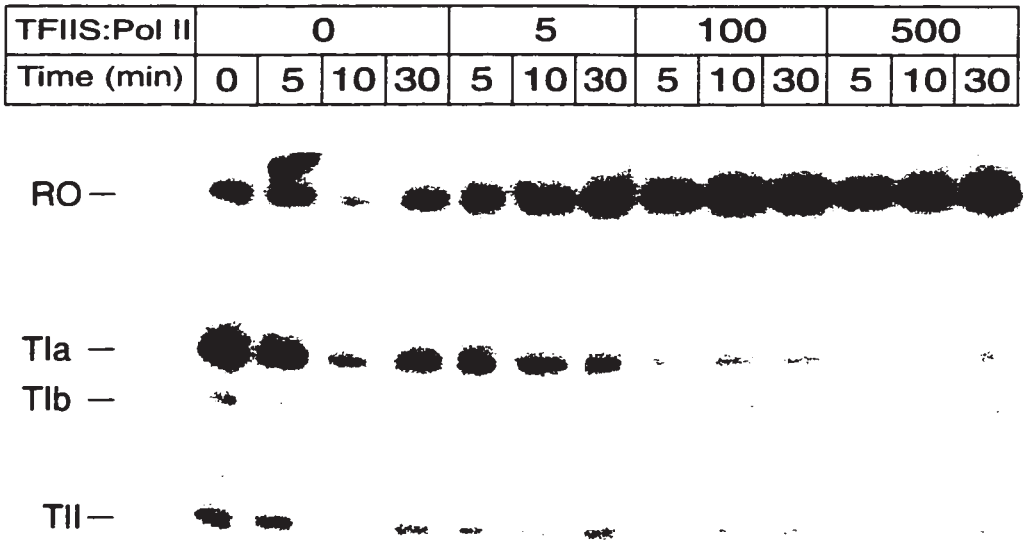

% of transcripts at Tla
70
60
50
40
30
20
10
0
0
10
20
30
Time (min)

Figure 4: Read through assay with $TFIIS_{131-309}$. A) Arrested ternary complexes were incubated for the designated times in the either the absence or presence of the indicated molar ratios of $TFIIS_{131-309}$: RNA polymerase II and resolved as described in the Materials and Methods. Transcript products identified to the right of the autoradiogram are; RO, run off product; TIa, major arrest site; TIb, minor arrest site; and TII, minor arrest site. B) Graphical representation of phosphorimager quantified transcript products at TIa and run off. % at TIa is as described in the text (TIa/(TIa+RO) x 100%). Diamonds represent no TFIIS added; filled squares represent products of 5:1 molar ratio of $TFIIS_{131-309}$:RNA polymerase II; filled triangles represent products of 100:1 molar ratio; and x represent the products of 500:1 molar ratio.

A


TFIIS:Pol II
0
5
100
500
Time (min)
0 5 10 30 5 10 30 5 10 30 5 10 30
RO
TIa
TIb
TII

B

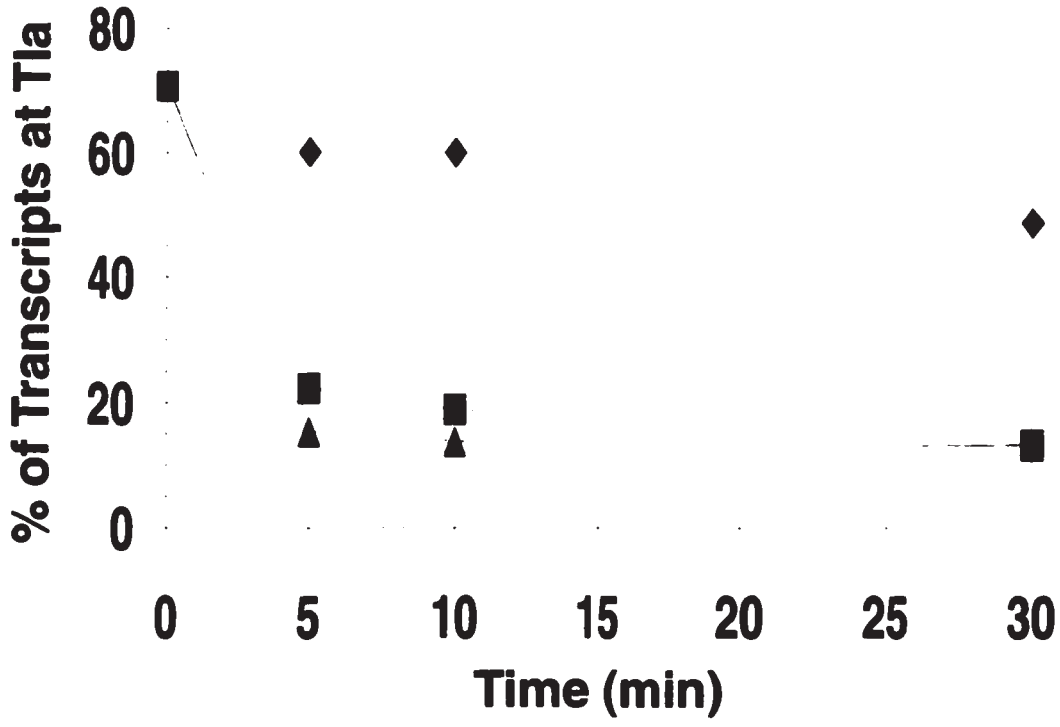

% of Transcripts at TIa
80
60
40
20
0
0 5 10 15 20 25 30
Time (min)

Figure 5: Cleavage assay with wild type TFIIS. Transcript cleavage products were obtained and resolved as described in Materials and Methods and visualized by autoradiography. Arrested ternary complexes were stimulated with a 3 fold molar excess of TFIIS : RNA polymerase II for the designated times. In the indicated lanes nucleotides were added and incubation allowed to proceed for an additional 10 minutes. Transcript elongation products are identified at the left as in Figure 3 with the addition of two cleavage products C1, and C2.

TFIIS:Pol II	0		3				
Time (min)	0	30	1	5	15	30	30
Nucleotides	-		-				+

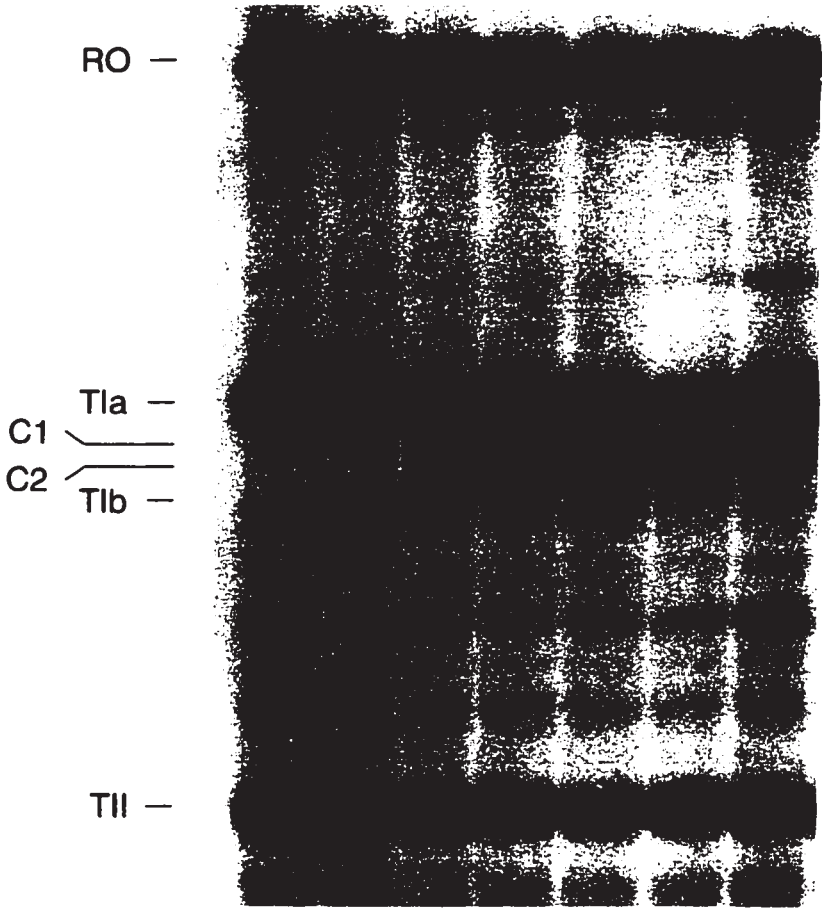


Figure 6: Cleavage assay with $TFIIS_{131-309}$. Transcript cleavage products were obtained and resolved as described in Materials and Methods and visualized by autoradiography. Arrested ternary complexes were stimulated with $TFIIS_{131-309}$ for the designated times with the indicated amounts of $TFIIS_{131-309}$. Nucleotides were added in the indicated lanes and incubation allowed to proceed for an additional 10 minutes. Transcript elongation products are identified at the left as in Figure 5.

TFIIS:Pol II	0					5				100				500			
Time (min)	0	1	5	10	10	1	5	10	10	1	5	10	10	1	5	10	10
Nucleotides	-				+	-			+	-			+	-			+

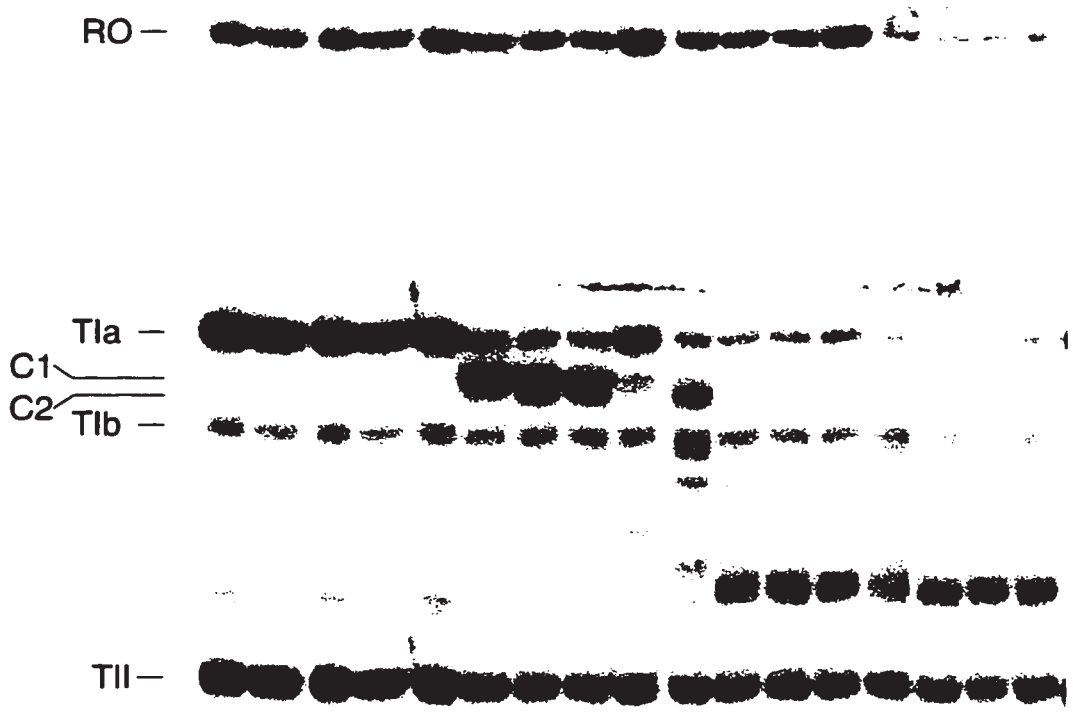
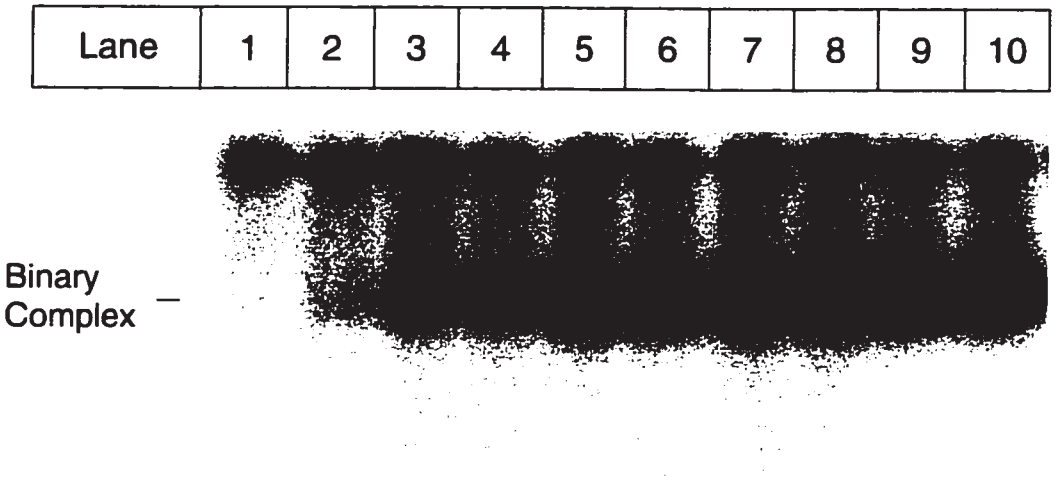


Figure 7: Binding assay between $TFIIS_{131-309}$ and RNA polymerase II. A) Autoradiogram of a gel mobility shift assay between radiolabelled $TFIIS_{131-309}$ and RNA polymerase II. The binary complex of $TFIIS_{131-309}$ and RNA polymerase II is identified to the left; uncomplexed $TFIIS_{131-309}$ migrates towards the cathode, out of the gel wells. Lanes 1-10 contain 0, 0.08, 0.15, 0.31, 0.46, 0.69, 1.15, 1.54, 2.31, and 4.62 pmoles RNA polymerase II, respectively. All lanes contain 1 pmole ^{32}P HMK-$TFIIS_{131-309}$ B) Binding assays were quantified by phosphorimager and plotted graphically. Error bars represent the standard deviation of three binding experiments.

A


Lane
1
2
3
4
5
6
7
8
9
10
Binary
Complex

B

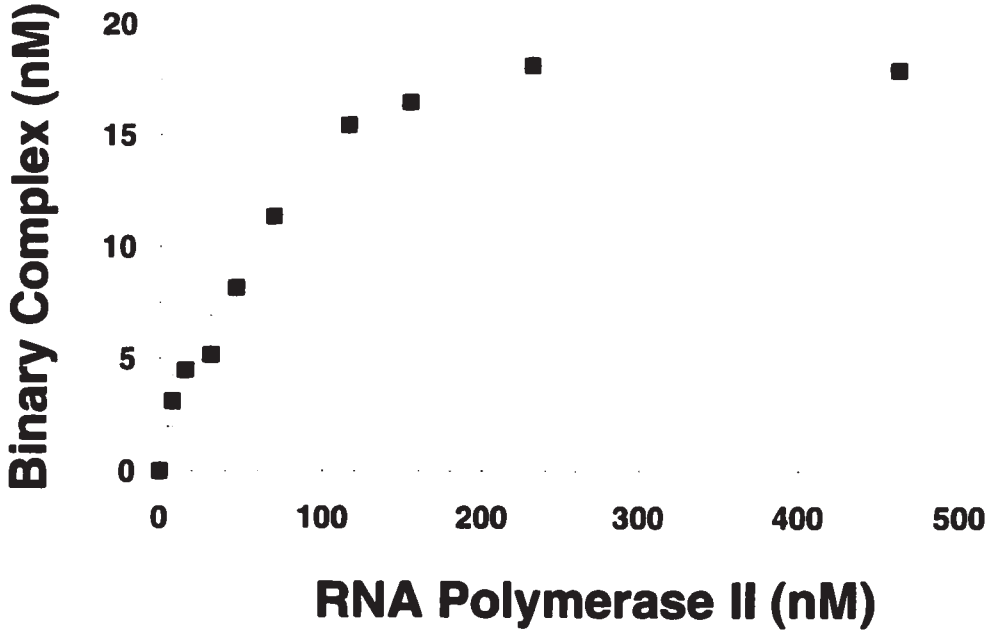
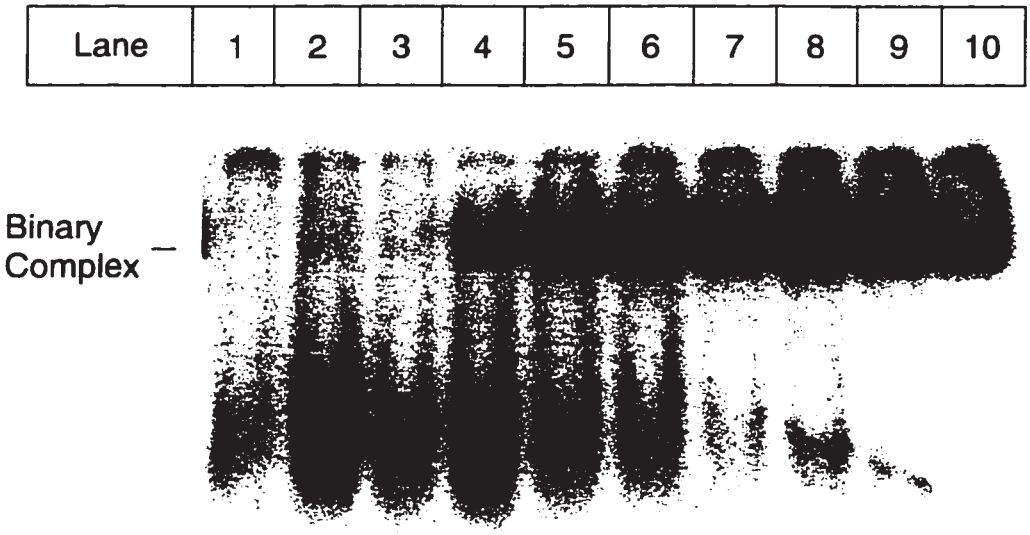
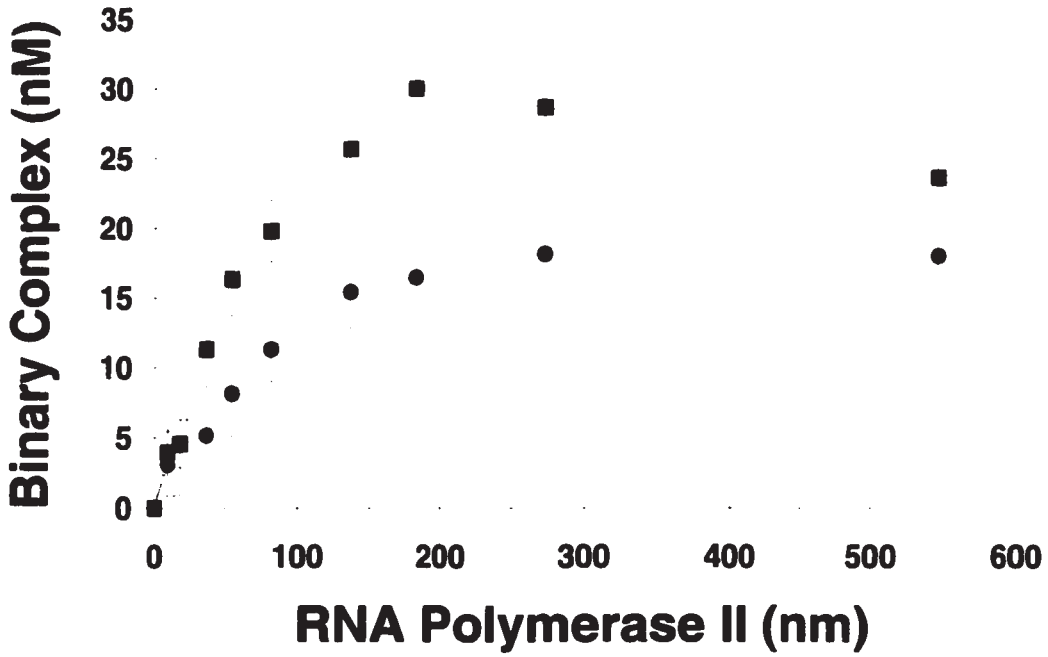

Binary Complex (nM)
20
15
10
5
0
0
100
200
300
400
500
RNA Polymerase II (nM)

Figure 8: Binding assay between $TFIIS_{131-240}$ and RNA polymerase II. A) Autoradiogram of a gel mobility shift assay between radiolabelled $TFIIS_{131-240}$ and RNA polymerase II. The binary complex of $TFIIS_{131-240}$ and RNA polymerase II is identified to the left; uncomplexed $TFIIS_{131-240}$ migrates towards the anode, into the gel, and appears as a higher mobility diffuse band. Lanes 1-10 contain 0, 0.08, 0.15, 0.31, 0.46, 0.69, 1.15, 1.54, 2.31, and 4.62 pmoles RNA polymerase II, respectively. All lanes contain 1 pmole ^{32}P HMK-$TFIIS_{131}$.240 B) Binding assays were quantified by phosphorimager and plotted graphically. Filled circles represent $TFIIS_{131-309}$ and filled squares represent $TFIIS_{131-240}$. Error bars represent the standard deviation of three binding experiments for $TFIIS_{131-309}$ and four binding experiments for $TFIIS_{131-240}$.

A


Lane
1
2
3
4
5
6
7
8
9
10
Binary Complex

B


Binary Complex (nM)
35
30
25
20
15
10
5
0
0
100
200
300
400
500
600
RNA Polymerase II (nm)

V. Analysis of site-directed mutants of TFIIS

To determine the identity of amino acid residues of TFIIS responsible for TFIIS activities, 36 site-directed mutants were constructed in the 131-309 deletion mutant, or obtained as gifts from Nell Shimasaki (K196A, R198A, and R200A). Site-directed mutants were purified and analyzed by transcript elongation assays, protein-protein interaction assays with RNA polymerase II, and NMR spectroscopy. Site-directed mutants which displayed large structural alterations, determined by comparing the NMR spectra of $TFIIS_{131-309}$ with site-directed mutants, could not be interpreted as being important for TFIIS activities and were therefore not analyzed further (Table 3). The large structural alterations resulting from mutations within a particular domain were localized and did not affect the structure of the remainder of $TFIIS_{131-309}$, indicating that each domain is an independent folding unit.

TFIIS residues important for mediating the interaction with RNA polymerase II

Mutants containing amino acid changes at Y195H:E291L, K196A, R198A, R200A, and K209A, exhibited both a decreased affinity for RNA polymerase II and a reduced stimulatory activity in transcript elongation assays, compared to wild type TFIIS (Table 2). The reduced stimulatory ability of TFIIS mutants, exemplified by mutant R198A, to promote transcript elongation through the TIa arrest site (Figure 9A) as well as induce transcript cleavage (Figure 9B), is inferred from the increased molar ratios of mutant TFIIS : RNA polymerase II required to promote these activities. The

residues involved in protein association with RNA polymerase II are therefore localized to the region 195-209. The spatial location of site-directed mutants which affect the interaction between TFIIS and RNA polymerase II will be discussed later. In summary, I have discovered that domain II is the region of TFIIS responsible for mediating the interaction with RNA polymerase II. The locations of these residues with respect to the secondary structure of TFIIS are illustrated in Figure 12.

TFIIS residues important for read through and cleavage activities

Contiguous deletions and site-directed mutagenesis of amino acids within domain III identified residues important in mediating the stimulatory activity of TFIIS (Table 1 and Table 2). Site-directed mutagenesis of several residues had no effect on the stimulation of both transcript elongation through arrest sites and stimulation of cleavage activity (Table 2; mutants K273A:K275A, and R305A). For the remainder of the mutants constructed, a range of stimulatory activity was observed in read through assays, varying from slightly reduced activity (Figure 10) to complete abrogation (Figure 11). The mutants that displayed a reduced stimulation of RNA polymerase II to progress through the TIa site in read through assays also displayed a reduced stimulatory activity in cleavage assays, without a reduction in the ability of TFIIS to interact with RNA polymerase II (Table 2; mutants K242A:Q243A, F269A, Q283A:Q285A, R287Q, E291(N, H), R287Q:E291(H, L), F296A, and K307A).

I observed that the stimulation of the ternary complex through intrinsic arrest

sites is correlated with transcript cleavage at the second cleavage site. Transcript cleavage to generate the first cleavage product, C1, was often observed without a concomitant progression through the TIa arrest site (Figure 10). Mutants completely defective in stimulating transcription through the TIa intrinsic arrest site, even at the highest TFIIS:RNA polymerase II molar ratios (Table 2, Figure 11A) stimulated transcript cleavage to C1, and not to C2 (Figure 11B). The mutagenesis of several residues within the carboxyl terminal zinc ribbon domain abrogated the ability of TFIIS to promote transcript shortening at sites upstream of C1, but no site-directed mutants completely abolished the appearance of the cleavage product C1 (Table 2). Clearly, cleavage to C1 and C2 are kinetically and structurally different processes. The location of residues affecting the stimulatory activity of TFIIS are illustrated with respect to the secondary structure in Figure 12.

Table 2: Transcript elongation and binding analysis of TFIIS mutants

Mutant	Read Through	Cleavage	Binding K_d (nM)
131-309	3	3	58.3
131-240	0	0	80
131-263	0	0	180
222-309	0	0	ND
264-309	0	0	>1500
K154A: D158A	2	2	196
K176A	2	2	34.4
N187A: C188A: D189A	2	2	56
N191A, E192A	3	2	184
K196A	2	2	>1500†
R198A	2	2	>1500†
R200A	2	1	>1500†
K209A	2	2	1284
N210A: N211A: D213A	3	2	163
N220A: D222A	2	2	97
D232A: K234A	3	2	47
D235A	3	3	ND†
K242A: Q243A	2	2	142
N252A: N255A: Q257A	1	1	19.4
T266A: D267A	2	2	16.3
F269A	1	2	ND
K273A: K275A	2	2	76.4
Q283A: Q285A	0	0	704
T286A: S288A	2	2	2.5
R287Q	1	1	174.7
E291N	0	0	79.8
E291H	0	0	50.7
E291L	0	0	1851
R287Q: E291N	0	1	292
R287Q: E291H	0	0	103
R287Q: E291H: _P292	0	0	166
R287Q: E291L	0	0	35
F296A	1	2	ND
R305A	2	2	ND†
K307A	0	2	ND
F308A	2	2	ND

Legend: 0, inactive or partially active at 500:1; 1, active at 500:1; 2, active at 100:1; 3, active at 5:1; ND, Not Determined; † Read through and cleavage were determined for the 1-309 mutant and binding for 131-309 mutant.

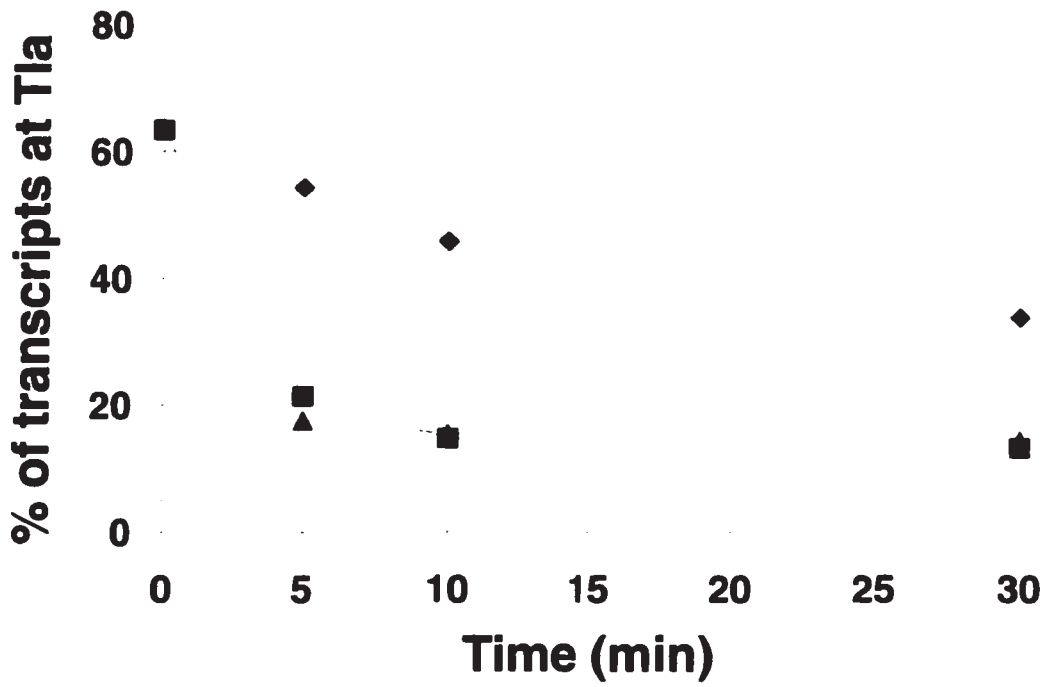
Table 3: TFIIS mutants displaying large structural alterations

Mutant	Read Through	Cleavage	Binding K_d (nM)
D150A	2	0	252
K215A: K217A	1	2	>1500
Y281A: Y282A	0	1	ND
W306A	0	0	ND

Legend: 0, inactive or partially active at 500:1; 1, active at 500:1; 2, active at 100:1; 3, active at 5:1; ND, Not Determined.

Figure 9: Analysis of TFIIS mutant R198A in transcript elongation assays. A) Graphical representation of phosphorimager quantified TIa and run off transcript products prepared and resolved as described in the Materials and Methods. % at TIa is as described in the text (TIa/(TIa+RO) x 100%). Diamonds represent products of incubation with 5:1 molar ratio; filled squares represent products of 100:1 molar ratio; and triangles represent the products of 500:1 molar ratio of $TFIIS_{131-309}$(R198A):RNA polymerase II. B) Transcript cleavage products were obtained and resolved as described in the Materials and Methods and visualized by autoradiography. Arrested ternary complexes were stimulated with the indicated amounts of $TFIIS_{131-309}$(R198A) for the designated times. Nucleotides were added in the indicated lanes and incubation allowed to proceed for an additional 10 minutes. Transcript elongation products are identified at the left as in Figure 5.

A



B

TFIIS:Pol II	5				100				500			
Time (min)	1	5	10	10	1	5	10	10	1	5	10	10
Nucleotides	-			+	-			+	-			+

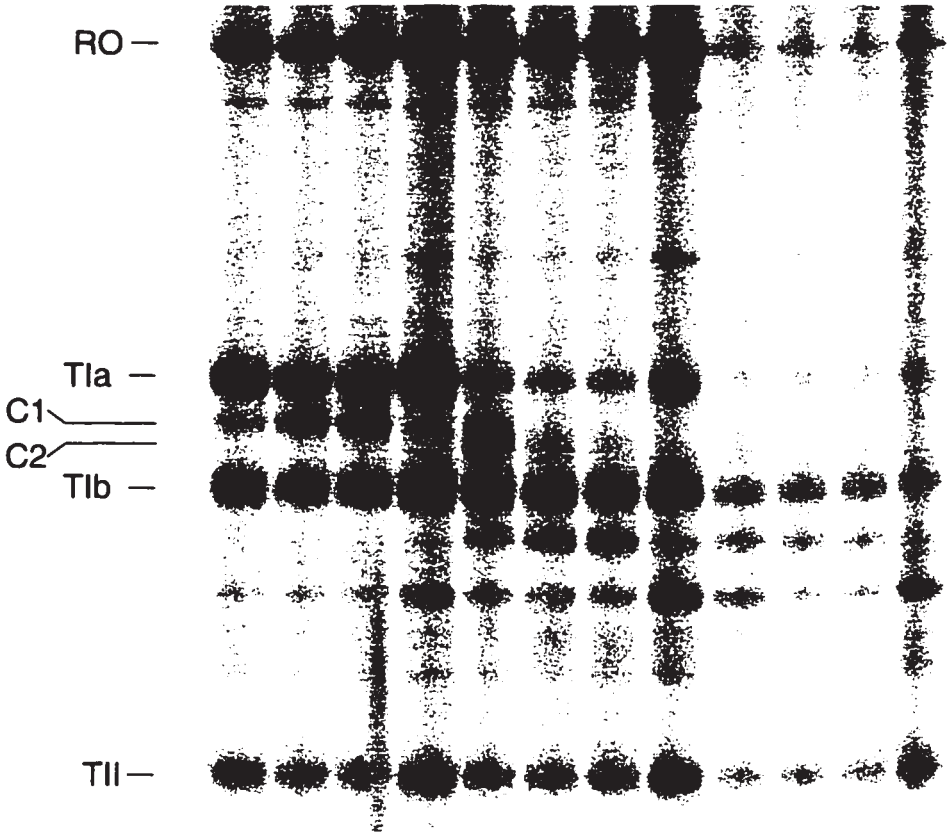
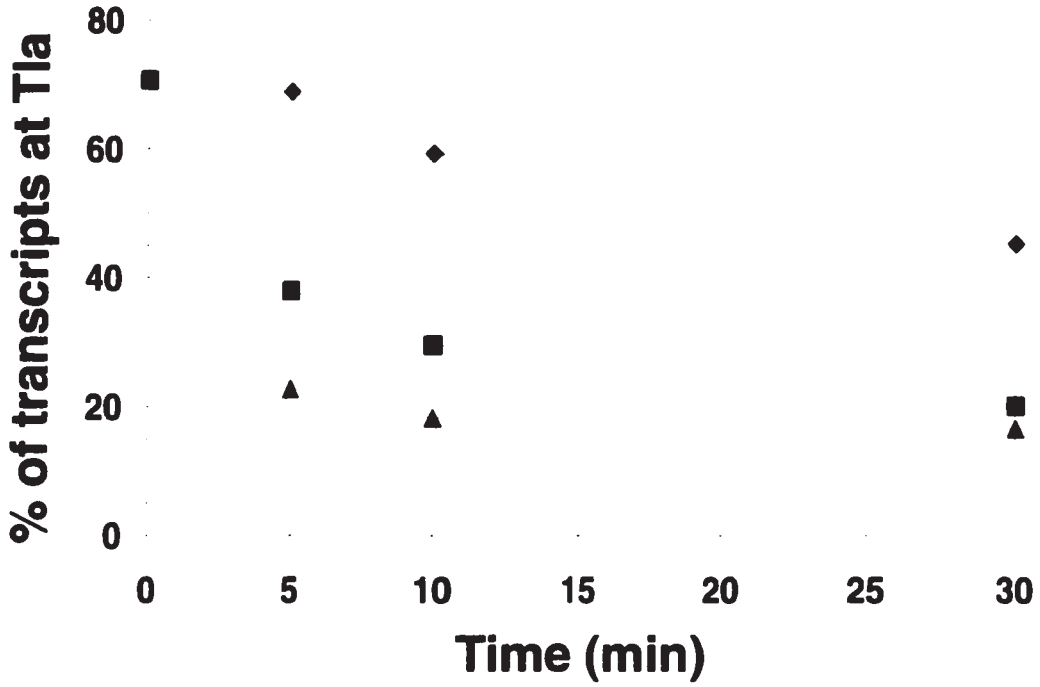


Figure 10: Analysis of TFIIS mutant F296A in transcript elongation assays. A) Graphical representation of phosphorimager quantified TIa and run off transcript products prepared and resolved as described in the Materials and Methods. % at TIa is as described in the text (TIa/(TIa+RO) x 100%). Diamonds represent products of incubation with 5:1 molar ratio; filled squares represent products of 100:1 molar ratio; and triangles represent the products of 500:1 molar ratio of $TFIIS_{131-309}$(F296A):RNA polymerase II. B) Transcript cleavage products were obtained and resolved as described in the Materials and Methods and visualized by autoradiography. Arrested ternary complexes were stimulated with the indicated amounts of $TFIIS_{131-309}$(F296A) for the designated times. Nucleotides were added in the indicated lanes and incubation allowed to proceed for an additional 10 minutes. Transcript elongation products are identified at the left as in Figure 5.

A



B

TFIIS:Pol II	5					100				500			
Time (min)	0	1	5	10	10	1	5	10	10	1	5	10	10
Nucleotides	-				+	-			+	-			+

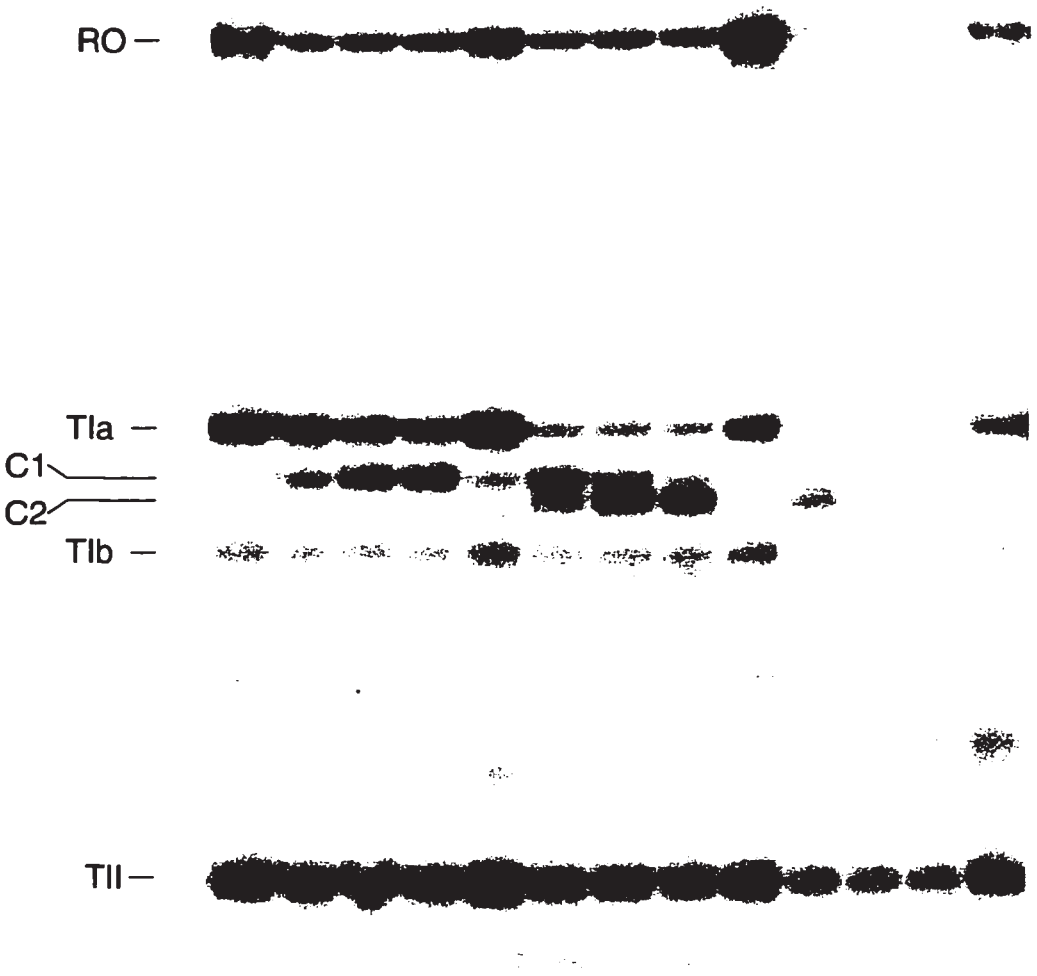
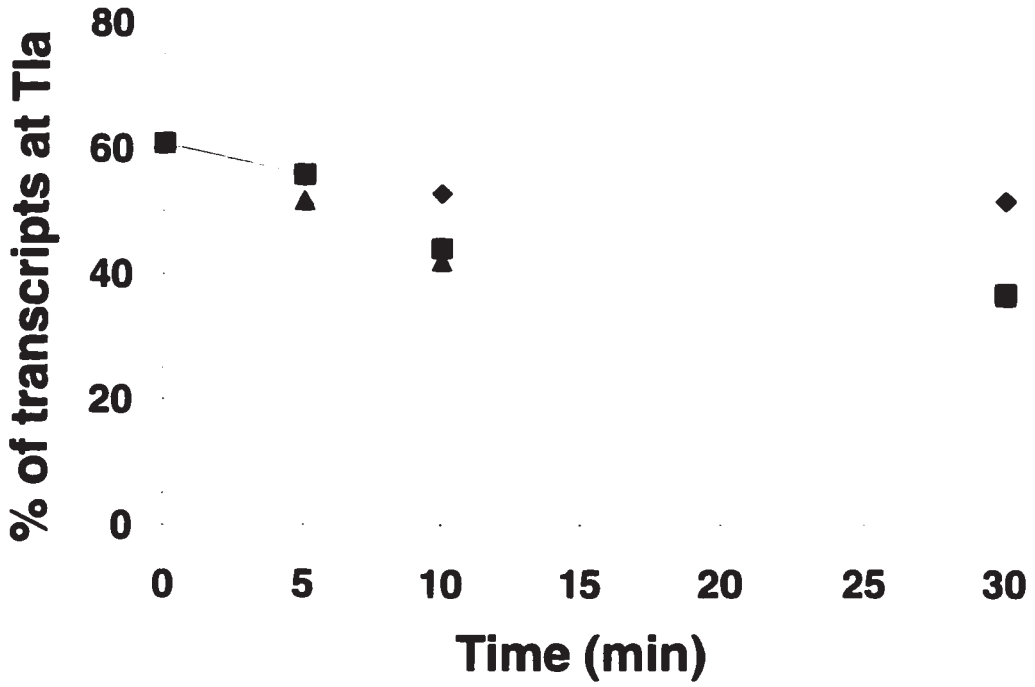


Figure 11: Analysis of TFIIS mutant Q283A:Q285A in transcript elongation assays. A) Graphical representation of phosphorimager quantified TIa and run off transcript products prepared and resolved as described in the Materials and Methods. % at TIa is as described in the text (TIa/(TIa+RO) x 100%). Diamonds represent products of incubation with 5:1 molar ratio; filled squares represent products of 100:1 molar ratio; and triangles represent the products of 500:1 molar ratio of $TFIIS_{131-309}$(Q283A:Q285A):RNA polymerase II. B) Transcript cleavage products were obtained and resolved as described in the Materials and Methods and visualized by autoradiography. Arrested ternary complexes were stimulated with the indicated amounts of $TFIIS_{131-309}$(Q283A:Q285A) for the designated times. Nucleotides were added in the indicated lanes and incubation allowed to proceed for an additional 10 minutes. Transcript elongation products are identified at the left as in Figure 5.

A



B

TFIIS:Pol II	5				100				500			
Time (min)	1	5	10	10	1	5	10	10	1	5	10	10
Nucleotides	-			+	-			+	-			+

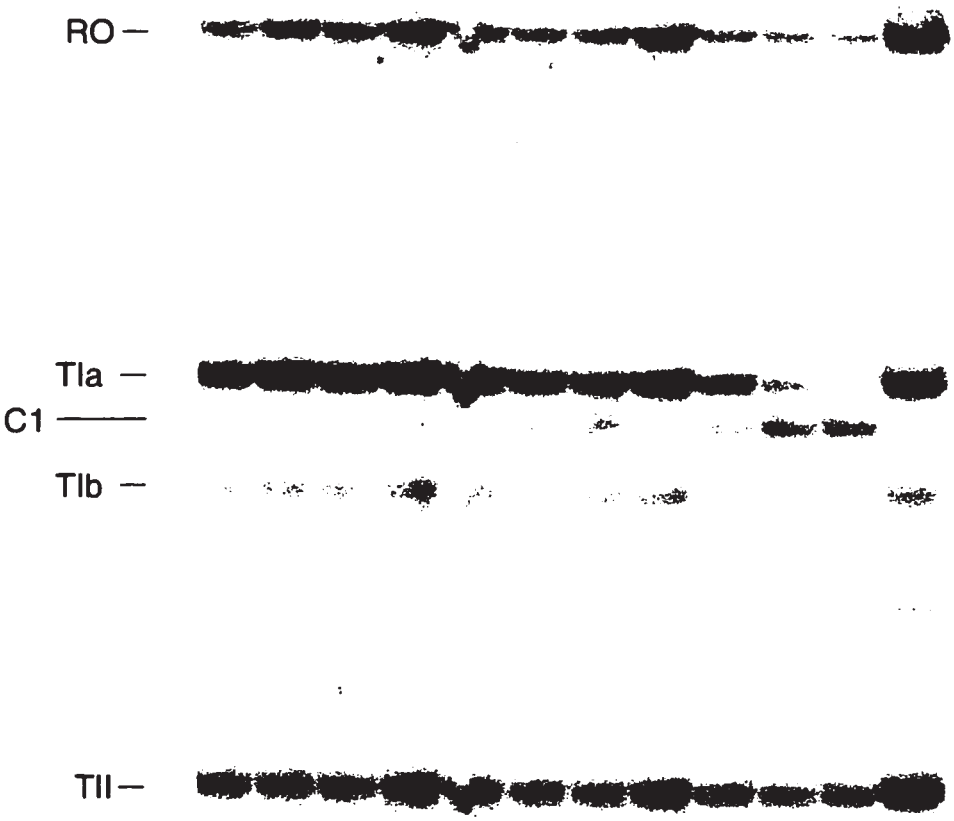
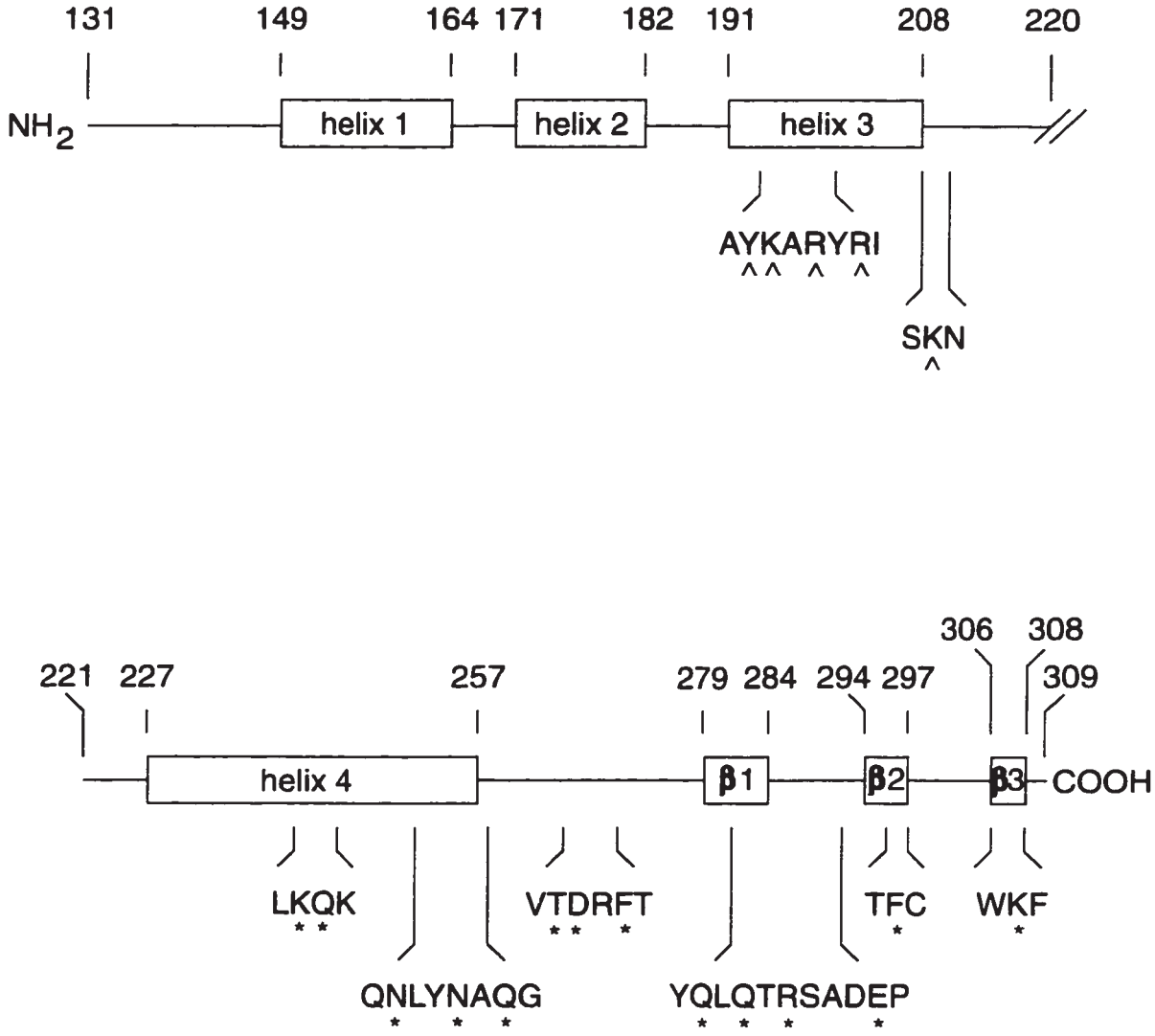


Figure 12: Schematic representation of TFIIS primary and secondary structure. A) Primary amino acid sequence of yeast TFIIS. B) Amino acid residues 131-309 are shown with the corresponding secondary structure indicated by open boxes. The residues defining the start and end of secondary structure elements are indicated. Amino acid residues identified by site-directed mutagenesis to be important for binding activity are indicated below the sequence 131-220, and labelled with ^. Residues important for stimulation activity were similarly identified and labelled with asterisks below the sequence 221-309.

A

1	MDSKEVLVHV	KNLEKNKSND	AAVLEILHVL	DKEFFPTEKL	40
41	LRETKVGVEV	NKFKKSTNVE	ISKLVKKMIS	SWKAQLIKNK	80
81	RCRQPQQHHQ	DHAPGNAEDK	TTVGESVNGV	QQPASSQSDA	120
121	MKQDKYVSTK	PRNSKNDGVD	TAIYHHKLRD	QVLKALYDVL	160
161	AKESEHPPQS	ILHTAKAIES	EMNKVNNCDT	NEAAYKARYR	200
201	IIYSNVISKN	NPDLKHKIAN	GDITPEFLAT	CDAKDLAPAP	240
241	LKQKIEEIAK	QNLYNAQGAT	IERSVTDRFT	CGKCKEKKVS	280
281	YYQLQTRSAD	EPLTTFCTCE	ACGNRWKFS		309

B


131
149
164
171
182
191
208
220
NH2
helix 1
helix 2
helix 3
AYKARYRI
SKN
221
227
257
279
284
294
297
306
308
309
helix 4
β1
β2
β3
COOH
LKQK
VTDRFT
TFC
WKF
QNLYNAQG
YQLQTRSADEP

VI. Structure of Domain II and Domains II + III

The purpose of the TFIIS structure determination is to provide a framework for biochemical and genetic studies of the stimulation of transcript elongation by TFIIS. Attempts at crystallization of the full length protein (TFIIS $_{1\text{-}309}$) or the smallest active protein (TFIIS $_{143\text{-}309}$) were unsuccessful, and since the entire protein is too large for NMR structure determination, I set out to identify structural domains that might be amenable for NMR structure determination. Therefore, efforts initially focused on identifying a fragment of TFIIS that harboured elongation activity and displayed the requisite stability and solubility characteristics for NMR.

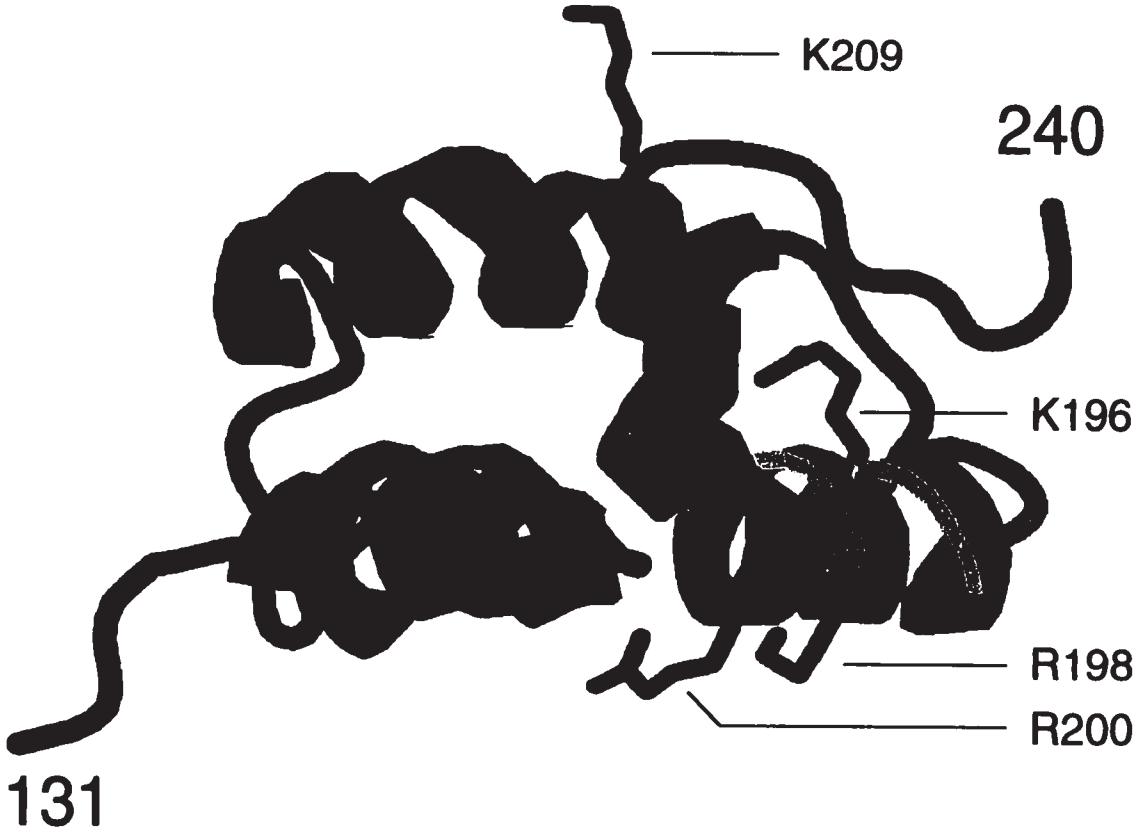
The interpretation and structure determination of TFIIS was performed in collaboration with Dr. C. Arrowsmith (OCI, Toronto) and included here to provide the framework for interpreting mutagenic studies. Preliminary NMR spectra were collected on the smallest active fragment of TFIIS$_{143\text{-}309}$. Although this fragment was active, the ^{15}N-HSQC NMR spectra indicated that TFIIS $_{143\text{-}309}$ had regions of conformational heterogeneity and/or disorder and was not suitable for structural analysis. NMR analysis of contiguous deletion mutants was performed to identify regions that yielded suitable NMR spectra. Two derivatives were identified, yeast TFIIS $_{131\text{-}309}$ and yeast TFIIS $_{131\text{-}240}$, that displayed NMR spectra indicative of stable domains in a single conformational state. Yeast TFIIS $_{131\text{-}240}$ is contained within the proteolytically defined domain II; yeast TFIIS $_{131\text{-}309}$ contains domains II + III, including the complete, naturally occurring carboxyl-terminus. Comparison of the ^{15}N-

HSQC and the methyl region of ^{13}C-HSQC spectra derived from these two protein fragments showed that the chemical shifts of yeast TFIIS $_{131\text{-}240}$ represented a subset of those found in the yeast TFIIS $_{131\text{-}309}$ spectra; approximately 72% of the amide ^{15}N-^{1}H peaks and approximately 76% of the methyl ^{13}C-^{1}H peaks of residues in the structured regions of domain II have identical chemical shifts in yeast $TFIIS_{131\text{-}240}$ and in yeast $TFIIS_{131\text{-}309}$. Thus, to a first approximation, it appears that no appreciable disruption of the structure of domain II is caused by the absence of the carboxyl terminal domain III.

Amino acid residues 139-215 form a well defined three helix bundle composed of helix 1: R149-L160, helix 2: I171-M182, and helix 3: A193-S204 with a short, amino terminal, extended strand D137-T142 (Figure 13). The hydrophobic core of this bundle is built around three tyrosine residues (Y195, Y199, Y203) that protrude from the same side of helix 3. The short amino terminal strand folds V139 back onto a hydrophobic patch on helix 1 formed largely by Y157. The absence of this interaction in yeast $TFIIS_{143\text{-}240}$ may explain its decreased stability in solution. Although the carboxyl terminal 25 amino acids of yeast $TFIIS_{131\text{-}240}$ do appear to contain some small level of helical secondary structure, the region displays fewer NOEs, somewhat narrower line widths and faster amide exchange properties than does the 3 helix bundle region, residues 139-215. The amino terminal eight residues (GSHM from the cloning vector and P131-S134 of yeast TFIIS) contain no regular secondary nor tertiary structure as evidenced by the low abundance of NOEs in this region.

Simultaneous alignment of helices 1, 2, and 3 for the ensemble of structures to its mean yields a root mean squared (r.m.s.) deviation of 0.52 Å for backbone atoms (N, C', C^{α}) and 0.91 Å for all heavy atoms in the helices. R.m.s. deviations are 0.98 Å for backbone and 1.43 Å for all heavy atoms in the structured region (V139-K215). Secondary structure assignments of yeast $TFIIS_{131-309}$ indicate the presence of a fourth helix comprised of residues 227-257 as well as a three strand beta sheet structure, similar to the secondary structure previously determined for the human TFIIS carboxyl terminal domain.

Figure 13: Three dimensional solution structure of $TFIIS_{131-240}$. Ribbon representation of the structure of $TFIIS_{131-240}$ determined by NMR. Residues implicated in mediating protein-protein contacts with RNA polymerase II are indicated. N, amino terminus; C, carboxyl terminus.


K209
240
K196
R198
R200
131

VII. Identification of the site of interaction between RNA polymerase II and TFIIS

Yeast strains lacking TFIIS displayed a sensitivity to 6 azauracil that can be suppressed by restoration of TFIIS expression (Hubert et al., 1983). Based upon these observations a genetic screen was used to select for RNA polymerase II mutants. Seven yeast strains, containing mutants of yeast RNA polymerase II, were identified which displayed a sensitivity to 6 azauracil which was suppressed by restoration of TFIIS expression (Archambault et al., 1992).

Two mutant yeast strains, RPO21-18 and RPO21-24, were obtained from Dr. J.D. Friesen (University of Toronto, Canada), containing amino acid alterations in the largest subunit of RNA polymerase II (I1237TRARV and E1230K, respectively) were characterized in transcript elongation and protein-protein interaction assays with purified yeast TFIIS. Transcript elongation assays were performed with purified RNA polymerase II mutants at 3:1 and 600:1 molar ratios of TFIIS to RNA polymerase II. Although both mutant and wild type polymerases recognized the TIa arrest site with a similar efficiency, the ability of TFIIS to stimulate mutant RNA polymerase II to read through the TIa site at low molar ratios of TFIIS was completely abrogated. Molar ratios of TFIIS:RPO21-24 of 600:1 were required to stimulate the ternary complexes at a level comparable to a 3:1 molar ratio of TFIIS : wild type RNA polymerase II (data not shown). A similar requirement for high molar ratios of TFIIS to stimulate transcript shortening was observed in transcript cleavage assays (Figure 14). The transcriptional activity of mutants RPO21-18 and RPO21-24 were similar to wild type

RNA polymerase II with the exception of their ability to be stimulated by TFIIS (Wu et al., 1996).

One explanation to account for the reduced sensitivity of mutant RNA polymerases to TFIIS is a diminished ability of the RNA polymerase to cleave the nascent transcript, an activity intrinsic to the polymerase. Previous studies using *E. coli* RNA polymerase indicated that arrested ternary complexes could be induced to cleave the nascent transcript in alkaline conditions, without addition of transcription factors (Orlova et al., 1995). I have demonstrated that yeast RNA polymerase II can also cleave its nascent transcript under alkaline conditions in the absence of TFIIS (Figure 15). The transcripts shortened by intrinsic cleavage are extended back to the TIa arrest site upon addition of nucleotides, indicating that the ternary complexes remain intact and transcriptionally active. The capacity of ternary complexes containing wild type and mutant RNA polymerases to undergo intrinsic transcript cleavage were indistinguishable, demonstrating that the two mutant RNA polymerases were fully capable of transcript cleavage, yet unresponsive to TFIIS (Figure 15).

A second explanation for the inability of RNA polymerase II to respond to TFIIS is that the interaction between TFIIS and RNA polymerase II may be disrupted. This indeed was the case since a binary complex was not observed in native gel mobility shift assays. I have demonstrated that the residues of RNA polymerase II responsible for mediating the interaction with TFIIS include I1237TRARV and E1230K.

Figure 14: Analysis of transcript cleavage by RNA polymerase II mutant RPO21-24. Arrested ternary complexes containing RPO21-24 were formed as described in Materials and Methods. Stimulation of transcript cleavage was performed with $TFIIS_{131-309}$:RNA polymerase II molar ratios of 0, 3:1, and 600:1 as indicated, and nucleotides added at the designated intervals for 10 minutes. Transcripts were resolved by gel electrophoresis and visualized by autoradiography. Transcript elongation products are identified at the left as in Figure 5.

TFIIS:Pol II	0					3					600				
Time (hr)	0	.2	1	2	2	0	.2	1	2	2	0	.2	1	2	2
Nucleotides	-				+	-				+	-				+

Figure 15: Intrinsic cleavage activity of wild type and RPO21-24 mutant of RNA polymerase II. Transcripts were formed and visualized as previously described in Materials and Methods. Intrinsic cleavage activity was monitored for up to 60 minutes, and incubation with nucleotides performed for 10 minutes. Transcript elongation products are identified at the left as in Figure 5.

Polymerase	wt Pol II						RPO21-18					RPO21-24				
Time (min)	0	5	15	60	60		0	5	15	60	60	0	5	15	60	60
Nucleotides	-				+		-				+	-				+

Analysis of contiguous deletion and site-directed mutants of yeast TFIIS in transcript elongation and protein interaction assays, in conjunction with the structure, allowed for the functional dissection of TFIIS. Both the amino acid residues of TFIIS important in mediating the protein-protein interaction with RNA polymerase II and the residues that stimulate transcript cleavage were identified. I have also determined at least some of the amino acid residues of the largest subunit of RNA polymerase II that participate in the interaction with TFIIS. This report also contains the first demonstration that the transcript cleavage activity of RNA polymerase II can be stimulated by alkaline conditions.

CHAPTER 4

Discussion

TFIIS stimulates arrested ternary complexes to resume transcription, leading to the production of full length transcripts (Reines et al., 1989). TFIIS activity requires RNA polymerase II binding, via domain II, and a stimulation function within domain III that promotes the ternary complex to cleave the transcript. I have localized the amino acid residues important for these functions by site-directed mutagenesis, and have constructed a model, based upon the NMR solution structure, describing how TFIIS performs its activities.

The formation of an arrested complex is proposed to occur by a multi-step process (Nudler et al., 1997). Upon encountering a block to elongation, the RNA polymerase can either transcribe through or convert to a paused complex. Paused complexes are transitory transcription intermediates for which there are three possible outcomes; continuation of transcript elongation, premature termination, or conversion to an arrested complex. Arrested ternary complexes are incapable of further transcription in the absence of accessory factors. The rate of conversion of a paused complex to an arrested complex or to a complex that resumes transcript elongation is determined by the nature of the block to transcription (Gu and Reines, 1995), and can be affected by alterations of the template DNA sequence (Arndt and Chamberlin, 1990).

Reactivation of arrested ternary complexes occurs in part through a conserved endonucleolytic activity of multisubunit RNA polymerases. The transcript cleavage event, which produces oligonucleotides derived from the 3' terminal end of the nascent

transcript, is stimulated by TFIIS homologues in eukaryotes, and GreA and GreB in bacteria. TFIIS is the only known eukaryotic transcription factor that can reactivate arrested ternary complexes containing RNA polymerase II. I initiated this investigation on TFIIS because I felt that the analysis of the structure / function relationship of TFIIS would be required to determine the mechanism by which TFIIS regulates transcript elongation.

I. Domain structure of yeast TFIIS

I discovered that TFIIS is composed of three domains of which domain I is dispensable for activity *in vitro*, and both domains II and III are required for stimulation of an arrested ternary complex. TFIIS activities could be reduced by altering key residues within either of domains II or III. Domain II is required for TFIIS binding to purified RNA polymerase II; site-directed mutagenesis of charged residues within this domain reduced the affinity of TFIIS for RNA polymerase II. Domain III is required for stimulating the endonucleolytic cleavage event but cleavage also required domain II. The region between domains II and III was also important; alteration of the spatial arrangement of domains II and III by insertion and deletion of amino acid residues caused decreased stimulatory activity for read through and cleavage as observed for mutants 131-309 (ins. @ 240:LKQKI) and 131-309 (del. @ 240-245:LKQKI) (see Table 1).

Activity of Domain II

Domain II of yeast TFIIS is composed of amino acid residues 105 (124) to amino acid 246 as defined by limited proteolytic digestion. However, I determined by deletion analysis that the minimal active fragment was amino acid 144 to the carboxyl terminus suggesting that not all of domain II is essential. A study by Nakanishi et al. (1995) corroborated these results by demonstrating that residues amino terminal to 148 were not required for activity *in vitro*. These results indicate that although domain II is resistant to protease digestion at residues carboxyl terminal to amino acid 105, the region 106 to 147 is dispensable for activity *in vitro*. Residues 131 to 240, which are contained within domain II, are required for the protein-protein interaction between TFIIS and RNA polymerase II.

I also localized the region of domain II responsible for interacting with RNA polymerase II. Site-directed mutagenesis at residues Y195H:E291L, K196A, R198A, R200A, and K209A reduced the affinity of the interaction between 131-309 and RNA polymerase II. The solution structure of $TFIIS_{131-240}$ revealed that the majority of residues implicated in binding to RNA polymerase are localized to the third helix (amino acids 193-204).

My site-directed mutagenesis data corroborates a previous analysis of human TFIIS by Ciprés-Palacín and Kane (1994, 1995). In their work three important clusters of charged residues in domain II were identified by alanine scanning mutagenesis. Mutations in two of these clusters, termed cluster 1 (containing mutations R141A:K143A:R145A:E146A) and cluster 7b (containing mutations

R188A:R190A:R192A) abolish both read through and transcript cleavage activities. The amino acid residues altered in cluster 1 correspond to the region of yeast TFIIS containing the site-directed mutant D150A, which by NMR analysis is misfolded. I suggesting that the human cluster 1 mutant may also be misfolded and therefore interpretation of this mutant in biochemical assays is difficult.

The human cluster 7b mutant corresponds to the yeast mutants R198A and R200A, which display a reduced affinity for RNA polymerase II. The yeast site-directed mutants display small structural alterations indicating the introduction of these mutations do not affect the protein structure. I suggest that the human cluster 7b mutant has a reduced affinity for RNA polymerase II, which remains to be tested.

Mutations in a third cluster within domain II, cluster 7 (containing mutations K186A:K188A), caused a loss of read through activity yet 50% of wild type cleavage activity was retained. The human TFIIS mutant containing the cluster 7 alterations corresponds to the yeast site-directed mutant K196A, which also displayed a reduced affinity for RNA polymerase II. Based upon the homology between yeast and human TFIIS I suggest that the human cluster 7 mutant may have a reduced affinity for RNA polymerase II.

Ciprés-Palacín and Kane (1994, 1995) interpreted their observations to suggest that the read through and cleavage activities of human TFIIS can be uncoupled. I suggest that their observations might be explained simply by a decreased affinity for RNA polymerase II. A reduction of the affinity between TFIIS and RNA polymerase II causes a decreased rate of transcript cleavage, and I suggest that the affinity

between the cluster 7 and RNA polymerase II may have been sufficiently reduced to stimulate cleavage at only the first cleavage site. Indeed in their paper only cleavage to the first site occurred. Since the cleavage events appear to be sequential, and the formation of the second cleavage product is correlated with read through by the ternary complex (which will be discussed later) I suggest the observation that read through and cleavage can be uncoupled may be from not discriminating between the two cleavage products.

A study by Horikoshi et al. (1990) indicated that the region of human TFIIS that mediates the interaction with RNA polymerase II is not contained within domain II. In their study a site-directed mutant of human TFIIS containing residues R247, E248 and H249 mutated to leucine showed reduced binding to RNA polymerase II and a reduction in the stimulation of transcription in non-specific transcription assays. A yeast TFIIS mutant, N252A:N255A:Q257A, also demonstrated a reduction in transcript elongation activity, however the affinity with RNA polymerase II was unaffected. The discrepancy between our results might be due to their method of analysis or due to species-specific differences. In their study, binding of TFIIS to RNA polymerase II was assayed by glycerol gradient centrifugation in a non-quantitative manner. Second, the tertiary structure of the human R247L:E248L:H249L mutant is likely disrupted since it displayed an altered sensitivity to the protease chymotrypsin.

Activity of Domain III

Domain III, which is composed of residues 247 to 309, was identified by limited proteolytic digestion of TFIIS. The presence of a protease sensitive site at residue 246 indicated that there is a solvent accessible region that separates domain II and domain III. Domain III is resistant to limited proteolytic digestion, highly conserved among TFIIS homologues, encompasses a zinc binding motif, and is required for TFIIS mediated transcript cleavage.

Residues important for the stimulation of transcript cleavage were revealed by site-directed mutagenesis within domain III. Site-directed mutants Q283A:Q285A, E291N, E291H, and R287Q:E291H are completely defective in stimulating transcription through the TIa arrest site. Site-directed mutants that display some residual stimulatory activity in read through assays include K242A:Q243A, N252A:N255A:Q257A, T266A:D267A, F269A, T286A:S288A, R287Q:E291N, F296A, K307A, and F308A. I suggest that these residues may be involved in interactions with either RNA polymerase II or the nascent RNA transcript. Interactions between domain III of TFIIS and RNA polymerase were not detected in native gel mobility shift assays, indicating that any interaction between this domain and purified RNA polymerase has a K_d of greater than 2 mM. This does not rule out the possibility that domain III interacts with the polymerase in the context of an arrested ternary complex.

The requirement for domain III of yeast TFIIS in the stimulation of transcript cleavage is supported by previous studies on a human TFIIS homologue. Site-directed

mutants targeting conserved residues, from 256 to 264, contained within the zinc ribbon of human TFIIS (which correspond to residues 285 to 293 of the yeast homologue) were deficient in transcript cleavage activity (Jeon et al., 1994). Transcript elongation assays with mutant human TFIIS were performed at equimolar ratios of mutant TFIIS:RNA polymerase II and therefore both small structural and functional group alterations within the zinc ribbon could be interpreted as causing a drastic reduction in stimulatory activity. The mutations that I introduced into the yeast TFIIS homologue at these same conserved residues revealed that these mutations do not abrogate activity; stimulation of transcript cleavage is evident at high molar ratios of TFIIS:RNA polymerase II. Characterization of the human mutants at higher molar ratios may likely have resulted in similar observations as described here for the yeast TFIIS homologue.

The region of human TFIIS homologous to domain III of yeast TFIIS has been suggested to interact with single stranded and duplex nucleic acids. This interaction is observed following the deletion of both domains I and II, and the affinity between domain III and nucleic acid is low with a K_d on the order of 10^{-3}M. This weak interaction may be important in the stimulation of transcript cleavage, however, studies comparing the nucleic acid binding and the ability to stimulate transcript cleavage of mutants within domain III have not been reported.

II. Domain III of TFIIS demonstrates structural similarity to an RNA binding motif

The structure of domain III of human TFIIS, determined by Qian et al. (1993a), shows similarity to evolutionarily conserved RNA binding motifs. These binding motifs are composed of a beta sheet region, and a helical stand that appears to play a role in structural stabilization. The main conserved features are aromatic residues which protrude from the beta strands, and are solvent accessible. The exposed aromatic residues are important for interacting with the RNA strand by contacting the nucleotide bases. Although the binary structure of domain III and nucleic acid has not been determined, I have demonstrated by site-directed mutagenesis that alteration of aromatic residues F269, F296, and F308, causes a reduction in TFIIS activity. These aromatic residues are exposed to the solvent as determined from the NMR derived structure by Qian et al. (1993a), and are conserved in TFIIS homologues. I suggest that domain III of TFIIS binds to the nascent RNA transcript via the exposed aromatic residues, and the RNA binding plays a role in the stimulation of transcript cleavage.

III. Cleavage and read through by the ternary complex is stimulated by TFIIS

The transcript cleavage factors, GreA, GreB, and TFIIS, all appear to function by stimulating the endogenous cleavage activity of their respective RNA polymerases. Three observations suggest that RNA polymerase is responsible for choosing the site of cleavage and performing the transcript cleavage event. First, transcript cleavage can be induced by alkaline solution conditions resulting in the generation of cleavage products of a similar size as those generated by stimulation with transcription factors

TFIIS (Awrey et al., 1997), GreA and GreB (Orlova et al., 1995). Second, the transcript can be cleaved by pyrophosphorolysis, the reverse reaction of nucleotide addition, resulting in cleavage products identical to factor stimulated cleavage (Orlova et al., 1995; Rudd et al., 1994). Third, the transcript cleavage event is observed upon GreB stimulation of transcript cleavage in binary complexes containing RNA and *E. coli* RNA polymerase (Altmann et al., 1994). The transcript cleavage property of binary complexes indicates that the cleavage site selection may be determined in part by the interactions between the nascent RNA and the polymerase.

Ternary complexes arrested at TIa are stimulated by TFIIS to cleave at two distinct sites within the nascent RNA chain (Reines, 1992). Alteration of residues within both domain II and domain III can affect the rate of transcript cleavage at the first and second cleavage sites. The rate of transcript cleavage at the second site was drastically reduced by site-directed mutagenesis of acidic residues within domain III, but transcript cleavage at the first site was not completely abolished. This observation indicates that there is a functional and structural difference between these two cleavage events.

There is a structural difference between ternary complexes that participate in cleavage at the first and second sites. Analysis of the DNA footprint of RNA polymerase II at TIa and both cleavage sites showed that the footprint before and after the first cleavage event does not change (Gu et al., 1993). The observation that translocation of the polymerase is not required prior to cleavage at the first site indicates that either the polymerase contains a mobile active site or the formation of

an arrested complex involves a repositioning of the polymerase on the template. Cleavage at the second site is associated with a translocation of the polymerase on the template DNA demonstrating that the ternary complexes involved in cleaving the transcript at C1 and C2 may be functionally distinct (Gu et al., 1993). Additional evidence that supports the hypothesis that cleavage at C1 and C2 are performed via different mechanisms is described below.

I propose that the cleavage events are sequential based on three observations. First, in TFIIS mutants displaying a reduced rate of transcript cleavage, near quantitative transcript cleavage occurs at the first cleavage site prior to the second cleavage event. Second, the formation of the second cleavage product does not result in a reduction in the residual amount of arrested complex at TIa, but does result in a reduction in the amount of the first cleavage product. Finally, the second cleavage product does not appear in the absence of the first cleavage product.

I have demonstrated that the first cleavage event is insufficient to allow progression through the arrest site. Transcript cleavage at the second site is correlated with reactivation of the ternary complex to progress through the TIa site. However, a previous study that analyzed the formation of cleavage products demonstrated that only one cleavage event was necessary for progression through the arrest site (Gu et al., 1993). Two observations support this assumption; first, TFIIS stimulated transcript cleavage by arrested complexes was monitored in the presence of nucleotides and although the first cleavage product was observed, no transcript cleavage at the second cleavage site was observed. Second, Gu et al. (1993) also trapped the cleavage

products by incorporation of chain terminating nucleotides, demonstrating convincingly that the second cleavage event is not required for reactivation of ternary complexes. In summary, transcript cleavage at the first site is necessary, but not sufficient to promote transcription through arrest sites.

Two explanations exist for the discrepancy in regard to the requirement for a second cleavage event. First, Gu et al. (1993) formed arrested ternary complexes in the presence of the general initiation factors including TFIIF, a factor known to affect transcript elongation. It is possible that the presence of TFIIF affected the ternary complex and promoted transcription through the arrest site after cleavage to the first site. Second, the correlation that I observed between cleavage at the second site and read through of the TIa arrest site is consistent with the following hypothesis. Cleavage at the second site has been demonstrated to be correlated with translocation of the ternary complex on the DNA template (Gu et al., 1993). After cleavage of the transcript at the first site, a conformational change may be required that allows the ternary complex to translocate. The ability for the ternary complex to translocate allows for either read through of the arrest site in the presence of nucleotides or transcript cleavage at the second site in the absence of nucleotides. In my analysis, cleavage would be favoured, due to the lack of nucleotides whereas in the study by Gu et al. (1993) transcription through the arrest site is favoured. Also, mutant K307A displayed an inability to stimulate progression through the arrest site even after inducing cleavage at the second site. An interpretation of this data surmises that TFIIS is responsible not only for stimulating the nascent transcript cleavage but also

for inducing a conformational change within the ternary complex prior to transcription through the arrest site.

VI. Structural comparison of GreA and GreB with TFIIS

TFIIS is functionally similar to the bacterial transcript elongation factors GreA and GreB. Each of these factors binds to its respective RNA polymerase, promotes read through of arrest sites and stimulates transcript cleavage. The functional similarity does not extend to sequence conservation; although the primary sequences of GreA and GreB share 50% sequence similarity, they are completely unrelated to TFIIS (Borukhov et al., 1993). A comparison of the structures of TFIIS and GreA (Stebbins et al., 1995) revealed a global similarity; each factor is composed of two domains of approximately equal size, an amino terminal helical domain and a carboxyl terminal β-sheet domain. In addition, the base of the helical domain in both GreA and TFIIS contain basic residues that are critical for transcription activity. However, in detail, the structures of TFIIS and GreA are significantly different; the helical domain of GreA is an extended, anti-parallel coiled coil, while that of TFIIS is a more compact helix bundle, also the beta strand domain of GreA does not bind Zn^{2-} as does domain III of TFIIS.

VII. Function of Intervening Region between Conserved Domains G and H of RNA Polymerase II Largest Subunit

The domain of RNA polymerase II responsible for TFIIS binding was initially identified through the use of a yeast genetic screen (Archambault et al., 1992). Purification of RNA polymerase II containing the mutant residues implicated in binding to TFIIS followed by biochemical characterization of these mutants unequivocally established the region responsible for this interaction.

Three observations indicate that TFIIS interacts with the 3' terminus of the RNA in an arrested complex. First, TFIIS interacts with the amino acid region between the conserved domains G and H, which is evolutionarily conserved among eukaryotic RNA polymerase II isoforms of the largest subunit. Second, domain G of *E. coli* RNA polymerase β' subunit interacts with the 3' end of the RNA chain (Markovtsov et al., 1996). The interaction between the 3' terminus of the transcript and domain G is more apparent in an arrested ternary complex. Third, TFIIS bound to an arrested ternary complex is reactive to photo cross-linking reagents incorporated into the 3' terminus of the nascent RNA (Powell et al., 1996).

TFIIS has also been implicated in binding to other regions of RNA polymerase II; the CTD (Sawadogo et al., 1980) and conserved region B of the largest subunit (Rappaport et al., 1988). However, a direct interaction between TFIIS and either the CTD or conserved region B has not been observed. In fact, transcript elongation assays with mutants of RNA polymerase II demonstrated that TFIIS does not require either the CTD of the largest subunit or the fourth and seventh largest subunits for

activity (Christie et al., 1994). Competition experiments performed with a fusion protein containing conserved region B of the largest subunit partially inhibited TFIIS activity in promoter independent, non-specific transcription assays. TFIIS activity was also partially reduced by the presence of an antibody directed against conserved domain B. This evidence implies that conserved region B may be in close spatial proximity to the region between conserved regions G and H, but does not demonstrate a direct interaction between TFIIS and conserved domain B.

VIII. Modulation of transcript elongation

Two models have been put forth to explain the properties of an elongating ternary complex. An earlier model, presented in the introduction, depicts the ternary complex as an 'inchworm' (Chamberlin, 1995) that can progress along the DNA template in a discontinuous manner. The variations in the size of the RNA polymerase footprint on the DNA are dependent on the template sequence and proposed to result from large incremental movements of the front of the polymerase. A newer model explains several more recent results and describes the polymerase as a sliding clamp (Nudler et al., 1997) that does not undergo the large structural alterations predicted by the Chamberlin model. Although caveats exist for both models, and in no manner does either model explain all of the experimental data, I will describe the sliding clamp model of transcript elongation as a framework for the presentation of my model of TFIIS activation. Upon encountering a template sequence that impedes the forward progression of the ternary complex, there is a reduction in

the size of the RNA polymerase footprint on the DNA template. The observed reduction is due to a decrease in the distance from the leading edge of the polymerase to the 3' terminus of the nascent RNA and was postulated to occur in response to encountering a pause signal. Progression through the pause site correlates with a large incremental increase in the length of the footprint at the front of the polymerase (Nudler et al., 1994; Wang et al., 1995). The different footprint sizes of RNA polymerase led Chamberlin to suggest his inchworm model. However, the different sizes of the RNA polymerase footprint could also be explained if the paused polymerase slid along the template; a large degree of mobility would cause an apparent decrease in the footprint length. RNA polymerase can indeed move in a backward direction which implies that a large conformational shift may not occur, and the reduced footprint size may be attributed to a large degree of mobility by the polymerase on the template (Nudler et al., 1997).

The central contention between the inchworming and the sliding clamp models is the length of the DNA:RNA hybrid and the role that the hybrid has in stabilizing the ternary complex. The sliding clamp model predicts that the DNA:RNA hybrid maintains a constant length of eight to nine base pairs. Sliding of the polymerase is predicted to occur as a result of a destabilization of the DNA:RNA hybrid at the 3' terminus of the nascent transcript, while maintaining the eight to nine base pair hybrid. The backward movement of the polymerase promotes the release of the 3' terminal nucleotide from the catalytic site (Nudler et al., 1997) and this conformation of the ternary complex is predicted to be a paused complex. In this model, DNA sequences

upstream of the leading edge of RNA polymerase that impede transcription may influence the polymerase to move in a backward direction, thereby promoting the formation of a paused complex. Paused complexes, formed by sliding of the polymerase on the template, may be stabilized by secondary structure elements within the nascent RNA chain that interact with the ternary complex at an unknown site.

In the sliding clamp model, the formation of an arrested complex is predicted to occur through extensive sliding of the ternary complex on the DNA template followed by the interaction of the 3' terminal region of the nascent RNA with a site on the polymerase. The interaction between the 3' terminal nucleotides and the polymerase is suggested to result in an arrested complex, that can neither slide on the template nor catalyze nucleotide addition. The conformation of the nascent RNA within an arrested complex is supported by the observation that the size of the oligonucleotides released from the 3' terminus of the transcript by GreB stimulation and Fe^{2+} hydroxyl radical cleavage is the same. The length of the released oligonucleotide corresponds to the number of nucleotides by which the leading edge of the ternary complex has moved closer to the 3' hydroxyl of the RNA (Wang et al., 1995; Nudler et al., 1997). The release from an arrested ternary complex is therefore predicted to occur by the removal of the region of the RNA that stabilizes the arrested complex.

IX. Model of TFIIS Stimulated Transcript Cleavage

TFIIS is responsible for the stimulation of an intrinsic cleavage activity of RNA polymerase II. The mechanism by which this occurs involves the interaction of TFIIS

with either the nascent RNA chain or the polymerase catalytic site. As described above, arrested complexes are proposed to contain the 3' terminus of the nascent RNA chain downstream of the active site. TFIIS interacts with the polymerase, via the domains isolated on both the polymerase and TFIIS, to position domain III of TFIIS in proximity to the 3' terminus of the nascent RNA.

TFIIS stimulates the nascent transcript cleavage at the active site of RNA polymerase II through an unknown mechanism(s). A common requirement between the intrinsic cleavage of RNA polymerase and TFIIS directed cleavage (Reines, 1992) is the requirement for Mg^{2+}. This cation is also required for nucleotide addition (Kadesch and Chamberlin, 1982) and suggests the involvement of the catalytic active site in transcript cleavage. Alternatively, Mg^{2+} may be required for the association of the 3' terminal RNA with the ternary complex in a manner that stimulates the cleavage event. The proximity of TFIIS to the 3' terminus of the nascent RNA (Powell et al., 1996) indicates that either TFIIS binds to the RNA and induces the intrinsic cleavage activity of RNA polymerase, or TFIIS interacts with the RNA polymerase near the 3' terminus of the RNA and effects a change in the polymerase, which in turn results in a cleavage event.

Residues within the disordered loop and three beta strands of the zinc ribbon motif are important for the stimulation of the cleavage activity. I propose that this domain is responsible for two activities required for the stimulation of transcript cleavage. First, residues in the zinc ribbon interact with the RNA contained within the ternary complex, resulting in an alteration in the RNA conformation at the catalytic

site. The interaction may be mediated by aromatic residues contained within the three beta sheets, which comprise a structural motif commonly found in single strand RNA binding proteins (Burd and Dreyfuss, 1994; Bochkarev et al., 1997). Second, the acidic residues within the disordered loop (located between the first and second beta sheet strands) interact with the catalytic site, and facilitate the deprotonation of a amino acid residue(s). The ability of the ternary complex to cleave the nascent transcript under alkaline conditions indicates that a deprotonation event may occur at a residue located near the catalytic site. A conserved histidine residue, H490, is located four residues carboxyl terminal to the predicted catalytic site and since the pKa of the intrinsic transcript cleavage reaction is approximately 8.5, I suggest that this residue may play a role in transcript cleavage. This model, which describes two activities for domain III, also explains why transcript cleavage is only completely abrogated upon removal of domain III, but not by independently disrupting either the proposed RNA binding activity (aromatic residues) or the catalytic site stimulation (acidic loop residues). The inability to completely abolish transcript cleavage, short of the deletion of domain III, might indicate that the RNA binding and catalytic site stimulatory activities act in a cooperative manner.

Cleavage of the nascent transcript is insufficient to promote transcription through an arrest site. Following the stimulation of transcript cleavage to the first cleavage site by several mutants of TFIIS, RNA synthesis occurs back to the arrest site, but the ternary complex does not proceed through it. I propose that a second TFIIS stimulatory event is required following transcript cleavage to promote a

conformational change prior to the resumption of transcript elongation. Although the conformational changes that occur in the conversion of an arrested complex to an elongation competent complex have not been thoroughly investigated, I speculate that removal of oligonucleotides from the 3' terminus of the transcript is insufficient to promote the resumption of transcription through arrest sites. TFIIS may play a role in either the stabilization of the DNA:RNA hybrid within the transcription bubble or the dissociation of inhibitory RNA secondary structure elements from the ternary complex, allowing transcription to resume through arrest sites.

REFERENCES

Agarwal, K., Baek, K., Jeon, C., Miyamoto, K., Ueno, A., and Yoon, H. (1991). Stimulation of transcript elongation requires both the zinc finger and RNA polymerase II binding domains of human TFIIS. *Biochemistry* **30**: 7842-7851.

Allison, L. A., Moyle, M., Shales, M., and Ingles, C. J. (1985). Extensive homology among the largest subunits of eukaryotic and prokaryotic RNA polymerases. *Cell* **42**: 599.

Altmann, C. R., Solow-Cordero, D. E., and Chamberlin, M. J. (1994). RNA cleavage and chain elongation by *Escherichia coli* DNA dependent RNA polymerase in a binary enzyme-RNA complex. *Proccedings of the National Academy of Sciences USA* **91**: 3784-3788.

Archambault, J., Lacroute, F., Ruet, A., and Friesen, J. D. (1992). Genetic interactio between transcription elongation factor TFIIS and RNA polymerase II. *Molecular Cellular Biology* **12**: 4142-4152.

Arndt, K. M., and Chamberlin, M. J. (1990). RNA chain elongation by *E. coli* RNA polymerase. Factors affecting the stability of elongating ternary complexes. *Journal of Molecular Biology* **213**: 79-108.

Aso, T., Lane, W. S., Conaway, J. W., and Conaway, R. C. (1995). Elongin (SIII): A multisubunit regulator of elongation by RNA polymerase II. *Science* **269**: 1439-1443.

Austin, R. J., and Biggin, M. D. (1996). Purification of the Drosophila RNA polymerasee II general transcription factors. *Proceedings of the National Academy of Sciences USA* **93**: 5788-5792.

Awrey, D. E., Weilbaccher, .R.G., Hemming, S. A., Orlicky, S. M., Kane, C. M., and Edwards, A. M. (1997). Transcription elongation through DNA arrest sites: A multi-step process involving both RNA polymerase II subunit RPB9 and TFIIS. *Journal of Biological Chemistry* : in press.

Bartsch, I., Schonenberg, C., and Grummt, I. (1988). Purification and characterization of TTFI, a factor that mediates termination of mouse ribosomal DNA transcription. *Molecular and Cellular Biology* **8**: 3891-3897.

Bengal, E., Flores, O., Krauskopf, A., Reinberg, D., and Aloni, Y. (1991). Role of the mammalian transcription factors IIF, IIS, and IIX during elongation by RNA polymerase II. *Molecular Cellular Biology* **11**: 1195-1206.

Blatter, E. E., Ross, W., Tang, H., Gourse, R. L., and Ebright, R. H. (1994). Domain organization of RNA polymerase alpha subunit: C-terminal 85 amino acids constitute a domain capable of dimerization and DNA binding. *Cell* **78**: 889-896.

Bochkarev, A., Pfuetzner, R. A., Edwards, A. M., and Frappier, L. (1997). Crystal structure of the single-stranded DNA binding domain of replication protein A bound to DNA. *Nature* **385**: 176-181.

Borukhov, S., and Goldfarb, A. (1993). Recombinant *Escherichia coli* RNA polymerase: purification of individually overexpressed subunits and *in vitro* assembly. *Protein Expr. Purif.* **4**: 503-511.

Borukhov, S., Polyakov, A., Nikiforov, V., and Goldfarb, A. (1992). GreA protein: a transcription elongation factor from *Escherichia coli. Proceedings of the National Academy of Science* **89**: 8899-8902.

Borukhov, S., Sagitov, V., and Goldfarb, A. (1993). Transcript cleavage factors from *E. coli. Cell* **72**: 459-466.

Bradsher, J. N., Jackson, K. W., Conaway, R. C., and Conaway, J. W. (1993). RNA polymerase II transcription factor SIII. I. Identification, purification, and properties. *Journal of Biological Chemistry* **268**: 25587-25593.

Bradsher, J. N., Tan, S., McLaury, H. J., Conaway, J. W., and Conaway, R. C. (1993). RNA polymerase II transcription factor SIII. II. Functional properties and role in RNA chain elongation. *Journal of Biological Chemistry* **268**: 25594-25603.

Brennan, C. A., Dombroski, A. J., and Platt, T. (1987). Transcription termination factor rho is an RNA-DNA helicase. *Cell* **48**: 945-952.

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

Burd, C., and Dreyfuss, G. (1994). Conserved structures and diversity of functions of RNA-binding proteins. *Science* **265**: 615-621.

Burgess, R. (1976). Purification and physical properties of *E. coli* RNA polymerase. In *RNA Polymerase*. Cold Spring Harbor, NY, Cold Spring Harbor Laboratory Press. 69-100.

Burgess, R. R., Travers, A. A., Dunn, J. J., and Bautz, E. (1969). Factor stimulating transcription by RNA polymerase. *Nature* **221**: 43-46.

Burton, Z. F., Killeen, M., Sopta, M., Ortolan, L. G., and Greenblatt, J. (1988). RAP30/74: a general initiation factor that binds to RNA polymerase II. *Molecular and Cellular Biology* **8**: 1602-1613.

Cadena, D., and Dahmus, M. (1987). Messenger RNA synthesis in mammalian cells is catalyzed by the phosphorylated form of RNA polymerase II. *Journal of Biological Chemistry* **262**: 12468-12474.

Chamberlin, M. J. (1995). New models for the mechanism of transcription elongation and its regulation. *Harvey Lectures* **Series 88**: 1-21.

Chan, C., and Landick, R. (1994). New perspectives on RNA chain elongation and termination. In *Transcription: Mechanism and Regulation*. New York, Raven Press. 297-320.

Chapman, K. A., and Burgess, R. R. (1987). Construction of bacteriophage T7 late promoters with point mutations and characterization by in vitro transcription properties. *Nucleic Acids Research* **15**: 5413-5432.

Chen, H. C., England, L., and Kane, C. M. (1992). Characterization of a HeLa cDNA clone encoding the human SII protein, an elongation factor for RNA polymerase II. *Gene* **116**: 253-258.

Christie, K., Awrey, D., Edwards, A. M., and Kane, C. M. (1994). Purified Yeast RNA Polymerase II Reads Through Intrinsic Blocks to Elongation in Response to the Yeast TFIIS Analogue. *Journal of Biological Chemistry* **269**: 936-943.

Churchill, M. E. A., Tullius, T. D., and Klug, A. (1990). Mode of interaction of the zinc finger protein TFIIIA with a 5S RNA gene of *Xenopus*. *Proceedings of the National Academy of Sciences USA* **87**: 5528-5532.

Cipres-Palacin, G., and Kane, C. M. (1995). Alanine scanning mutagenesis of human transcription elongation factor TFIIS. *Biochemistry* **34**: 15375-15380.

Ciprés-Palacín, G., and Kane, C. M. (1994). Cleavage of the nascent transcript induced by TFIIS is insufficient to promote read-through of intrinsic blocks to elongation by RNA polymerase II. *Proceedings of the National Academy of Sciences USA* **91**: 8087-8091.

Clark, A. B., Dykstra, C. C., and Sugino, A. (1991). Isolation, DNA sequence and regulation of a *Saccharomyces cerevisiae* gene that encodes DNA strand transfer protein α. *Molecular Cellular Biology* **11**: 2576-2582.

Comai, L., Tanese, N., and Tjian, R. (1992). The TATA-binding protein and associated factors are integral components of the RNA polymerase I trascription factor, SL1. *Cell* **68**: 965.

Conaway, J. W., Reines, D., and Conaway, R. C. (1990). Transcription Initiated by RNA Polymerase II and Purified Transcription Factors from Liver. *The Journal of Biological Chemistry* **265**: 7552-7558.

Darst, S. A., Edwards, A. M., Kubalek, E. W., and Kornberg, R. D. (1991). Three-Dimensional Structure of Yeast RNA Polymerase II at 16 A Resolution. *Cell* **66**: 1-20.

Darst, S. A., Kubalek, E. W., and Kornberg, R. D. (1989). Three-dimensional structure of *Escherichia coli* RNA polymerase holoenzyme determined by electron crystallography. *Nature* **340**: 730-732.

Das, A., Merril, C., and Adhya, S. (1978). Interaction of RNA polymerase and rho in transcription termination: coupled ATPase. *Proceedings of the National Academy of Sciences USA* **75**: 4828-4832.

Daube, S. S., and von Hippel, P. H. (1992). Functional transcription elongation complexes from synthetic RNA-DNA bubble duplexes. *Science* **258**: 1320-1324.

Donahue, B. A., Yin, S., Taylor, J.-S., Reines, D., and Hanawalt, P. C. (1994). Transcript cleavage by RNA polymerase II arrested by a cyclobutane pyrimidine dimer in the DNA template. *Proceedings of the National Academy of Sciences USA* **91**: 8502-8506.

Drapkin, R., Sancar, A., and Reinberg, D. (1994). Where transcription meets repair. *Cell* **77**: 9-12.

Duan, D. R., Pause, A., Burgess, W. H., Aso, T., Chen, D. Y. T., Garrett, K. P., Conaway, R. C., Conaway, J. W., Linehan, W. M., and Klausner, R. D. (1995). Inhibition of transcription elongation by the VHL tumor suppressor protein. *Science* **269**: 1402-1406.

Edwards, A. M., Darst, S. A., Feaver, W. J., Thompson, N. E., Burgess, R. R., and Kornberg, R. D. (1990). Purification and lipid-layer crystallization of yeast RNA polymerase II. *Proceedings of the National Academy of Sciences USA* U.S.A. **87**: 2122-2126.

Edwards, A. M., Kane, C. M., Young, R. A., and Kornberg, R. D. (1991). Two dissociable subunits of yeast RNA polymerase II stimulate the initiation of transcription at a promoter *in vitro*. *Journal of Biological Chemistry* **266**: 71-75.

Erie, D. A., Hajiseyedjavadi, O., Young, M. C., and von Hippel, P. H. (1993). Multiple RNA polymerase conformations and GreA: control of the fidelity of transcription. *Science* **262**: 867-873.

Exinger, F., and Lacroute, F. (1992). 6-Azauracil inhibition of GTP biosynthesis in *Saccharomyces cerevisiae*. *Current Genetics* **22**: 9-11.

Felsenfeld, G. (1992). Chromatin: an essential part of the transcriptional mechanism. *Nature* **355**: 219-224.

Flores, O., Maldonado, E., and Reinberg, D. (1988). Factors involved in specific transcription by mammalian RNA polymerase II: RNA polymerase II associating protein 30 is an essential component of transcription factor IIF. *Journal of Biological Chemistry* **263**: 10812-10816.

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

French, S. (1992). Consequences of replication fork movement through transcription units in vivo. *Science* **258**(1362-1365).

Gamper, H., and Hearst, J. (1982). A topological model for transcription based on unwinding angle analysis of *E. coli* RNA polymerase binary, initiation and ternary complexes. *Cell* **29**: 81-90.

Gardella, T., Moyle, H., and Susskind, M. M. (1989). A mutant *Escherichia coli* sigma 70 subunit of RNA polymerase with alterred promoter specificity. *Journal of Molecular Biology* **206**: 579-590.

Geiduschek, E. P., and Tocchini-Valentini, G. P. (1988). Transcription by RNA polymerase III. *Annual Review of Biochemistry* **57**: 873.

Gerard, M., Fischer, L., Moncollin, V., Chipoulet, J. M., Chambon, P., and Egly, J. M. (1991). Purification and interaction properties of the human RNA polymerase B(II) general transcription factor BTF2. *Journal of Biological Chemistry* **266**: 20940-20945.

Grachev, M. A., Kolocheva, T. I., Lukhtanov, E. A., and Mustaev, A. A. (1987). Studies on the functional topography of Escherichia coli RNA polymerase. Highly selective affinity labelling by analogues of initiating substrates. *European Journal of Biochemistry* **163**: 113-121.

Gu, W., Powell, W., Mote, J., J., and Reines, D. (1993). Nascent RNA cleavage by arrested RNA polymerase II does not require upstream translocation of the elongating complex on DNA. *Journal of Biological Chemistry* **268**: 25604-25616.

Gu, W., and Reines, D. (1995). Identification of a decay in transcription potential that results in elongation factor dependence of RNA polymerase II. *Journal of Biological Chemistry* **270**: 11238-11244.

Gu, W., and Reines, D. (1995). Variation in the size of nascent RNA cleavage products as a function of transcript length and elongation competence. *Journal of Biological Chemistry* **270**: 30441-30447.

Hagler, J., and Shuman, S. (1992). Nascent RNA Cleavage by Purified Ternary Complexes of Vaccinia RNA Polymerase. *Journal of Biological Chemistry* **268**: 2166-2173.

Hanna, M. M., and Meares, C. F. (1983). Topography of transcription path: the leading end of nascent RNA through the *Escherichia coli* transcription complex. *Proceedings of the National Academy of Sciences USA* **80**: 4238-4242.

Hengge-Aronis, R. (1993). Survival of hunger and stress: the role of *rpoS* in early stationary phase gene regulation in *E. coli*. *Cell* **72**: 165-168.

Horikoshi, M., Sekimizu, K., Hirashima, S., Mitsuhashi, Y., and Natori, S. (1985). Structural relationships of the three stimulatory factors of RNA polymerase II from Ehrlich ascites tumor cells. *Journal of Biological Chemistry* **260**: 2166-2173.

Horikoshi, N., Sekimizu, K., and Natori, S. (1990). Site-directed mutagenesis of arginine 246, glutamic acid 247, and histidine 248 in the eukaryotic transcription factor S-II. *Journal of Biological Chemistry* **265**: 11854-11857.

Hsu, L., Vo, N., and Chamberlin, M. (1995). Escherichia coli transcript cleavage factors GreA and GreB stimulate promoter escape and gene expression in vivo and in vitro. *Proceedings of the National Academy of Sciences* USA **92**: 11588-11592.

Hubert, J.-C., Guyonvarch, A., Kammerer, B., Exinger, F., Liljelund, P., and Lacroute, F. (1983). Complete sequence of a eukaryotic regulatory gene. *EMBO Journal* **2**: 2071-2073.

Iida, C. T., Kownin, P., and Paule, M. R. (1985). Ribosomal RNA transcription: proteins and DNA sequences involved in preinitiation complex formation. *Proceedings of the National Academy of Sciences USA* **82**: 1668-1672.

Ishihama, A. (1981). Subunit assembly of *Escherichia coli* RNA polymerase. *Adv. Biophys.* **14**: 1-35.

Izban, M. G., and Luse, D. S. (1992). The RNA polymerase II ternary complex cleaves the nascent transcription in a 3'-> 5' direction in the presence of elongation factor SII. *Genes & Development* **6**: 1342-1356.

Izban, M. G., and Luse, D. S. (1993). The increment of SII-facilitated transcript cleavage varies dramatically between elongation competent and incompetent RNA polymerase II ternary complexes. *Journal of Biological Chemistry* **268**: 12874-12885.

Izban, M. G., and Luse, D. S. (1993). SII-facilitated transcript cleavage in RNA polymerase II stalled early after initiation occurs in primarily dinucleotide increments. *Journal of Biological Chemistry* **268**: 12864-12873.

Jeon, C. J., and Agarwal, K. (1996). Fidelity of RNA polymerase II transcription controlled by elongation factor TFIIS. *Proceedings of the National Academy of Sciences USA* **93**: 13677-13682.

Jeon, C. J., Yoon, H. S., and Agarwal, K. (1994). The transcription factor TFIIS zinc ribbon dipeptide Asp-Glu is critical for stimulation of elongation and RNA cleavage by RNA polymerase II. *Proceedings of the National Academy of Sciences USA* **91**: 9106-9110.

Kadesch, T. R., and Chamberlin, M. J. (1982). Studies of in vitro transcription by calf thymus RNA polymerase II using a novel duplex DNA template. *Journal of Biological Chemistry* **257**: 5286.

Kambadur, R., Culotta, V., and Hamer, D. (1990). Cloned yest and mammalian transcription factor TFIID gene products support basal but not activated metallotheinein gene transcription. *Proceedings of the National Academy of Sciences USA* **87**: 9168-9172.

Kane, C. (1994). Transcript elongation and gene regulation in eukaryotes. In *Transcription: Mechanism and Regulation.* New York, Raven Press. Ltd. 279-296.

Kash, S. F., Innis, J. W., Jackson, A. U., and Kellems, R. E. (1989). Functional analysis of a stable transcription arrest site in the first intron of the murine adenosine deaminase gene. *Molecular and Cellular Biology* **13**: 2718-2729.

Kassavetis, G. A., Braun, B. R., Nguyen, and Geiduschek, E. P. (1990). *S. cerevisiae* TFIIIB is the transcription initiation factor proper of RNA polymerase III, while TFIIIA and TFIIIC are assembly factors. *Cell* **60**: 235.

Kassavetis, G. A., and Geiduschek, E. P. (1993). RNA polymerase marching backward. *Science* **259**: 944-945.

Kassavetis, G. A., Bardeleben, C., Bartholomew, B., Braun, B. R., Joazeiro, C. A. P., Pisano, M., and Geiduschek, E. P. (1994). Transcription by RNA polymerase III. In *Transcription: Mechanism and Regulation.* New York, Raven Press. Ltd. 107-126.

Kelly, W., G,, Dahmus, M. E., and Hart, G. W. (1993). RNA polymersae II as a glycoprotein: modification of the C-terminal domain by O-GlcNAc. *Journal of Biological Chemistry* **268**: 10416-10424.

Kerppola, T. K., and Kane, C. M. (1990). Analysis of the signals for transcription termination by purified RNA polymerase II. *Biochemistry* **29**: 269-278.

Kibel, A., Iliopoulos, O., DeCaprio, J. A., and Kaelin, W. G. (1995). Binding of the von Hippel-Lindau tumor suppressor protein to elongin B and C. *Science* **269**: 1444-1446.

Kobayashi, M., Nagata, K., and Ishihama, A. (1990). Promoter selectivity of Escherichia coli RNA polymerase: effect of base substitutions in the promoter -35 region on promoter strength. *Nucleic Acids Research* **18**: 7367-7372.

Koleske, A. J., and Young, R. A. (1994). An RNA polymerase II holoenzyme responsive to activators. *Nature* **368**: 466-469.

Kollmar, R., and Farnham, P. (1993). Site-specific initiation of transcription by RNA polymerse II. *P.S.E.B.M.* **203**: 127-139.

Kolodziej, P. A., and Young, R. A. (1989). RNA polymerase II subunit RPB3 is an essential component of the mRNA transcription apparatus. *Molecular Cellular Biology* **9**: 5387.

Krummel, B., and Chamberlin, M. (1989). RNA chain Initiation by Escherichia coli RNA polymerase. Structural transitions of the early ternary complex. *Biochemistry* **28**: 7829-7842.

Krummel, B., and Chamberlin, M. J. (1992). Structural analysis of ternary complexes of *Escherichia coli* RNA polymerase. *Journal of Molecular Biology* **225**: 221-237.

Kuhn, A., Bartsch, I., and Grummt, I. (1990). Specific interaction of the murine transcription termination factor TTF I with class-I RNA polymersaes. *Nature* **344**: 559-562.

Kuhn, A., and Grummt, I. (1989). 3' end formation of mouse pre-rRNA involves both transcription termination and a specific processing reaction. *Genes & Development* **3**: 224-231.

Kutsu, S., Santero, E., Keener, J., Popham, D., and Weiss, D. (1989). Expression of sigma 54 -(*ntrA*)-dependent genes is probably united by a common mechanism. *Microbiol. Rev.* **53**: 367-376.

Labhart, P. (1997). Transcript cleavage in an RNA polymerase I elongation complex. *Journal of Biological Chemistry* **272**: 9055-9061.

Landick, R., and Yanofsky, C. (1987). Transcription attenuation. In *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*. Washington, D.C., American Society for Microbiology. 1276-1301.

Lee, D. N., Phung, L., Stewart, J., and Landick, R. (1990). Transcription Pausing by *Escherichia coli* RNA Polymerase Is Modulated by Downstream DNA Sequences. *Journal of Biological Chemistry* **265**: 15145-15153.

Levin, J. R., Krummel, B., and Chamberlin, M. J. (1987). Isolation and properties of transcribing ternary complexes of *Escherichia coli* RNA polyemrase positioned at a single nucleotide base. *Journal of Molecular Biology* **196**: 85-100.

Maicas, E., and Friesen, J. (1990). A sequence pattern that occurs at the transcription initiation region of yeast RNA polymerase II promoters. *Nucleic Acids Research* **18**: 3387-3393.

Maraia, R. J., Kenan, D. J., and Keene, J. D. (1994). Eukaryotic transcription termination factor La mediatess transcript release and facilitates reinitiation by RNA polymerase III. *Molecular and Cellular Biology* **14**: 2147-2158.

Markovtsov, V., Mustaev, A., and Goldfarb, A. (1996). Protein-RNA interactions in the active center of transcription elongation complex. *Proceedings of the National Academy of Sciences USA* **93**: 3221-3226.

Masters, B. S., Stohl, L. L., and Clayton, D. A. (1987). Yeast mitochondrial RNA polymerase is homologous to those encoded by bacteriophages T3 and T7. *Cell* **51**: 89.

McCracken, S., Fong, N., Yankulov, K., Ballantyne, S., Pan, G., Greenblatt, J., Patterson, S. D., Wickens, M., and Bentley, D. L. (1997). The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. *Nature* **385**: 357-361.

McKune, K., Richards, K. L., Edwards, A. M., Young, R. A., and Woyckik, N. A. (1993). RPB7, One of Two Dissociable Subunits of Yeast RNA Polymerase II, is Essential for Cell Viability. *Yeast* **9**: 1993.

Mote, J., Ghanouni, P., and Reines, D. (1994). A DNA minor groove-binding ligand both potentiates and arrests transcription by RNA polymerase II. Elogation factor SII enables readthrough at arrest sites. *Journal of Molecular Biology* **236**: 725-737.

Nakanishi, T., Nakano, A., Nomura, K., Sekimizu, K., and Natori, S. (1992). Purification, gene cloning, and gene disruption of the transcription elongation factor S-II in *Saccharomyces cerevisiae*. *Journal of Biological Chemistry* **267**: 13200-13204.

Nakanishi, T., Shimoaraiso, M., Kubo, T., and Natori, S. (1995). Structure-function relationship of yeast S-II in terms of stimulation of RNA polymerase II, arrest relief, suppression of 6-azauracil sensitivity. *Journal of Biological Chemistry* **270**: 8991-8995.

Nam, S.-C., and Kang, C. (1988). Transcription initiation site selection and abortive initiation cycling of phage SP6 RNA polymerase. *Journal of Biological Chemistry* **263**: 18123-18127.

Nonet, M., Sweetser, D., and Young, R. A. (1987). Functional redundancy and structural polymorphism in the large subunit of RNA polymerase II. *Cell* **50**: 909.

Nudler, E., Ekaterina, A., Markovtsov, V., and Goldfarb, A. (1996). Transcription processivity: Protein-DNA interactions holding together the elongation complex. *Science* **273**: 211-217.

Nudler, E., Goldfarb, A., and Kashlev, M. (1994). Discontinuous Mechanism of Transcription Elongation. *Science* **265**: 793-799.

Nudler, E., Mustaev, A., Lukhtanov, E., and Goldfarb, A. (1997). The RNA-DNA hybrid maintains the register of transcription by preventing backtracking of RNA polymerase. *Cell* **89**: 33-41.

Oakley, J. L., and Coleman, J. E. (1977). Structure of a promoter for T7 RNA polymerase. *Proceedings of the National Academy of Sciences USA* **74**: 4266-4270.

Orlova, M. (1995). *Intrinsic transcript cleavage activity of RNA polymerase.*, Ph.D., Institute of molecular genetics, Russian Academy of Science, Moscow.

Orlova, M., Newlands, J., Das, A., Goldfarb, A., and Borukhov, S. (1995). Intrinsic transcript cleavage activity of RNA polymerase. *Proceedings of the National Academy of Sciences USA* **92**: 4596-4600.

Parsons, M. C., and Weil, P. A. (1990). Purification and characterization of Saccharomyces cerevisiae transcription factor TFIIIC. *Journal of Biological Chemistry* **265**: 5095.

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

Pizzagalli, A., Valsasnini, P., Plevani, P., and Lucchini, G. (1988). DNA polymerase I gene of *Saccharomyces cerevisiae*: nucleotide sequence, mapping of a temperature-sensitive mutation, and protein homology with other DNA polymerases. *Proceedings of the National Academy of Sciences USA* **85**: 3772-3776.

Polyakov, A., Severinova, E., and Darst, S. A. (1995). Three-dimensional structure of E. coli core RNA polymerase: promoter binding and elongation conformations of the enzyme. *Cell* **83**: 365-373.

Powell, W., Bartholomew, B., and Reines, D. (1996). Elongation factor SII contacts the 3'-end of RNA in the RNA polymerase II elongation complex. *Journal of Biological Chemistry* **271**: 22301-22304.

Pugh, B. F., and Tjian, R. (1991). Transcription from a TATA-less promoter requires a multisubunit TFIID complex. *Genes & Development* **5**: 1935-1945.

Puhler, G., Leffers, H., Gropp, F., Klenk, H.-P., Lottspeich, F., Garrett, R. A., and Zillig, W. (1989). Archaebacterial DNA-dependent RNA polymerases testify to the evolution of the eukaryote nuclear genome. *Proceedings of the national Academy of Sciences USA* **86**: 4569-4573.

Qian, X., Gozani, S. N., Yoon, H., Jeon, C., Agarwal, K., and Weiss, M. A. (1993a). Novel Zinc Finger Motif in the Basal Transcription Machinery: Three-Dimensional NMR Studies of the Nucleic Acid Binding Domain of Transcriptional Elongation Factor TFIIS. *Biochemistry* **32**: 9944-9959.

Qian, X., Jeon, C., Yoon, H., Agarwal, K., and Weiss, M. (1993b). Structure of a new nucleic-acid-binding motif in eukaryotic transcriptional elongation factor TFIIS. *Nature* **365**: 277-279.

Rappaport, J., Cho, K., Saltzman, A., Prenger, A., Golomb, M., and Weinmann, R. (1988). Transcription elongation factor SII interacts with a domain of the large subunit of human RNA polymerase II. *Molecular Cellular Biology* **8**: 3136-3142.

Reinberg, D., and Roeder, R. G. (1987). Factors involved in specific transcription by mammalian RNA polymerase II. *Journal of Biological Chemistry* **262**: 3331-3337.

Reines, D. (1992). Elongation factor-dependant transcript shortening by template-engaged RNA polymerase II. *Journal of Biological Chemistry* **267**: 3795-3800.

Reines, D. (1994). Nascent RNA cleavage by transcription elongation complexes. In *Transcription: Mechanism and Regulation*. New York, Raven Press. 263-278.

Reines, D., Chamberlin, M. J., and Kane, C. M. (1989). Transcription elongation factor SII (TFIIS) enables RNA Polymerse II to elongate through a block to transcription in a human gene in vitro. *Journal of Biological Chemistry* **264**: 10799-10809.

Rice, G. A., Kane, C. M., and Chamberlin, M.J. (1991). Footprinting analysis of mammalian RNA polymerase II along its transcript: An alternate view of transcription elongation. *Proceedings of the national Academy of Sciences USA* **88**: 4245-4249.

Riva, M., Schaffner, A. R., Sentena, A., Hartmann, G. R., Mustaev, A. A., Zaychikov, E. F., and Grachev, M. A. (1987). Acitive site labelling of the RNA polymerases A, B, and C from yeast. *Journal of Biological Chemistry* **265**: 14377-14380.

Roeder, R. G., and Rutter, W. J. (1969). Multiple forms of DNA-dependent RNA polymerase in eukaryotic organisms. *Nature* **224**: 234.

Rudd, M. D., Izban, M. G., and Luse, D. S. (1994). The active site of RNA polymerase II participates in transcript cleavage within arrested ternary complexes. *Proceedings of the National Academy of Sciences USA* **91**: 8057-8061.

Sadhale, P. P., and Woychik, N. A. (1994). C25, an essential RNA polymerase II subunit related to the RNA polymerase subunit RPB7. *Molecular and Cellular Biology* **14**: 6164-6170.

Sawadogo, M., Lescure, M., Sentenac, A., and Fromageot, P. (1981). Native deoxyribonucleic acid transcription by yeast RNA polymerase-P37 complex. *Biochemistry* **20**: 3542-3547.

Sawadogo, M., Sentenac, A., and Fromageot, P. (1980). Interaction of a new polypeptide with yeast RNA polymerase B. *Journal Biological Chemistry* **255**: 12.

Schaefer, L., Roy, R., Humbert, S., Moncollin, V., Vermeulen, W., Hoeijmakers, J. H., Chambon, P., and Egly, J. M. (1993). DNA repair helicase: a component of BTF2 (TFIIH) basic transcription factor. *Science* **260**: 58-63.

Schultz, P., Celia, H., Riva, M., Sentenac, A., and Oudet, P. (1993). Three-dimensional model of yeast RNA polymerase I determined by electron microscopy of two-dimensional crystals. *EMBO* **12**: 2601-2607.

Sekimizu, K., Kobayashi, N., Mizuno, D., and Natori, S. (1976). Purification of a factor from Ehrlich ascites tumor cells specifically stimulating RNA polymerase II. *Biochemistry* **15**: 5064-5070.

Sentenac, A., Riva, M., Thuriaux, P., Buhler, J.-M., Treich, I., Carles, C., Werner, M., Ruet, A., Huet, J., Mann, C., Chiannilkulchai, N., Stettler, S., and Mariotte, S. (1992). Yeast RNA Polymerase Subunits and Genes. In *Transcriptional Regulation*. Cold Spring Harbor, NY, Cold Sprng Harbor Laboratory Press. 27-54.

Serizawa, H., Conaway, J. W., and Conaway, R. C. (1994). Transcription Initiation by Mammalian RNA Polymerase II. In *Transcription: Mechanism and Regulation.* New York, Raven Press. 27-43.

Shilatifard, A., Lane, W. S., Jackson, K. W., Conaway, R. C., and Conaway, J. W. (1996). An RNA polymerase II elongation factor encoded by the human ELL gene. *Science* **271**: 1873-1876.

Sousa, R., Chung, Y. J., Rose, J. P., and Wang, B. C. (1993). Crystal structure of bacteriophage T7 RNA polymerase at 3.3 A resolution. *Nature* **364**: 593-599.

Spencer, C. A., and Groudine, M. (1990). Transcription elongation and eukaryotic gene regulation. *Oncogene* **5**: 777-785.

Stebbins, C., Borukhov, S., Orlova, M., Polyakov, A., Goldfarb, A., and Darst, S. (1995). Crystal structure of the GreA transcript cleavage factor from *Escherichia coli*. *Nature* **373**: 636-640.

Svejstrup, J. Q., Feaver, W. J., LaPointe, J., and Kornberg, R. D.(1994). RNA Polymerase II Transcription Factor IIH Holoenzyme From Yeast. *Journal of Biological Chemistry* **269**: 28044-28048.

Wang, D., Meier, T., Chan, C., Feng, G., Lee, D., and Landick, R. (1995). Discontinuous movements of DNA and RNA in RNA polymerse accompany formation of a paused transcription complex. *Cell* **81**: 341-350.

Whitehall, S. K., Bardeleben, C., and Kassavetis, G. A. (1994). Hydrolytic cleavage of nascent RNA in RNA polymerase III ternary transcription complexes. *Journal of Biological Chemistry* **269**: 2299-2306.

Woychik, N., and Young, R. A. (1993). Yeast RNA Polymerase II Subunit RPB11 is Related to a Subunit Shared by RNA Polymerase I and III. *Gene Expression* **3**: 77-82.

Woychik, N. A., Liao, S.-M., Kolodziej, P. A., and Young, R. A. (1990). Subunits shared by eukaryotic nuclear RNA polymerases. *Genes & Development* **4**: 313-323.

Woychik, N. A., and Young, R. A. (1989). RNA polymerase II subunit RPB4 is essential for high- and low-temperature yeast cell growth. *Molecular and Cellular Biology* **9**: 2854-2859.

Woychik, R. A., Lane, W. S., and Young, R. A. (1991). Yeast RNA polymerase II subunit RPB9 is essential for growth at temperature extremes. *Journal of Biological Chemistry* **266**: 19053.

Wu, J., Awrey, D. E., Edwards, A. M., Archambault, J., and Friesen, J. D.(1996). In vitro characterization of mutant yeast RNA polymerase II with reduced binding for elongation factor TFIIS. *Proceedings of the National Academy of Sciences USA* **93**: 11552-11557.

Yankulov, K., Blau, J., Purton, T., Roberts, S., and Bentley, D. L. (1994). Transcriptional elongation by RNA polymerase II is stimulated by transactivators. *Cell* **77**: 749-759.

Yankulov, K. Y., Pandes, M., McCracken, S., Bouchard, D., and Bentley, D. L. (1996). TFIIH functions in regulating transcriptional elongation by RNA polymerase II in Xenopus oocytes. *Molecular and Cellular Biology* **16**: 3291-3299.

Zaychikov, E., Martin, E., Denissova, L., Kozlov, M., Markovtsov, V., Goldfarb, A., and Mustaev, A. (1996). Mapping of catalytic residues in the RNA polymerase active center. *Science* **273**: 107-109.

Zhou, Q., Lieberman, P. M., Boyer, T. G., and Berk, A. J. (1992). Holo-TFIID supports transcriptional stimulation by diverse activators and from a TATA-less promoter. *Genes & Development* **6**: 1-10.

Zhu, W., Zang, Q., Colangelo, C., Lewis, M., Summers, M., and Scott, R. (1996). The N-terminal domain of TFIIB from Pyrococcus furiosus forms a zinc domain. *Nature Structural Biology* **3**: 122-124.

Appendix I

Purified Yeast RNA Polymerase II Reads through Intrinsic Blocks to Elongation in Response to the Yeast TFIIS Analogue, P37.

Christie, K.R., Awrey, D.E., Edwards, A.M., and Kane, C.M. (1994).
Journal of Biological Chemistry, vol. 269, pp. 936-943.

THE JOURNAL OF BIOLOGICAL CHEMISTRY

Vol. 269, No. 2, Issue of January 14, pp. 936-943, 1994
Printed in U.S.A.

Purified Yeast RNA Polymerase II Reads through Intrinsic Blocks to Elongation in Response to the Yeast TFIIS Analogue, P37*

(Received for publication, July 26, 1993, and in revised form, September 15, 1993)

Karen R. Christie‡§, Donald E. Awrey¶, Aled M. Edwards¶‖, and Caroline M. Kane‡**

From the ‡Division of Biochemistry and Molecular Biology, University of California, Berkeley, California 94720 and the ¶Departments of Pathology and Biochemistry, McMaster University, Hamilton, Ontario, Canada L8N 3Z5

***Saccharomyces cerevisiae* has a TFIIS-related transcription elongation factor, originally called P37 (Sawadogo, M., Sentenac, A., and Fromageot, P. (1979) *J. Biol. Chem.* 255, 12–15; Nakanishi, T., Nakano, A., Nomura, K., Sekimizu, K., and Natori, S. (1992) *J. Biol. Chem.* 267, 13200–13204), which binds directly to RNA polymerase II and stimulates read-through of intrinsic blocks to elongation. To elucidate functional features of this protein:protein interaction, we tested the ability of several forms of RNA polymerase II to respond to either full-length or an amino-terminal truncation of TFIIS. The variants of the polymerase differed in the structure of the carboxyl-terminal domain of the largest subunit or lacked two of the smaller subunits. No differences in ability to recognize intrinsic blocks to elongation or to read through them in response to either form of TFIIS were detected among these variants.**

Furthermore, ternary complexes containing each variant form of RNA polymerase cleave the 3′ end of the nascent transcripts in response to TFIIS, a reaction previously reported for mammalian and *Drosophila* TFIIS (Kassavetis, G. A., and Geiduschek, E. P. (1993) *Science* 259, 944–945) and likely to be important in TFIIS function. Thus the carboxyl-terminal domain of the largest subunit and subunits four and seven of the polymerase, required *in vivo*, are not required *in vitro* for recognition of intrinsic blocks to elongation, read-through in response to TFIIS, or TFIIS-stimulated cleavage of the nascent transcript.

Many eukaryotic genes contain regulated blocks to transcript elongation within the transcription unit (12, 13). Generally, these blocks to elongation are found early in the transcription unit, often within the first intron in mammalian genes. When the polymerase stops at such sites, expression of the gene is prevented. For the gene product to be expressed, the polymerase must read through these sites and synthesize full-length RNA. One strategy that might permit synthesis of full-length RNA involves a protein factor (TFIIS) that promotes read-through of such blocks to elongation (12, 14–19).

TFIIS-related proteins which stimulate elongation by RNA polymerase II have been purified and the genes cloned from a variety of organisms (1–5, 18, 20–26).[1] Sequence analysis of these clones predicts significant amino acid similarities, especially in the COOH-terminal region of the proteins (20). In most cases, these TFIIS-related proteins have also been shown to promote read-through of specific blocks to elongation *in vitro* (14–19).

The mechanism by which these factors stimulate elongation and promote read-through is not understood, but some important features of the reaction have been described. The TFIIS-related proteins can bind to RNA polymerase II (1, 15, 18, 27, 28), although some data suggest that they do not remain associated throughout elongation (18). Mutants in mouse SII (29) and human TFIIS (27) unable to bind the polymerase neither stimulate elongation nor promote read-through.

Recent work with mammalian and *Drosophila* RNA polymerase II has shown that ternary elongation complexes blocked in elongation cleave the 3′ end of the nascent transcript in response to TFIIS (6–11). Following cleavage, the 5′ fragment is retained in the ternary complex and can be elongated. This cleavage precedes and may be necessary for efficient read-through.

To dissect the molecular mechanism by which this factor stimulates elongation and promotes transcript cleavage, it will be important to define the binding interaction between TFIIS and RNA polymerase. An identification of the subunits of RNA polymerase necessary for TFIIS activity would permit an analysis of the more specific contacts. This problem is complex; TFIIS apparently can function as a single polypeptide (11, 17, 26, 27, 30), but catalytically active RNA polymerase II purifies as a complex of 9–12 polypeptides (31, 32).

Immunological, genetic, and biochemical experiments have been reported which begin to assess the RNA polymerase subunits involved in TFIIS function. Much of this work has pointed to the largest subunit of RNA polymerase II. Antibodies directed against a fusion peptide containing sequences from the largest subunit of mammalian RNA polymerase II interfere with TFIIS function (33). Conditional mutations in a different region of the largest subunit of the yeast RNA polymerase II are complemented *in vivo* by overproduction of TFIIS protein (4). Also, the yeast TFIIS analogue apparently binds more efficiently to RNA polymerase II when its largest subunit contains the highly conserved COOH-terminal domain (CTD)[2] of heptapeptide repeats (1). This domain, unique to RNA polymerase II (34, 35), is often removed by proteolysis during purification to generate form IIb (31). In this work, we examine the TFIIS-stimulated read-through activity of purified RNA polymerase II with or without the conserved COOH-terminal domain of the largest subunit.

* This work was supported by Grant DM9205583 from the National Science Foundation (to C. M. K.) and a by grant from the Medical Research Council of Canada (to A. M. E.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ Supported in part by a National Defense Science and Engineering Graduate Fellowship.

‖ Medical Research Council of Canada Scholar.

** To whom correspondence should be addressed. Tel.: 510-642-4118; Fax: 510-642-7846.

[1] A. Guyonvarch, A. Ruet, and F. LaCroute, personal communication.

[2] The abbreviations used are: CTD, conserved carboxyl-terminal domain; kb, kilobase(s); ORF, open reading frame; PCR, polymerase chain reaction; DTT, dithiothreitol.

Subunits other than the largest subunit may also be important for TFIIS function. In particular, the yeast TFIIS analogue, P37, can protect against inhibition of transcription by antibodies to the 23-kDa subunit of yeast RNA polymerase II (2). Here, we test the requirement for two different small subunits, designated four and seven, in TFIIS function; deletion of the gene for subunit seven renders yeast cells inviable (36). Purified yeast RNA polymerase II contains substoichiometric amounts of these two subunits (37, 38). Although they are required for promoter-specific initiation *in vitro*, they are not required for efficient elongation or recognition of some specific blocks to elongation *in vitro* (38). However, it was not known whether TFIIS could stimulate polymerase lacking these subunits to read through these blocks to elongation.

One such block to elongation falls within the first intron of the human histone H3.3 gene (39), and mammalian TFIIS stimulates mammalian polymerase to read through this block (16, 17). Thus a comparison of the effect of TFIIS on read-through of this site with each variant of yeast RNA polymerase II could establish which structural features of the polymerase are required for TFIIS function.

EXPERIMENTAL PROCEDURES

Materials—Inhibit-ACE was obtained from 5′ → 3′. Phosphocellulose (P11) and DEAE-cellulose (DE52) were obtained from Whatman. CM-Sephadex, C-25 was obtained from Pharmacia LKB Biotechnology Inc. Nucleotides were obtained from Pharmacia. [α-^{32}P]CTP, >400 Ci/mmol, was obtained from Amersham Corp. Bio-Gel 30 columns were obtained from Bio-Rad.

RNA polymerase II forms IIa and Δ4,7 were purified by immunoaffinity as described (40, 41). RNA polymerase IIb, (ΔCTD) was the generous gift of Y. Li and R. Kornberg, Stanford University. It was prepared by treating enzyme containing a recombinant largest subunit with Factor Xa; this treatment removed the CTD from the largest subunit and generated form IIb.[3] The specific activities of the IIa and Δ4,7 polymerases were comparable. The specific activity of the ΔCTD form was approximately 2–4-fold lower, probably due in part to the Factor Xa cleavage conditions used to generate this polymerase (data not shown).

Plasmid Constructions—The 3.7-kb *Saccharomyces cerevisiae* genomic *Bam*HI fragment containing the PPR2 gene, encoding TFIIS, from pFL44D (generously provided by A. Guyonvarch, A. Ruet, and F. LaCroute) was ligated into the *Bam*HI site of pBS.KS$^+$ (Stratagene) to generate pKC3. Plasmids were then constructed encoding fusion proteins between an amino-terminal histidine-rich region with either the full-length open reading frame (ORF) of TFIIS or a 113-amino acid amino-terminal deletion of the ORF.

Fragments containing either the full-length or truncated ORF were prepared from pKC3 using the polymerase chain reaction (PCR). The 5′ primer sequences were (5′ CGCGCACCATATGGATAGTAAGGAAGTACTGGTA 3′) for the full-length and (5′ CCTCACCATATGGCCTCCTCCCAGTCAGAT 3′) for the truncated ORF. The 3′ primer sequence (5′ GCGAGCCGGATCCCCTTCCTCCTATTGTTTCCTT 3′) was identical for both PCR reactions.

The full-length reading frame PCR product was digested with *Nde*I and *Bam*HI, and the fragment was ligated into the 5.7-kb *Nde*I-*Bam*HI-digested pET15b (Novagen). This plasmid is referred to as pET15bPPR2.

The truncated reading frame PCR product was digested with *Bam*HI, and the fragment was ligated into the 3.0-kb *Bam*HI-*Sma*I-digested pGEM7Zf(+) (Promega). Next, the 0.6-kb *Nde*I-*Bam*HI fragment from the pGEM-derived construct was ligated into pET15b, and this plasmid is referred to as pET15bΔ1–113PPR2.

Purification of Recombinant TFIIS and Δ1–113TFIIS—pET15bPPR2 was transformed into *Escherichia coli* BL21 (DE3) cells which were then grown to an OD_{600} of 0.8–1.0 in 12 liters of Terrific broth (TB, 1.2% bactotryptone, 2.4% yeast extract, 0.4% glycerol, 90 mM phosphate, pH 7.8). Isopropyl 6-D-thiogalactopyranoside was than added to a final concentration of 0.2 mM. The cells were grown for an additional 3 h before harvesting and were resuspended in 50 mM Tris-HCl, pH 7.5, and 10% sucrose (2 ml/g of cells). The cells were lysed with a French press in 50 mM Tris-HCl, pH 7.5, 10% sucrose, 250 mM NaCl, 1 mM EDTA, 1 mM DTT, in the presence of protease inhibitors (1 mM phenylmethylsulfonyl fluoride, 1 mM benzamidine, 1 μg/ml leupeptin, 1 μg/ml pepstatin). The supernatant fluid was clarified by centrifugation and then diluted 2.5-fold with 50 mM Tris-HCl, pH 7.5, 1 mM EDTA, 1 mM DTT, in the presence of the protease inhibitors indicated above. Polyethyleneimine was added to a final concentration of 0.3% (v/v), and the solution was mixed gently for 15 min at 4 °C. The suspension was centrifuged in a Sorvall GSA rotor at 11,000 rpm for 20 min at 4 °C. The supernatant solution was loaded onto a 5 × 5-cm DE52 (Whatman) column, equilibrated in 20 mM Tris-HCl, pH 7.5, 100 mM NaCl, 10% sucrose, 1 mM EDTA, 1 mM DTT containing the protease inhibitors indicated above. The flow-through was made 200 mM in NaCl and loaded onto a 5 × 10-cm SP Fast Flow (Pharmacia) column. Protein was eluted with 20 mM Tris-HCl, pH 7.5, 500 mM NaCl, 1 mM EDTA, 1 mM DTT, in the presence of the protease inhibitors indicated above. The protein was precipitated with 65% ammonium sulfate, and the pellet was resuspended in 20 mM HEPES, pH 7.2, 30 mM ammonium sulfate, 5% glycerol, 1 mM EDTA, 1 mM DTT, containing protease inhibitors; the concentration of ammonium sulfate was adjusted to 1.5 M with saturated ammonium sulfate.

The protein suspension was clarified by centrifugation in a Sorvall SS34 rotor at 10000 rpm for 20 min at 4 °C and was loaded onto a phenyl-Sepharose column (Pharmacia XK 26/10). TFIIS was eluted with a 250-ml gradient of decreasing ammonium sulfate (1.5–0.03 M); the TFIIS peak appeared at 1.1 M ammonium sulfate. TFIIS-containing fractions were dialyzed to 20 mM HEPES, pH 7.9, 0.5 M NaCl, containing protease inhibitors. The dialysate was applied to a 5-ml Poros MC (metal chelate column, PerSeptive Biosystems, Cambridge, MA) charged with Ni^{2+} and was eluted with a 40-ml imidazole gradient (0.5–0.3 M imidazole). Fractions containing TFIIS were dialyzed against 20 mM HEPES, pH 7.5, 100 mM NaCl, 1 mM EDTA, 1 mM DTT, and 10% glycerol before loading onto a Mono S column (Pharmacia HR 5/5). TFIIS was eluted with a 10-ml gradient (100 mM to 1.0 M NaCl).

pET15bΔ1–113PPR2, encoding the truncated derivative of TFIIS, was transformed into *E. coli* BL21 (DE3) cells which were then grown to an OD_{600} 0.8–1.0 in 12 liters of Luria-Bertani broth. The cells were induced, harvested, and lysed under the conditions described above for the full-length TFIIS. The purification scheme of the truncated derivative was identical to that described above through the phenyl-Sepharose column. Fractions from that column containing truncated TFIIS were dialyzed against 20 mM Tris-HCl, pH 7.5, 100 mM NaCl, 5% glycerol, 1 mM EDTA, 1 mM DTT, containing protease inhibitors. The dialysate was loaded onto a Bio-Rad SP-5 PW column, and truncated TFIIS was eluted with the same buffer lacking EDTA and DTT. The eluted protein was brought to 0.5 M NaCl and loaded onto a 2.5 × 4.5-cm His-Bind column charged with Ni^{2+} (Novagen 69670–2). Fractions containing TFIIS were dialyzed into 20 mM Tris-HCl, pH 7.5, 100 mM potassium acetate, 10% glycerol, 1 mM EDTA, and 1 mM DTT and were loaded onto a heparin column (Supelco Heparin-5PW1, Oakville, Ontario, Canada). Truncated TFIIS was eluted with a 20-ml gradient of potassium acetate (0.1–1.0 M potassium acetate). Throughout the purification, the presence of full-length or truncated TFIIS was monitored using SDS-polyacrylamide gel electrophoresis.

Preparation from S. cerevisiae of TFIIS/P37-containing Fraction—A partially purified fraction of TFIIS/P37 protein was obtained essentially as described (1). Briefly, a clarified cell lysate was prepared from 80 g of *S. cerevisiae* strain CBO18 (the generous gift of Robert Fuller, Stanford University), and partial fractionation was carried out by ammonium sulfate precipitation. Proteins were fractionated by column chromatography as described (1). Active fractions were pooled, and glycerol was added to 15%. The pooled fraction was stored at −80 °C. Activity was monitored using the stimulation assay as described (1).

SDS-Polyacrylamide Gel Electrophoresis—RNA polymerase II preparations were resolved on 7–12% acrylamide gradient gels. Preparations of TFIIS were resolved on 12% acrylamide gels. Proteins were visualized by silver staining.[4]

Read-through Assay—The ability of TFIIS protein to stimulate read-through by RNA polymerase II stalled at intrinsic blocks to elongation was assayed using a 3′-deoxycytidine-extended template (42) of the *Taq*I fragment containing the human histone H3.3 gene (39). In this template are three characterized blocks to elongation, TII, TIb, and TIa (39, 43). Reactions (50 μl) in transcription buffer (60 mM Tris-OAc, pH 8; 5 mM MgOAc; 5% glycerol; 100 mM $(NH_4)_2SO_4$; 6 mM spermidine HCl; 0.8 mM each ATP, GTP, and UTP; 1 μM [α-^{32}P]CTP; 1 unit of Inhibit-ACE; and yeast RNA polymerase II) were incubated for 1 min at 30 °C. A 5-μl aliquot was diluted into 100 μl of stop buffer (20 mM EDTA, 0.5 M NH_4OAc, and 12 μg/ml carrier RNA) on ice for analysis. At 1 min 15

[3] Y. Li and R. Kornberg, personal communication.

[4] E. Rosenberg and H. Nikaido, unpublished procedure.

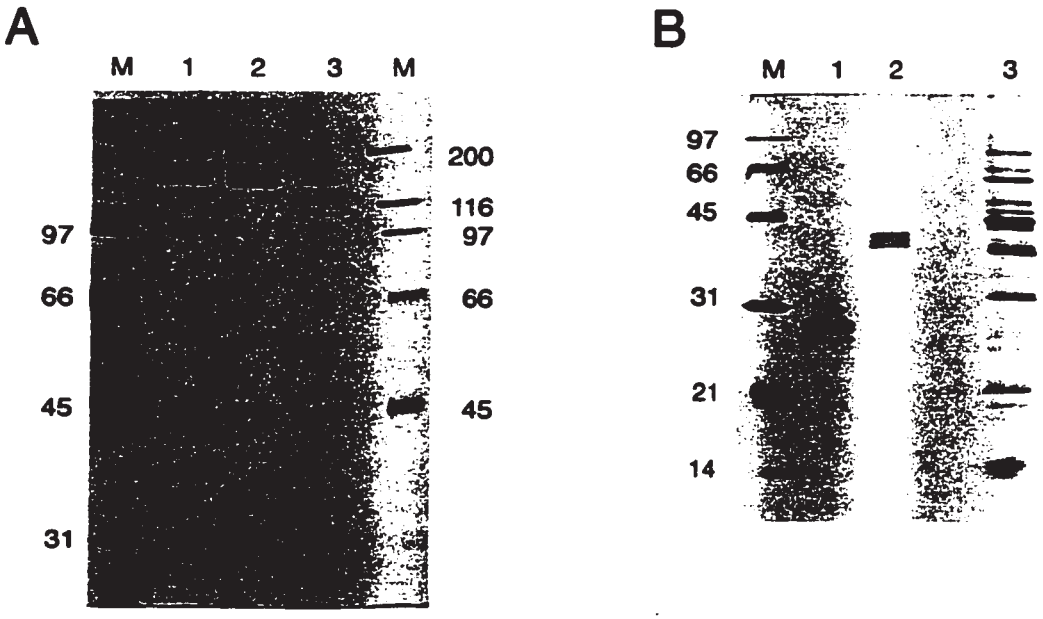


FIG. 1. **Protein preparations used in transcription assays.** *A*, various forms of yeast RNA polymerase II used in transcription assays were resolved on a 7–12% gradient SDS-PAGE gel. *Lane 1*, 1 μg of yeast RNA polymerase II, form Δ4,7. *Lane 2*, 1 μg of yeast RNA polymerase II, form IIa. *Lane 3*, 0.5 μg of yeast RNA polymerase II, form ΔCTD. Marker lanes contain either silver stain low molecular weight protein standards or silver stain high molecular weight protein standards from Bio-Rad, sizes are in kilodaltons, as indicated. *B*, various forms of yeast TFIIS used in transcription assays were resolved on a 12% SDS-PAGE gel. *Lane 1*, 0.5 μg of truncated (Δ1–113) TFIIS. *Lane 2*, 0.25 μg of full-length TFIIS. *Lane 3*, 1.2 μg of protein from partially purified TFIIS/P37-containing fraction, used in the experiment of Fig. 2. Marker lane contains silver stain low molecular weight protein standards from Bio-Rad, sizes are in kilodaltons, as indicated.

s, 30 μl was diluted 10 times into chase buffer (transcription buffer plus 100 μg/ml heparin and 0.1 mM CTP) and incubated for 45 s at 30 °C. This procedure labels the 5′-proximal region of all transcripts with equivalent amounts of ^{32}P such that quantitation of cpm in each transcript allows a direct comparison of the number of transcripts of each size that are present, regardless of length. Next, a 30-μl aliquot was diluted into 100 μl of stop buffer. The remainder of the reaction was divided into two parts (120 μl each), and incubation at 30 °C was continued with the addition of either TFIIS protein or TFIIS storage buffer (200 mM Tris-HCl, pH 8, 10 mM β-mercaptoethanol, 5 mM EDTA, 250 mM $(NH_4)_2SO_4$, and 15% glycerol). Aliquots (30 μl) were removed to 100 μl of stop buffer at various time points, and RNAs were ethanol-precipitated, resuspended in formamide load buffer (90 mM Tris borate, pH 8.0, 2 mM EDTA, 80% formamide, 0.13% bromphenol blue, 0.13% xylene cyanol, 0.06% SDS) and resolved by electrophoresis through 5% acrylamide, 8.3 M urea gels.

Ternary Complex Cleavage Assay—Bio-Gel 30 columns were equilibrated by overlaying each column with 200 μl of chase buffer lacking nucleotides and centrifuging for 2 min in the swinging bucket rotor of an IEC clinical centrifuge at a setting of 4 (approximately 1000 relative centrifugal force); this was done four times. Reactions (25 μl) were set up as described for the read-through assay and were incubated for 1 min at 30 °C. The reactions were then diluted 5-fold into chase buffer at 30 °C, and the incubation was continued for 1 additional min. This procedure produces a mixture of complexes stalled at TII, TIb, and TIa as well as run-off transcripts. These complexes were applied to Bio-Gel 30 columns prepared as described above and centrifuged for 2 min 30 s at a setting of 4. This was repeated with a second column, a procedure which almost completely removed unincorporated nucleotides from the isolated ternary complexes.

Isolated complexes were incubated at 30 °C in the presence of either TFIIS protein or TFIIS storage buffer, with or without nucleotides (0.8 mM). When α-amanitin was used, it was added to a final concentration of 100 μg/ml for at least 1 min prior to TFIIS addition. Aliquots were removed at various times and mixed with an equal volume of SDS/urea load buffer (10 M urea, 0.5% SDS, 90 mM Tris borate, pH 8.0, 10 mM EDTA, 0.03% bromphenol blue, 0.03% xylene cyanol). RNA products were resolved by electrophoresis through 5% acrylamide, 8.3 M urea gels.

Quantitation of Transcripts in Polyacrylamide Gels—Quantitation of radioactivity in polyacrylamide gels was performed using the Molecular Dynamics PhosphorImager system. Transcripts with 3′ ends at TII, TIb, TIa, and the run-off transcript were identified, and the cpm in each was determined. The total cpm in these four types of transcripts was determined for each time point and is called (Sum). The fraction of complexes stopped at TII was calculated as (TII cpm)/(Sum). The fraction stopped at TIb was calculated as (TIb cpm)/[(Sum) − (TII cpm)). The fraction stopped at TIa was calculated as (TIa cpm)/[(Sum) − (TIb cpm) − (TII cpm)].

RESULTS

TFIIS Activity in S. cerevisiae—A partially purified protein from *S. cerevisiae* (P37) had been shown to stimulate elongation by both RNA polymerase I and II in nonspecific assays (1–3). This activity was analogous to that of SII, a protein purified from mouse Ehrlich ascites cells, which stimulated the elongation reaction of mouse RNA polymerase II in nonspecific assays (44). The SII protein is a member of the TFIIS family of proteins, cloned from mouse (21), human (20, 26), and *Drosophila* (24). Each of these proteins promotes transcript read-through as well as nascent transcript cleavage (10). Protein sequences predicted from the gene encoding P37, PPR2 (5, 22),[1] place it into the TFIIS family. To see whether this protein was a functional homolog of TFIIS, the previously reported P37 purification (1) was used to generate a fraction enriched in TFIIS/P37 activity; about 10–15% of the protein in this preparation migrates as expected for the P37 protein (Fig. 1*B*, *lane 3*). This preparation stimulates yeast RNA polymerase II 2–5-fold in nonspecific transcription assays (data not shown). Furthermore, the stimulatory activity in this preparation is specific for yeast RNA polymerase II and does not stimulate *Drosophila* or calf thymus RNA polymerase II (data not shown).

The P37-containing fraction was tested for promoting read-through by RNA polymerase II using a 3′-deoxycytidine-extended template containing the three intrinsic blocks to elongation from the human histone H3.3 gene. Purified yeast RNA polymerase IIa recognizes and stops at these sites (38). Transcripts with 3′ ends at TII, TIb, and TIa are observed 45 s after the addition of chase mixture to initiated complexes (Fig. 2, *lane 2*). The TII and TIb sites stop the polymerase with low efficiency, 10% or less. However, the polymerase recognizes the TIa site with approximately 50% efficiency. During continued incubation, some ternary complexes stalled at the site resume elongation. However, over half of the complexes originally

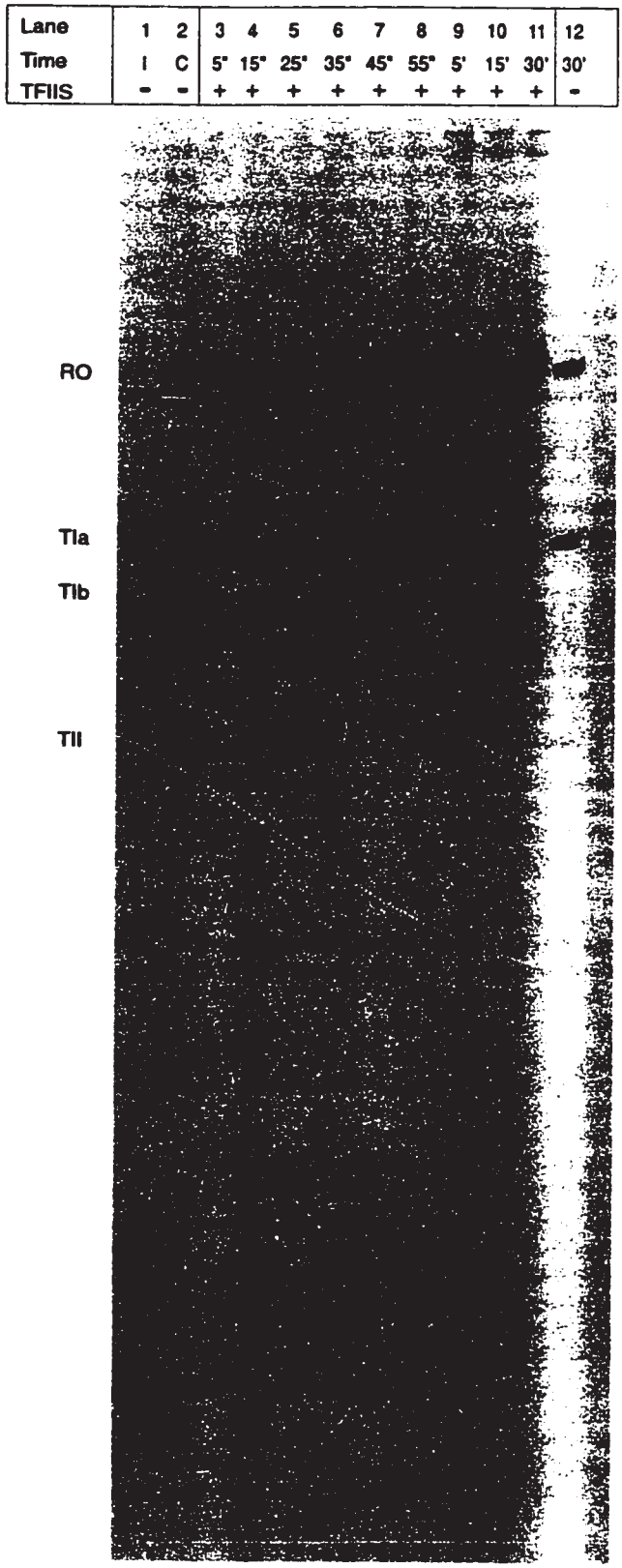


FIG. 2. **Time course of TFIIS-stimulated read-through of intrinsic blocks to elongation.** Ternary complexes containing the IIa form of RNA polymerase II were formed as described under "Experimental Procedures." An aliquot was removed and incubated with TFIIS storage buffer. The partially purified TFIIS/P37 containing fraction (300 ng of protein) was added to the remainder of the reaction, and aliquots were stopped at various times. *Lane 1*, RNA from initiated complexes; *lane 2*, RNA from ternary complexes chased to the TIa, TIb, and TII sites and the run-off transcript; *lanes 3–11*, RNA from ternary complexes incubated with the TFIIS/P37-containing fraction for the indicated times; *lane 12*, RNA from ternary complexes incubated in the absence of TFIIS/P37 for 30 min.

stopped at TIa do not read through the site even after 30 min of incubation at 30 °C (Fig. 2, *lane 12*). When TFIIS/P37 protein is added to ternary complexes stalled at these sites, transcript elongation begins within 5 min (Fig. 2, *lane 9*), and most complexes initially stalled at TIa elongate transcripts to the runoff within 30 min (Fig. 2, *lane 11*). Clearly, the P37-containing fraction exhibits a biochemical read-through activity similar to that observed with other members of the TFIIS family.

Read-through by Purified TFIIS—To more definitively assign the read-through activity to the TFIIS/P37 protein, the cloned sequences for this protein were expressed as an oligo-histidine fusion in *E. coli*. Similarly, a protein with a 113-amino acid amino-terminal deletion relative to the full-length TFIIS was expressed as an oligo-histidine fusion protein. The fusions facilitated purification and did not affect activity (data not shown).

The purified fusion proteins containing the full-length (Fig. 1*B*, *lane 2*) or the truncated TFIIS (Fig. 1*B*, *lane 1*) were used in combination with several variant forms of yeast RNA polymerase II (Δ4,7, IIa, ΔCTD; Fig. 1*A*). The ability of each variant to recognize the intrinsic sites TIa, TIb, and TII and to elongate through them in the presence or absence of the TFIIS proteins was compared. Data for the Δ4,7 form of the polymerase with both forms of TFIIS are displayed in Fig. 3; the quantitative analyses for all forms of the polymerase and TFIIS are displayed in Fig. 4. There are no major differences in the behavior of these different forms of polymerase. They are essentially identical in recognizing these intrinsic blocks to elongation and in read-through of these sites in response to TFIIS. All three polymerase variants recognize the TII and TIb sites with approximately 10% efficiency and the TIa site with about 50% efficiency. When either form of TFIIS is added, the complexes resume elongation, and within 30 min, 80% of the transcripts originally ending at TIa have been elongated. None of the transcripts which remain at TIa after 30 min is elongated during continued incubation, and it is likely that these transcripts have been released from the polymerase. In the absence of TFIIS, over half of the transcripts with 3′ ends at TIa have not been elongated, even after 50 min of incubation (Figs. 3 and 4).

Clearly, neither the COOH-terminal domain of the largest subunit nor subunits four or seven are required for TFIIS to interact with the polymerase and promote read-through of these intrinsic blocks to elongation *in vitro*. In addition, the amino-terminal 113 amino acids of the yeast TFIIS protein are unnecessary for read-through function *in vitro*. Proteins with amino-terminal truncations from human (27),[5] mouse (7, 45), and *Drosophila* (11) also have been shown to be functional in stimulating elongation by the cognate RNA polymerase II.

These results are both qualitative and quantitative. The kinetics of read-through as well as the stoichiometry of TFIIS to polymerase necessary to promote read-through are identical in all tested combinations (Fig. 4 and data not shown). The difference in kinetics observed in Fig. 4 is due to the different ratios of TFIIS to polymerase used in each series of experiments. When comparable ratios are used, the read-through kinetics are identical for the two forms of TFIIS. At a molar ratio of 5:1 TFIIS to RNA polymerase II protein, the maximum rate of read-through is obtained. Additional TFIIS does not accelerate read-through (data not shown).

Cleavage of Nascent Transcripts in Ternary Complexes—Mammalian and *Drosophila* TFIIS proteins stimulate cleavage of the nascent transcript in stalled ternary complexes containing RNA polymerase II (6–11). A similar cleavage reaction has been observed in ternary complexes formed with vaccinia virus RNA polymerase (46) which contains a subunit with sequence

[5] G. Ciprés-Palacín and C. Kane, unpublished results.

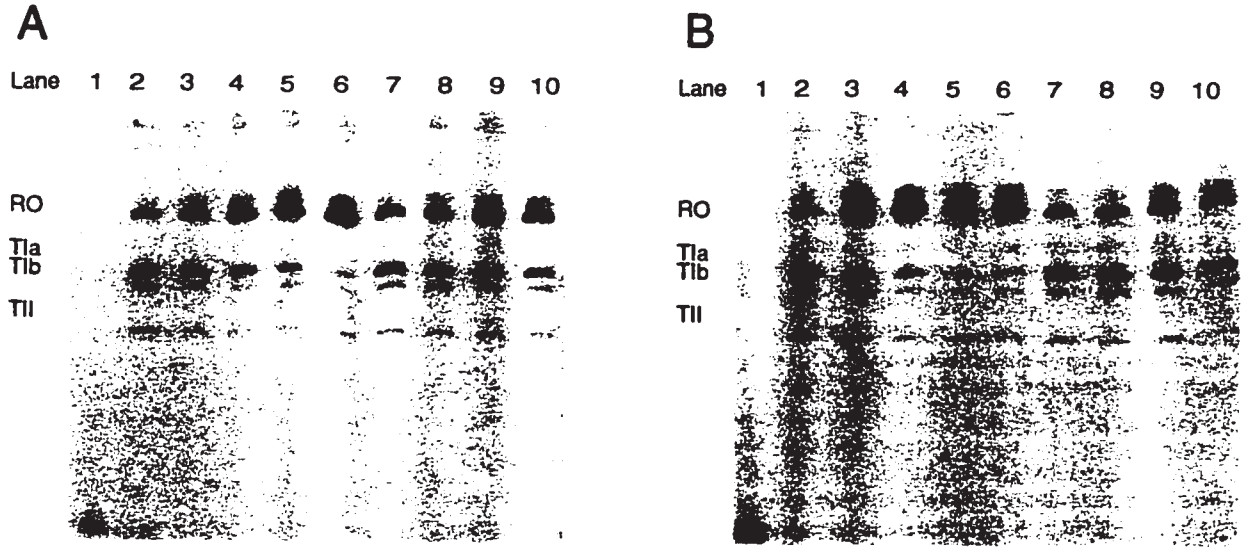


FIG. 3. **Comparison of ability of the full-length and truncated TFIIS proteins to stimulate read-through by RNA polymerase II.** Ternary complexes containing the Δ4,7 form of RNA polymerase II were formed as described under "Experimental Procedures." At 2 min 30 s, reactions containing ternary complexes stalled at intrinsic blocks to elongation were divided into two parts, and incubation was continued with the addition of either TFIIS or the storage buffer. Aliquots were stopped at various times. *A*, reactions containing full-length TFIIS. The molar ratio of full-length TFIIS to RNA polymerase II was approximately 2:1. *B*, reactions containing truncated TFIIS. The molar ratio of truncated TFIIS to RNA polymerase II was approximately 5:1. *Lane 1*, RNA from initiated complexes; *lane 2*, RNA from ternary complexes chased to the TIa, TIb, and TII sites and the runoff; *lanes 3–6*, RNA from ternary complexes incubated with TFIIS, stopped after 5, 10, 30, or 50 min of incubation, respectively; *lanes 7–10*, RNA from ternary complexes incubated in the absence of TFIIS, stopped after 5, 10, 30, or 50 min of incubation, respectively.

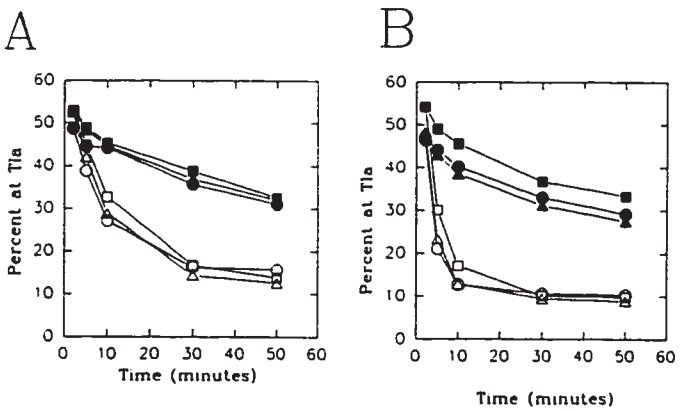


FIG. 4. **Quantitation of read-through activity by variant forms of RNA polymerase II.** Elongation assays with each form of RNA polymerase II were performed as described under "Experimental Procedures." Either full-length (*A*) or truncated (*B*) TFIIS was added, and reactions were incubated for the times indicated. RNAs were quantitated as described under "Experimental Procedures." *A*, percent of RNAs with 3′ ends at TIa with or without the full-length TFIIS protein with each form of RNA polymerase II at a molar ratio of 2:1, TFIIS:RNA polymerase II. *B*, percent of RNAs with 3′ ends at TIa with or without the truncated TFIIS protein with each form of RNA polymerase II at a molar ratio of 5:1, TFIIS:RNA polymerase II. Each point represents the average of two experiments. *Filled symbols* represent time points in the absence of TFIIS; *open symbols* represent time points in the presence of TFIIS. ▲ = IIa. ■ = Δ4,7. ● = ΔCTD.

similarity to the TFIIS family of proteins (47). In all of these cases, the 5′ end of the transcript is retained and elongated by the active complex after release of the 3′-terminal fragment (6–11). This cleavage has been observed with complexes stalled by controlling nucleotide levels (8, 46, 48, 49), by a DNA-binding protein (50), or by naturally occurring blocks to elongation (6, 7, 9, 49). In the case of complexes stalled at the TIa site in the human histone H3.3 gene, cleavage of the transcript promoted by mammalian TFIIS appears to occur before read-through of the site is detectable (7). The similarities in read-through activity between yeast TFIIS and the mammalian factor suggested that the yeast protein might also stimulate transcript cleavage.

No shortened transcripts resulting from cleavage had been seen in the experiments described above (Figs. 2 and 3), even with samples taken within 5 s after the addition of TFIIS (Fig. 2, *lane 3*). To test for cleavage in the absence of nucleotides, ternary elongation complexes containing RNA polymerase IIa stopped at the TIa, TIb, and TII sites were isolated. The incubation of either form of TFIIS with these complexes in the absence of nucleotides produces transcripts shorter than the TIa transcript. Data with truncated TFIIS are shown (Fig. 5, *lanes 2–6*). Full-length TFIIS promotes cleavage at the same rate to give the same pattern of cleaved products. Cleavage of transcripts with 3′ ends at TII can also be detected (data not shown). Complexes incubated in the absence of TFIIS produce no shortened transcripts (Fig. 5, *lane 10*). The shortened transcripts formed in the presence of either form of TFIIS can be elongated upon the addition of nucleotides (Fig. 5, *lanes 7–9*). These results are consistent with nascent transcript cleavage within active ternary complexes in the presence of TFIIS.

Ternary complexes containing each of the three variants of RNA polymerase II were compared for cleavage activity in response to the P37-containing fraction (data not shown). In this comparison, the kinetics and pattern of cleavage, the inhibition of cleavage by α-amanitin, and the elongation of truncated transcripts upon nucleotide addition were indistinguishable, regardless of the polymerase in the ternary complex.

Thus TFIIS stimulates cleavage of the nascent transcript in ternary complexes containing yeast RNA polymerase II, and the first 113 amino acids of the TFIIS protein are not required for this stimulatory activity *in vitro*. Experiments examining elongation subsequent to the cleavage reaction indicate that the nucleotides are being removed from the 3′ end of the transcript (data not shown), a result similar to that seen with the cleavage reaction in ternary complexes of vaccinia virus RNA polymerase (46) or complexes with mammalian or *Drosophila* RNA polymerase II in response to TFIIS (7–9, 11). The exact sizes of the primary cleavage products removed from the 3′ end in yeast ternary complexes are not yet known.

Within 10 s after the addition of TFIIS to stalled isolated ternary complexes, cleavage has occurred, and shortened transcripts can be detected (Fig. 5, *lane 2*). At a TFIIS to RNA

Lane	1	2	3	4	5	6	7	8	9	10	11	12
Time	0'	10"	1'	5'	10'	20'	10"	1'	20'	20'	20'	22'
TFIIS	-	+	+	+	+	+	+	+	+	-	+	+
NTPs	-	-	-	-	-	-	+	+	+	-	-	+
α-Amanitin	-	-	-	-	-	-	-	-	-	-	+	+

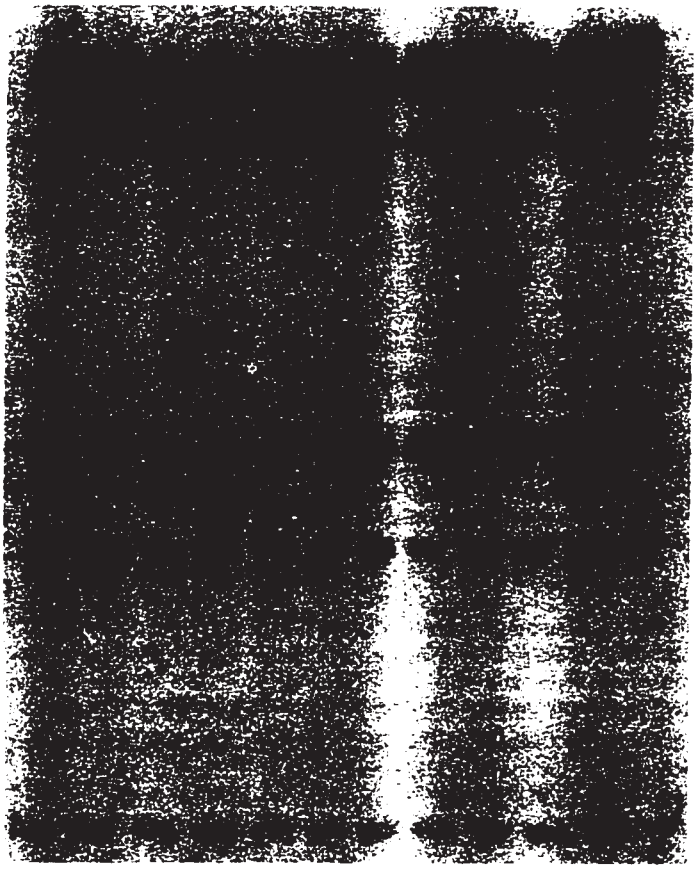


FIG. 5. **TFIIS-stimulated cleavage of nascent RNA in ternary complexes.** Isolated ternary complexes were formed with RNA polymerase IIa as described under "Experimental Procedures." *Lane 1*, RNAs from isolated complexes without further incubation; *lanes 2–6*, truncated TFIIS was added to isolated ternary complexes, and aliquots of the reaction were stopped following incubation for 10 s, 1 min, 5 min, 10 min, or 20 min after TFIIS addition; *lanes 7–9*, nucleotides (0.8 mM) were added to complexes that had been incubated with TFIIS for 5 min 30 s. Aliquots of the reaction were stopped at 10 s, 1 min, or 20 min after nucleotide addition. *Lane 10*, RNAs from isolated ternary complexes (*lane 1*) incubated 20 min with no further addition. *Lanes 11* and *12*, α-amanitin was preincubated with isolated ternary complexes for 2 min prior to the addition of TFIIS. An aliquot of the reaction was stopped 20 min after TFIIS addition (*lane 11*). Nucleotides were added to the remainder and stopped after 2 min further incubation (*lane 12*).

polymerase II ratio of 5:1, most of the transcripts with 3′ ends at TIa have been shortened within 5 min (Fig. 5, *lane 4*). As mentioned above, this ratio results in the maximum rate of read-through of this site by the polymerase. Longer incubation in the presence of TFIIS produces further shortening of transcripts within active ternary elongation complexes (Fig. 5, *lanes 5* and *6*). These shortened transcripts are still associated with ternary complexes, because within 10 s after nucleotide addition, shortened transcripts have been elongated back to the TIa site (Fig. 5, *lane 7*). However, detectable read-through of the TIa site is somewhat slower, reaching a limit within 10 min of incubation (Fig. 4*B*).

It remains unclear whether TFIIS itself generates the cleavage of the nascent transcript or if the function of TFIIS might be to stimulate an inherent cleavage activity of RNA polymerase II. To examine this, isolated ternary complexes were treated with 100 μg/ml α-amanitin prior to the addition of TFIIS. RNA polymerase II is the most sensitive of the nuclear polymerases to this toxin which interferes with transcript elongation (51). This α-amanitin treatment significantly decreases, but does not completely prevent, transcript cleavage in response to TFIIS (Fig. 5, *lane 11 compared* with *lane 6*), nor does it prevent elongation of the cleaved transcripts back to the TIa site. However, extensive shortening of cleaved transcripts appears to be prevented; the first detectable cleavage intermediate is apparently not shortened further (Fig. 5, compare *lane 11* with *lanes 2–4*). These results could indicate that a fraction of the RNA polymerase II-containing complexes are insensitive to α-amanitin. Alternatively, transcript elongation is a multistep process (52), and α-amanitin may not interfere with the cleavage reaction, although it may block elongation at some other step such as polymerase translocation (51, 53). A more trivial explanation, that the α-amanitin-insensitive cleavage is due to either RNA polymerase I or III, is very unlikely; the polymerases used in these experiments have been purified using immunoaffinity (41), taking advantage of antibody directed against the COOH-terminal repeated sequence unique to the largest subunit of RNA polymerase II (31, 35). In addition, no large subunits coincident with those expected for either RNA polymerase I or III have been visualized upon sensitive silver staining of polymerase preparations (Fig. 1*A*).

DISCUSSION

The conserved COOH-terminal domain of the largest subunit of RNA polymerase II does not have a significant effect on the ability of the polymerase to recognize intrinsic blocks to elongation or to read through them in response to the elongation factor TFIIS. Similarly, polymerase lacking subunits four and seven recognizes these blocks to elongation and reads through them in response to TFIIS. However, *in vivo*, the CTD of the largest subunit is essential (54). Furthermore, deletion of subunit four causes slow growth and temperature sensitivity of the cells (55), whereas deletion of subunit seven is lethal (36). The *in vitro* results would suggest that these *in vivo* phenotypes are unrelated to TFIIS function.

Ternary complexes containing all three variant forms of RNA

polymerase II exhibit the ability to cleave their nascent transcripts and read-through blocks to elongation in the presence of TFIIS. However, none of these experiments details how the cleavage reaction might promote read-through nor what the mechanism of the cleavage reaction might be. α-Amanitin dramatically inhibits the TFIIS-stimulated cleavage (Refs. 6, 8, 9, 11, and 53 and this report). Since this toxin binds to the polymerase, this inhibition suggests that transcript cleavage might be catalyzed by the polymerase protein with TFIIS stimulating an activity intrinsic to the polymerase. Alternatively, the TFIIS protein may be involved more directly in catalyzing the cleavage reaction, and the binding of α-amanitin may interfere with the interaction between TFIIS and the polymerase in the ternary complex.

Cleavage of nascent transcripts seems to be a feature in a general mechanism that allows RNA polymerases to resume elongation after stalling before reaching the end of a transcription unit. Such stalled complexes may terminate transcription, resulting in transcript release, if the reactions promoting read-through are not carried out quickly enough to defeat transcript release. Ternary complexes stopped at the sites described here can cleave nascent transcripts within 10 s after the addition of TFIIS, yet productive read-through of these sites requires an incubation of many minutes in the presence of both TFIIS and nucleotides. Approximately 50% of RNA polymerase II molecules stop at the TIa site. Perhaps, following cleavage in the presence of TFIIS, 50% of those ternary complexes elongating the shortened transcripts stop once again at the TIa site, even in the presence of TFIIS and nucleotides, whereas the other 50% pass through to the runoff. This cycle of stopping, cleaving, and elongating would continue until essentially all complexes pass through the block to elongation.

When mammalian RNA polymerase II is stopped at these sites *in vitro*, transcript release can be detected under physiological salt concentrations, although this release is not quantitative and is likely to be a slow process (39).[6] On the other hand, in the presence of TFIIS cleavage is very rapid (6, 7), and elongation through these sites can be nearly quantitative (7, 16, 17). Such sites may be distinct from some intrinsic *termination* sites where the polymerase would stop and transcript release would be rapid and from *pause* sites where the polymerase would stop and quantitatively resume elongation in a "finite" period of time. Alternatively, such sites which cause transcription arrest or a block to elongation might represent sequences on a continuum for which termination or read-through is determined by both the stability of the ternary complex at that specific site and the ability of that complex to respond to accessory elongation regulatory factors. If this were the case, the properties of ternary complexes would be expected to vary at different sites.

Indeed, the structures of ternary elongation complexes for both bacterial RNA polymerase and RNA polymerase II can change during elongation (53, 56–60). Each distinct structure might represent a distinct and specific target for regulation, targets that might terminate transcription or continue elongation in response to changing conditions within the cell.

There are a growing number of characterized eukaryotic genes containing "conditional" blocks to elongation or regulated transcription arrest sites within the transcription unit (12, 13). It is probable that not all of these regions are regulated by TFIIS, and likewise, not all of the regions that have been identified as blocks to elongation *in vivo* block elongation by purified RNA polymerase II *in vitro*.[7] More than one mechanism promotes terminator read-through by elongation complexes of bacterial RNA polymerase (61), and several different mechanisms might contribute to the balance between termination and read-through at sites for RNA polymerase II that fall within genes.

Transcript cleavage was first described in ternary complexes of *E. coli* RNA polymerase (62), and this cleavage reaction is promoted by at least two protein factors, greA and greB (63, 64). Stalled ternary complexes formed with mammalian (6–9), *Drosophila* (11), and yeast RNA polymerase II (this report) can cleave and then elongate nascent transcripts in response to TFIIS. Cleavage of the nascent transcript has also been observed in stalled ternary complexes formed with vaccinia RNA polymerase which contains a subunit with sequence similarity to TFIIS (46, 47). Cleavage of the nascent transcript has also been seen in isolated ternary complexes containing yeast RNA polymerase III.[8] Furthermore, purified yeast RNA polymerase I stopped by the mouse template binding termination factor TTF-1 produces a transcript shortened from the primary transcription product by about 10 nucleotides (see Fig. 3 in Ref. 65); in this case, this shortened product may be an intermediate in the eventual 3′ end processing of the ribosomal RNA transcript. However, as TTF-1 itself does not generate this cleavage product (66), it is possible that RNA polymerase I also cleaves its nascent transcript when stopped during elongation, although this cleavage may not be part of a read-through mechanism.

Certainly, the diversity of polymerases and organisms in which nascent transcript cleavage within ternary complexes is seen supports the idea that this reaction is physiologically relevant. Identifying the molecular details of this reaction will be an important next step in understanding how such a reaction can modulate the use of regulated blocks to elongation in the cell.

[6] T. Kerppola and C. Kane, unpublished results.

[7] C. Kane, unpublished results.

[8] G. Kassavetis, personal communication.

REFERENCES

1. Sawadogo, M., Sentenac, A., and Fromageot, P. (1979) *J. Biol. Chem.* **255**, 12–15
2. Sawadogo, M., Huet, J., and Fromageot, P. (1980) *Biochem. Biophys. Res. Commun.* **96**, 258–264
3. Sawadogo, M., Lescure, B., Sentenac, A., and Fromageot, P. (1981) *Biochemistry* **20**, 3542–3547
4. Archambault, J., Lacroute, F., Ruet, A., and Friesen, J. D. (1992) *Mol. Cell. Biol.* **12**, 4142–4152
5. Nakanishi, T., Nakano, A., Nomura, K., Sekimizu, K., and Natori, S. (1992) *J. Biol. Chem.* **267**, 13200–13204
6. Reines, D. (1992) *J. Biol. Chem.* **267**, 3795–3800
7. Reines, D., Ghanouni, P., Li, Q., and Mote, J., Jr. (1992) *J. Biol. Chem.* **267**, 15516–15522
8. Izban, M. G., and Luse, D. S. (1992) *Genes & Dev.* **6**, 1342–1356
9. Wang, D., and Hawley, D. K. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90**, 843–847
10. Kassavetis, G. A., and Geiduschek, E. P. (1993) *Science* **259**, 944–945
11. Guo, H., and Price, D. H. (1993) *J. Biol. Chem.* **268**, 18762–18770
12. Kerppola, T. K., and Kane, C. M. (1991) *FASEB J.* **5**, 2833–2842
13. Spencer, C. A., and Groudine, M. (1990) *Oncogene* **5**, 777–785
14. Bengal, E., Flores, O., Krauskopf, A., Reinberg, D., and Aloni, Y. (1991) *Mol. Cell. Biol.* **11**, 1195–1206
15. Reinberg, D., and Roeder, R. G. (1987) *J. Biol. Chem.* **262**, 3331–3337
16. Reines, D., Chamberlin, M. J., and Kane, C. M. (1989) *J. Biol. Chem.* **264**, 10799–10809
17. SivaRaman, L., Reines, D., and Kane, C. M. (1990) *J. Biol. Chem.* **265**, 14554–14560
18. Sluder, A. E., Greenleaf, A. L., and Price, D. H. (1989) *J. Biol. Chem.* **264**, 8963–8969
19. Wiest, D. K., Wang, D., and Hawley, D. K. (1992) *J. Biol. Chem.* **267**, 7733–7744
20. Chen, C. H., England, L., and Kane, C. M. (1992) *Gene (Amst.)* **116**, 253–258
21. Hirashima, S., Hirai, H., Nakanishi, Y., and Natori, S. (1988) *J. Biol. Chem.* **263**, 3858–3863
22. Hubert, J.-C., Guyonvarch, A., Kammerer, B., Exinger, F., Liljelund, P., and Lacroute, F. (1983) *EMBO J.* **2**, 2071–2073
23. Kanai, A., Kuzuhara, T., Sekimizu, K., and Natori, S. (1991) *J. Biochem. (Tokyo)* **109**, 674–677
24. Marshall, T. K., Guo, H., and Price, D. H. (1990) *Nucleic Acids Res.* **18**, 6293–6298
25. Rappaport, J., Reinberg, D., Zandomeni, R., and Weinmann, R. (1987) *J. Biol. Chem.* **262**, 5227–5232
26. Yoo, O., Yoon, H., Baek, K., Jeon, C., Miyamoto, K., Ueno, A., and Agarwal, K. (1991) *Nucleic Acids Res.* **19**, 1073–1079

27. Agarwal, K., Baek, K., Jeon, C., Miyamoto, K., Ueno, A., and Yoon, H. (1991) *Biochemistry* **30**, 7842–7851
28. Horikoshi, M., Sekimizu, K., and Natori, S. (1984) *J. Biol. Chem.* **259**, 608–611
29. Horikoshi, N., Sekimizu, K., and Natori, S. (1990) *J. Biol. Chem.* **265**, 11854–11857
30. Natori, S. (1982) *Mol. Cell. Biochem.* **46**, 173–187
31. Sentenac, A. (1985) *CRC Crit. Rev. Biochem.* **18**, 31–90
32. Kolodziej, P. A., Woychik, N. A., Liao, S.-M., and Young, R. A. (1990) *Mol. Cell. Biol.* **10**, 1915–1920
33. Rappaport, J., Cho, K., Saltzman, A., Prenger, J., Golomb, M., and Weinmann, R. (1988) *Mol. Cell. Biol.* **8**, 3136–3142
34. Allison, L. A., Wong, J. K.-C., Fitzpatrick, V. D., Moyle, M., and Ingles, C. J. (1988) *Mol. Cell. Biol.* **8**, 321–329
35. Corden, J. L. (1990) *Trends Biochem. Sci.* **15**, 383–387
36. McKune, K., Richards, K. L., Edwards, A. M., and Kornberg, R. D. (1993) *Yeast* **9**, 295–299
37. Woychik, N. A., and Young, R. A. (1990) *Trends Biochem. Sci.* **15**, 347–351
38. Edwards, A. M., Kane, C. M., Young, R. A., and Kornberg, R. D. (1991) *J. Biol. Chem.* **266**, 71–75
39. Reines, D., Wells, D., Chamberlin, M. J., and Kane, C. M. (1987) *J. Mol. Biol.* **196**, 299–312
40. Thompson, N. E., Aronson, D. B., and Burgess, R. R. (1990) *J. Biol. Chem.* **265**, 7069–7077
41. Edwards, A. M., Darst, S. A., Feaver, W. J., Thompson, N. E., Burgess, R. R., and Kornberg, R. D. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87**, 2122–2126
42. Kadesch, T. R., and Chamberlin, M. J. (1982) *J. Biol. Chem.* **257**, 5286–5295
43. Kerppola, T. K., and Kane, C. M. (1990) *Biochemistry* **29**, 269–278
44. Sekimizu, K., Nakanishi, Y., Mizuno, D., and Natori, S. (1979) *Biochemistry* **18**, 1582–1588
45. Horikoshi, M., Sekimizu, K., Hirashima, S., Mitsuhashi, Y., and Natori, S. (1985) *J. Biol. Chem.* **260**, 5739–5744
46. Hagler, J., and Shuman, S. (1993) *J. Biol. Chem.* **268**, 2166–2173
47. Ahn, B.-Y., Gershon, P. D., Jones, E. V., and Moss, B. (1990) *Mol. Cell. Biol.* **10**, 5433–5441
48. Izban, M. G., and Luse, D. S. (1993) *J. Biol. Chem.* **268**, 12864–12873
49. Izban, M. G., and Luse, D. S. (1993) *J. Biol. Chem.* **268**, 12874–12885
50. Reines, D., and Mote, J., Jr. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90**, 1917–1920
51. Wieland, T., and Faulstich, H. (1991) *Experientia* **47**, 1186–1193
52. Chamberlin, M. J. (1993) *Harvey Lect.*, in press
53. Gu, W., Powell, W., Mote Jr., J., and Reines, D. (1993) *J. Biol. Chem.* **268**, 25604–25616
54. Nonet, M., Sweetser, D., and Young, R. A. (1987) *Cell* **50**, 909–915
55. Woychik, N. A., and Young, R. (1989) *Mol. Cell. Biol.* **9**, 2854–2959
56. Krummel, B., and Chamberlin, M. J. (1992) *J. Mol. Biol.* **225**, 221–237
57. Krummel, B., and Chamberlin, M. J. (1992) *J. Mol. Biol.* **225**, 239–250
58. Lee, D. N., and Landick, R. (1992) *J. Mol. Biol.* **228**, 759–777
59. Linn, S. C., and Luse, D. S. (1991) *Mol. Cell. Biol.* **11**, 1508–1522
60. Rice, G. A., Chamberlin, M. J., and Kane, C. M. (1993) *Nucleic Acids Res.* **21**, 113–118
61. Turnbough, C., and Landick, R. (1992) in *Transcriptional Regulation* (McKnight, S., and Yamamoto, K. eds) pp. 407–448, Cold Spring Harbor Press, Cold Spring Harbor, NY
62. Surratt, C. K., Milan, S. C., and Chamberlin, M. J. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88**, 7983–7987
63. Borukhov, S., Sagitov, V., and Goldfarb, A. (1993) *Cell* **72**, 459–466
64. Borukhov, S., Polyakov, A., Nikiforov, V., and Goldfarb, A. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89**, 8899–8902
65. Kuhn, A., Bartsch, I., and Grummt, I. (1990) *Nature* **344**, 559–562
66. Kuhn, A., and Grummt, I. (1989) *Genes & Dev.* **3**, 224–231

Appendix II

Elongation factor TFIIS contains three structural domains: Solution structure of domain II.

Morin, P. E., Awrey, D. E., Edwards, A. M., and Arrowsmith, C. H. (1996). Proceedings of the National Acadamy of Sciences USA vol. 93, pp. 10604-10608.

Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 10604–10608, October 1996
Biochemistry

Elongation factor TFIIS contains three structural domains: Solution structure of domain II

(transcription elongation/NMR/protein structure)

PAUL E. MORIN*, DONALD E. AWREY†, ALED M. EDWARDS†‡, AND CHERYL H. ARROWSMITH*‡

*Division of Molecular and Structural Biology, Ontario Cancer Institute, and Department of Medical Biophysics, University of Toronto, 610 University Avenue, Toronto, ON Canada M5G 2M9; and †Cancer Research Group, Institute for Molecular Biology and Biotechnology, McMaster University, 1200 Main Street West, Hamilton, ON Canada L8N 3Z5

Communicated by Michael J. Chamberlin, University of California, Berkeley, CA, June 10, 1996 (received for review March 16, 1996)

ABSTRACT **Transcription elongation by RNA polymerase II is regulated by the general elongation factor TFIIS. This factor stimulates RNA polymerase II to transcribe through regions of DNA that promote the formation of stalled ternary complexes. Limited proteolytic digestion showed that yeast TFIIS is composed of three structural domains, termed I, II, and III. The two C-terminal domains (II and III) are required for transcription activity. The structure of domain III has been solved previously by using NMR spectroscopy. Here, we report the NMR-derived structure of domain II: a three-helix bundle built around a hydrophobic core composed largely of three tyrosines protruding from one face of the C-terminal helix. The arrangement of known inactivating mutations of TFIIS suggests that two surfaces of domain II are critical for transcription activity.**

Transcription elongation is not a monotonic process. During transcription elongation, RNA polymerase stutters along the DNA template, occasionally pausing for extended periods of time at sequences that block progress of the enzyme. Transcription through such blocks is facilitated by accessory elongation protein factors, including TFIIS and elongin A in eukaryotes and GreA and GreB in prokaryotes (1–4). TFIIS is conserved in all eukaryotes and can stimulate elongation by RNA polymerase II (pol II) through blockages caused by DNA-binding proteins and drugs that bind DNA (5). The yeast factor (yTFIIS) contains 309 amino acids, of which the region 143–309 ($yTFIIS_{143-309}$) is sufficient for biochemical activity *in vitro* (6, 7). This fragment of TFIIS is also functional *in vivo*; $yTFIIS_{143-309}$ can complement the 6-azauracil-sensitive phenotype displayed by yeast strains lacking the TFIIS gene (6). Although residues 1–142 are not essential for transcription activity and have no known function at this time, their role does appear conserved; the rat elongation factor, elongin A, contains a stretch of 115 amino acids near its N terminus that is 29% identical and 53% similar to the N-terminal domain of human TFIIS (8).

Here, we use limited proteolytic digestion to show that TFIIS is composed of three structural domains, termed I, II, and III, of which domains II and III constitute the minimal transcriptionally active fragment. The solution structure of human TFIIS domain III has been solved previously by Weiss and coworkers (9, 10) who refer to it as a "Zn ribbon" due to its ability to bind Zn and its β-sheet secondary structure composition. Domain III of yTFIIS is expected to be structurally very similar to human domain III because they share 61% identity in this region. We now report the solution structure of domain II ($yTFIIS_{131-240}$) which, in contrast to domain III, contains only α-helical secondary structural elements.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

METHODS

Cloning of TFIIS in Bacterial Expression Vectors. Full-length TFIIS ($TFIIS_{1-309}$) and a truncation mutant encoding residues 131–240 ($TFIIS_{131-240}$) were inserted into bacterial expression vectors. The coding sequence for each mutant was amplified from the TFIIS gene (6) by using Vent polymerase (New England Biolabs) by the polymerase chain reaction (PCR). The restriction sites *Nde*I and *Bam*HI were added to the 5′ and 3′ ends of the PCR product, respectively, and used to insert the TFIIS fragments between the *Nde*I and *Bam*HI sites of the T7 polymerase expression vector pET15b (Novagen). The proteins were expressed in *Escherichia coli* BL21(DE3) as C-terminal fusions to an N-terminal six-histidine tag and a thrombin protease site. The thrombin-digested protein retained four N-terminal residues, GSHM-, which are not found in the wild-type TFIIS sequences.

Bacterial cells expressing yTFIIS were grown in Luria broth for the proteolysis studies ($TFIIS_{1-309}$) or in M9 minimal medium containing ^{13}C-labeled glucose and/or $^{15}NH_4Cl$ for NMR studies ($TFIIS_{131-240}$). Cells were grown at 37°C to an optical density of 0.8 at 600 nm, and TFIIS expression was induced with 1.0 mM isopropyl β-D-thiogalactopyranoside. Three hours after induction, the cells were harvested by centrifugation, resuspended in 50 mM Tris·HCl, pH 7.5/10% sucrose/1.0 mM benzamidine/1.0 mM phenylmethylsulfonyl fluoride/1.0 mM EDTA (3 ml per gram of cells) and frozen at −70°C. TFIIS was purified from the frozen cells as described by Christie *et al.* (11). $TFIIS_{1-309}$ was stored frozen at 3.5 mg/ml in a solution containing 20 mM Hepes at pH 7.5, 50 mM NaCl, 10 mM dithiothreitol (DTT), 10% (vol/vol) glycerol, and 10 μM $ZnSO_4$. For NMR studies, isotopically labeled $TFIIS_{131-240}$ was dialyzed against 5 mM potassium phosphate/10 mM DTT/10 μM $ZnSO_4$, pH 6.5. The protein was then concentrated by ultra-filtration to 2 mM protein. During concentration, protonated DTT was replaced with deuterated DTT and ^2H_2O (D_2O) was added to provide a deuterium lock signal. The final solution contained 5 mM potassium phosphate, 5 mM deuterated DTT, pH 6.8–6.9, in 7% D_2O/93% H_2O (vol/vol). For the observation of amide exchange rates, and for certain nuclear Overhauser effect (NOE) spectroscopy (NOESY) spectra (see below), protein samples were lyophilized from the solution conditions above and resuspended in 99.996% D_2O.

Partial Proteolysis Studies. For analytical purposes, 5 μg of purified yTFIIS-poly-histidine fusion protein was incubated with 10 ng of endoproteinase Glu-C from *Staphylococcus* V8 (V8 protease; Boehringer Mannheim), chymotrypsin (Sigma), Pronase (Boehringer Mannheim), or trypsin (Sigma) for up to

Abbreviations: pol II, RNA polymerase II; NOE, nuclear Overhauser effect; NOESY, NOE spectroscopy.

‡To whom reprint requests should be addressed.

60 min in protease buffer (20 mM Hepes, pH 7.5/150 mM NaCl/5 mM DTT/10 μM $ZnSO_4$/10% glycerol). The proteolytic products were resolved by SDS/polyacrylamide gel electrophoresis and visualized by staining with Coomassie blue. For preparative purposes, 5 mg of purified yTFIIS fusion protein was incubated with 10 μg of V8 protease for 30 min in protease buffer. The proteolytic products were resolved by Ni-affinity chromatography and purified by reverse-phase chromatography on a C_4 column developed with an acetonitrile gradient. The molecular weights of the purified products were determined by laser desorption mass spectroscopy.

NMR Spectroscopy and Spectral Assignments. All NMR spectra were collected at 23°C on either a Varian Unity Plus 500-MHz or Unity 600-MHz spectrometer equipped with pulse-field gradient units and actively shielded *z*-gradient triple-resonance probes. NMR experiments involving correlations of amide protons were conducted with gradient-enhanced versions (12) of the originally published pulse sequences. NMR data used in these studies were processed by using nmrPipe software (13). Spectral analysis was assisted with the programs PIPP and CAPP (14).

Sequential assignment of ^1H, $^{13}C^\alpha$, ^{15}N, and $^{13}C^\beta$ resonances for 99 of the 109 nonproline residues making up $yTFIIS_{131-240}$ was achieved through correlations from CBCA(CO)NNH (15), CBCANH (16), ^{15}N NOESY-HMQC (17, 18), and ^1H-^{15}N heteronuclear sequential quantum correlation (HSQC) (19, 20) spectra. Eight of the remaining 10 nonproline residues from the N terminus (N-terminal GSHM and residues P131–S134 of yTFIIS) display no observable NH cross peaks in HSQC spectra. Residues K209 and N210, located in the C-terminal loop of the structured region, behave in a similar manner in the HSQC experiment. This behavior is likely due to line broadening arising from conformational heterogeneity in these portions of the polypeptide chain.

Side-chain ^1H and ^{13}C resonances of aliphatic residues were assigned from the following spectra: HCCH-TOCSY (19, 21) recorded with mixing times (τ_m) of 7 ms and 14 ms; a two-dimensional (2D) constant-time ^{13}C HSQC (22); CCC-TOCSY; and ^{15}N TOCSY-HMQC (23) with a 41-ms τ_m. Side-chain ^1H resonances of aromatic residues were assigned from 100-ms τ_m homonuclear 2D NOESY (24) and TOCSY (25) spectra measured in D_2O. For the 99 residues whose backbone amide protons were assigned, correlation of the above spectra allowed assignment of 93% of side-chain proton and 94.5% of side-chain ^{13}C resonances (excluding 11 lysines and 3 arginines with highly degenerate methylene resonances). Lysines and arginines with highly degenerate methylene resonances are charged surface residues whose side chains are likely to have a greater mobility than those of the protein core.

^1H-^1H NOEs were identified from the following spectra: ^{13}C-edited NOESY in D_2O (26); ^{13}C, ^{15}N-edited NOESY in H_2O (27); and 2D homonuclear NOESY in D_2O, which was employed primarily to identify NOEs to/from aromatic residues. NOE mixing times were 150 ms in all cases. Finally, a 2D $^1H/^{15}N$ HMQC-J spectrum (28) was acquired with a final digital resolution of 0.035 Hz per point in F1, allowing the determination of $^3J_{NH\text{-}H\alpha}$ coupling constants.

Structural Restraints and Calculations. The NMR-derived structural information used to calculate the structure of $yTFIIS_{131-240}$ included four types of restraints: (*i*) NOE-derived distance restraints were used with limits of 0 to 2.7 Å, 0 to 3.3 Å, and 0 to 5.0 Å, corresponding to strong, medium, and weak intensity NOEs, respectively (29). Upper limits of distance restraints to amide protons found in helical segments were relaxed to 2.9 Å and 3.5 Å for strong and medium intensity NOEs, respectively. Pseudo-atom corrections were made to distance upper bounds involving methyl, methylene, and aromatic ring protons of Tyr and Phe residues (30, 31). An additional 0.5 Å was added to the upper bounds of NOEs from methyl groups (28, 31). (*ii*) ϕ angles of residues with $^3J_{NH\text{-}H\alpha}$ coupling constants between 8.0 and 9.0 Hz were constrained to $-120° \pm 40°$; $^3J_{NH\text{-}H\alpha}$ values greater than 9.0 Hz were constrained to $-120° \pm 30°$; values ≤ 5.5 Hz were constrained to $-60° \pm 30°$ (32, 33, 34). (*iii*) Additional backbone ϕ and ψ dihedral angle restraints were implemented on the basis of backbone chemical shift indices (35). Specifically, residues with chemical shift index consensus values of +2 for $^{13}C^\alpha$ and $^1H^\alpha$ resonances were restrained to $-60° \pm 30°$ (ϕ) and $-40° \pm 40°$ (ψ); those with consensus values of −2 were restrained to $-120° \pm 30°$ (ϕ) and $130° \pm 40°$ (ψ). (*iv*) Main-chain hydrogen bond restraints (H—O, 2.5 Å; N—O, 3.5 Å) within α-helices were implemented for residue pairs that were clearly α-helical as determined by multiple criteria: significantly lower rates of amide hydrogen exchange, NOE patterns characteristic of α-helices, small $^3J_{NH\text{-}H\alpha}$ values, and α-helical secondary chemical shift indices for both donor and acceptor.

Three-dimensional structures were calculated by using the dynamic simulated annealing protocol (36) implemented within X-PLOR (37, 38). Initial calculations employed approximately 10 unambiguous restraints per residue in the structured region of the molecule. In a fashion similar to that described by Bagby *et al.* (39) and Forman-Kay *et al.* (40), structures were calculated iteratively to resolve ambiguities in the identity of additional NOEs through inspection of resultant structures. The final calculations relied upon 665 intraresidue, 358 sequential, 212 medium-range ($1 < i - j \leq 4$), and 104 long-range ($i - j \geq 5$) interproton restraints (Fig. 1*C*), 44 restraints from 22 main-chain hydrogen bonds, and 81 backbone ϕ and 74 backbone ψ torsion angle restraints. From a total of 50 simulated annealing structures calculated, the 20 lowest-energy structures were selected. These structures had no NOE violations > 0.3 Å and no dihedral violations > 5°. Structural statistics from this ensemble are listed in Table 1. The atomic coordinates will be submitted to the Brookhaven Protein Data Bank.

RESULTS AND DISCUSSION

Our goal is to derive a three-dimensional structure for the elongation factor TFIIS. Attempts at crystallization of the full-length protein ($TFIIS_{1-309}$) or the smallest active protein ($TFIIS_{143-309}$) have proven unsuccessful. Since the entire protein is too large for NMR structure determination, we set out to identify structural domains that might be amenable to NMR

Table 1. Structural statistics for the 20 simulated annealing structures

rms deviations from experimental distance restraints, Å	
All (1381)	0.041 ± 0.003
Sequential ($\|i - j\| = 1$) (358)	0.041 ± 0.004
Medium-range ($\|i - j\| \leq 5$) (212)	0.063 ± 0.004
Long-range ($\|i - j\| > 5$) (104)	0.046 ± 0.010
Intraresidue (665)	0.023 ± 0.003
Hydrogen bond (42)	0.042 ± 0.006
rms deviation from experimental dihedral restraints, °	0.39 ± 0.14
F_{noe}, *kcal·mol^{-1}	62.2 ± 6.80
F_{cdih} *kcal·mol^{-1}	4.93 ± 0.55
F_{repel} †kcal·mol^{-1}	61.39 ± 8.80
Deviations from covalent geometry	
Bond lengths, Å	0.0095 ± 0.0011
Angles, °	2.145 ± 0.0580
Impropers, °	0.755 ± 0.0520

The number of each type of restraint used in the structure calculations is given in parentheses.

*F_{noe} and F_{cdih} were calculated using force constants of 50 kcal·mol^{-1}·mol^{-1}·Å^{-4} and 200 kcal·mol^{-1}·rad^{-2} (1 kcal = 4.18 kJ).

†F_{repel} was calculated using a final value of 4.0 cal·mol^{-1}·Å^{-4} with the van der Waals hard-sphere radii set to 0.75 times those in the parameter set PARALLHSA supplied with X-PLOR 3.1 (38).

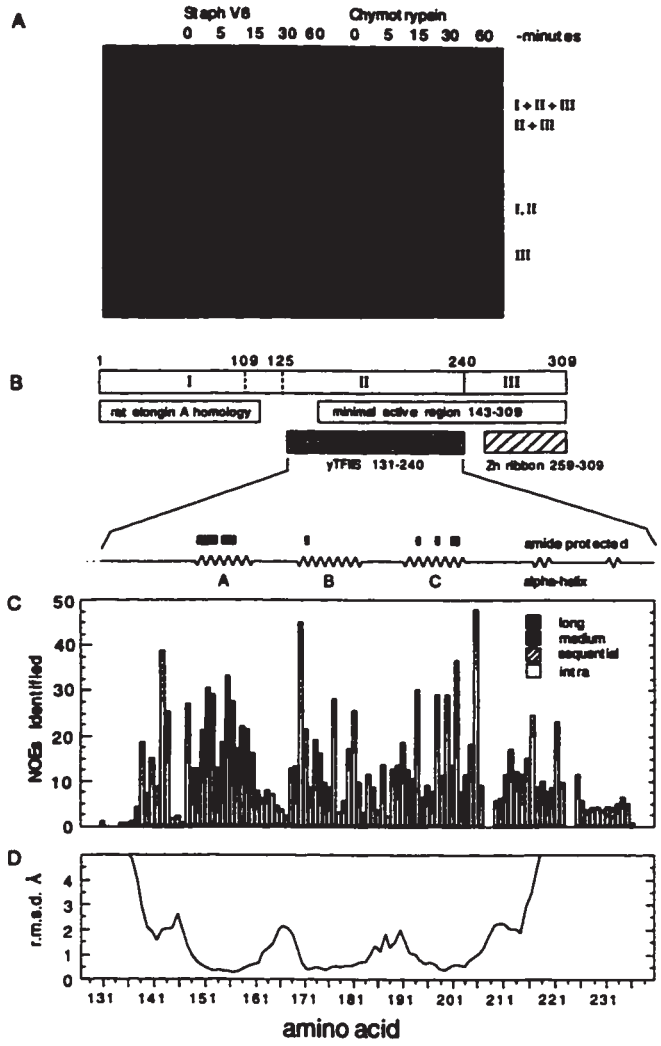


FIG. 1. Structural domains of yTFIIS. (*A*) SDS/PAGE gel showing the cleavage patterns for $yTFIIS_{1-309}$ treated with V8 protease and chymotrypsin. (*B*) Schematic of the TFIIS primary sequence with proteolytically defined domains I–III indicated. Also indicated are the region with homology to rat elongin A (8), the minimal active region, the construct used in the present study ($yTFIIS_{131-240}$), and the Zn ribbon reported by Qian *et al.* (9, 10). The α-helical regions in $yTFIIS_{131-240}$ are indicated by jagged lines and the residues which have solvent-protected amide hydrogens are indicated with black bars. (*C*) Histogram of the number and classification of NOEs for each residue. (*D*) Average rms deviation from the energy-minimized mean structure per residue.

structure determination. The boundaries of stable structural domains were determined by using limited proteolytic digestion. $yTFIIS_{1-309}$ was treated with a selection of different proteases (V8 protease, chymotrypsin, trypsin, and Pronase), and the digestion products were monitored by gel electrophoresis. Representative V8 protease and chymotrypsin cleavage patterns are shown in Fig. 1*A*. For the V8 protease digestion, complete cleavage occurred at residues 105 and 246 and partial cleavage was also observed at residue 124, suggesting that TFIIS contains three structural domains (Fig. 1*B*).

Domain I of yTFIIS, which extends from the N terminus to the region 105–124, corresponds well with the region recently identified in human TFIIS as being 29% identical and 53% similar to the N terminus of rat elongin A (8). Therefore, domain I likely has a conserved protease-resistant three-dimensional structure. Domain II extends from the 105–124 region to residue 246. The sequence of this domain is conserved among known TFIIS genes and, together with domain III, is essential for transcription activity. Domain III comprises the C-terminal 50 or so amino acids and is highly conserved in all TFIIS genes. The structure of domain III from human TFIIS has been determined by NMR spectroscopy and contains a Zn-ribbon domain (9, 10). This domain, which is composed entirely of β-sheet secondary structure, does not bind to nucleic acids as orignally reported (41). Amino acid sequence alignments have predicted Zn-ribbons to exist in a variety of transcription factors, including TFIIB (42, 43) and RPB9, a subunit of pol II (44). However, Zn-ribbons have no function known at this time.

The purpose of the TFIIS structure determination is to provide a structural framework for our biochemical and genetic studies of its transcriptional elongation activity. Therefore, we initially focused on identifying a fragment of TFIIS that harbored elongation activity and displayed the requisite stability and solubility characteristics for NMR. Preliminary NMR spectra were collected on the smallest active fragment of TFIIS (143–309). Although this fragment was active, the ^{15}N-HSQC NMR spectra indicated that $TFIIS_{143-309}$ had regions of conformational heterogeneity and/or disorder and was not suitable for structural analysis. We therefore turned to the cloning and expression of various ^{15}N-labeled fragments of the protein to identify those that would yield suitable NMR spectra. For these experiments, selection of the N and C termini of the fragments was guided by our knowledge of the structural domains determined by partial proteolysis. We identified two derivatives, $yTFIIS_{131-309}$ and $yTFIIS_{131-240}$, that displayed NMR spectra indicative of stable domains in a single conformational state. $yTFIIS_{131-240}$ is contained within our proteolytically defined domain II; $yTFIIS_{131-309}$ contains domains II and III, including the complete, naturally occurring C terminus. Comparison of the ^{15}N HSQC and the methyl region of ^{13}C HSQC spectra derived from these two protein fragments showed that the chemical shifts of $yTFIIS_{131-240}$ represented a subset of those found in the $yTFIIS_{131-309}$ spectra; approximately 72% of the amide ^{15}N–^{1}H peaks and approximately 76% of the methyl ^{13}C–^{1}H peaks of residues in the structured regions of domain II have identical chemical shifts in $yTFIIS_{131-240}$ and in $yTFIIS_{131-309}$. Thus, to a first approximation, it appears that no appreciable disruption of the structure of domain II is caused by the absence of the C-terminal domain III.

Fig. 2 shows the 20 lowest-energy structures that satisfy the NMR-derived structural restraints. Residues 139–215 form a well-defined three-helix bundle composed of helix A (R149–L160), helix B (I171–M182), and helix C (A193–S204) with a short, N-terminal, extended strand, D137–T142. The hydrophobic core of this bundle is built around three tyrosine residues (Y195, Y199, and Y203), all protruding from the same side of helix C. The short N-terminal strand folds V139 back onto a hydrophobic patch on helix A formed largely by Y157. The absence of this interaction in $yTFIIS_{143-240}$ may explain why the 143–240 construct was less stable in solution. Although the C-terminal 25 amino acids of $yTFIIS_{131-240}$ do appear to contain some small level of helical secondary structure, the region displays fewer NOEs, somewhat narrower line widths, and faster amide exchange rates than does the three-helix bundle region, residues 139–215. It is possible that this C-terminal region forms a flexible linker between the three-helix bundle and the Zn-ribbon of domain III. The N-terminal 8 residues (GSHM from the cloning vector and P131–S134 of yTFIIS) contain no regular secondary or tertiary structure, as evidenced by the low abundance of NOEs in this region (Fig. 1*C*). Simultaneous alignment of helices A, B, and C for the ensemble of structures to its mean yields an rms deviation of 0.52 Å for backbone atoms (N, C′, and C^{α}) and 0.91 Å for all heavy atoms in the helices. Deviations are 0.98 Å for backbone and 1.43 Å for all heavy atoms in the structured region (V139–K215 depicted in Figs. 2 and 3).

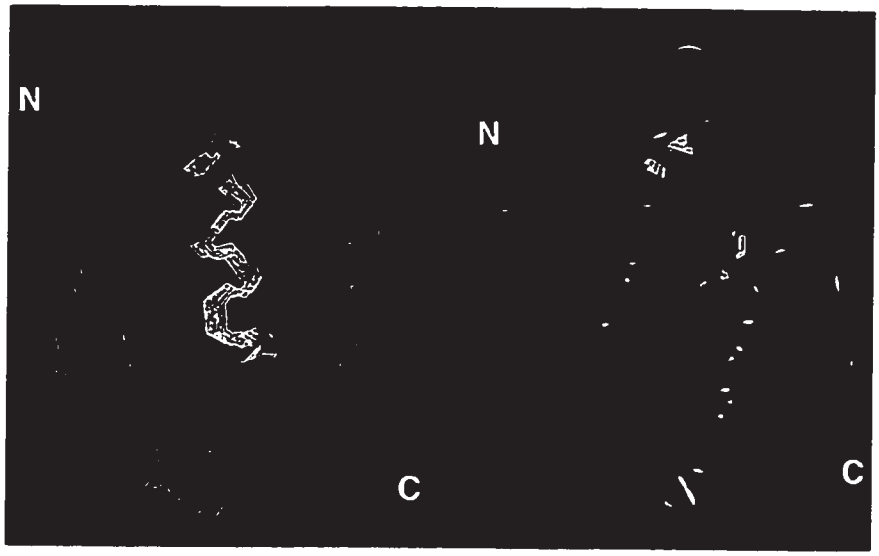


FIG. 2. (*Left*) A backbone superposition (helices only) of the 20 lowest-energy structures for residues V139–K215. Helices A, B, and C are shown in fuchsia, yellow, and red, respectively. (*Right*) A backbone ribbon representation of a single structure with the same view as in the left image. Selected hydrophobic residues from 10 structures are shown in the same colors as their corresponding helix in the left image.

Domain II is essential for pol II to read through transcription pause sites and for transcript cleavage. Cipres-Palacin and Kane (45, 46) have used alanine scanning mutagenesis to identify clusters of charged residues on the surface of human TFIIS that affect these elongation activities. Mutations in two of these clusters, termed cluster 1 and cluster 7b by those authors, are in domain II and abolish both read-through and transcript cleavage capabilities (46). Mutations in a third cluster within domain II, cluster 7, result in the loss of read-through activity but retain 50% of wild-type cleavage activity. The charged residues in these clusters, which are conserved between the yeast and human proteins, are K147/R149, K196, and R198/R200 for clusters 1, 7, and 7b, respectively. The surfaces implicated by these mutations map to two separate faces of the domain II helix bundle and are displayed in Fig. 3. Although the phenotypes of these human TFIIS mutations have yet to be verified in the yeast system, our structure suggests that two separate surfaces are critical for read-through and mRNA cleavage functions of TFIIS. Interestingly, clusters 7 and 7b, which display different phenotypes in human TFIIS, are found adjacent to one another in three-dimensional space.

TFIIS is functionally similar to the bacterial transcription elongation factors GreA and GreB. Each of these factors binds to its respective RNA polymerase, promotes read-through of transcription pause sites, and stimulates transcript cleavage. The functional similarity does not extend to sequence conservation; although the primary sequences of GreA and GreB share 50% sequence similarity, they are completely unrelated to TFIIS. A comparison of the structures of TFIIS domains II and III and GreA revealed a holistic similarity; each protein is composed of two domains of approximately equal size, an N-terminal helical domain and a C-terminal β-sheet domain. In addition, the base of the helical domain in both GreA and TFIIS contains basic residues that are critical for transcription activity. However, in detail, the structures of TFIIS and GreA are significantly different; the helical domain of GreA is an extended, antiparallel coiled coil, while that of TFIIS is a more

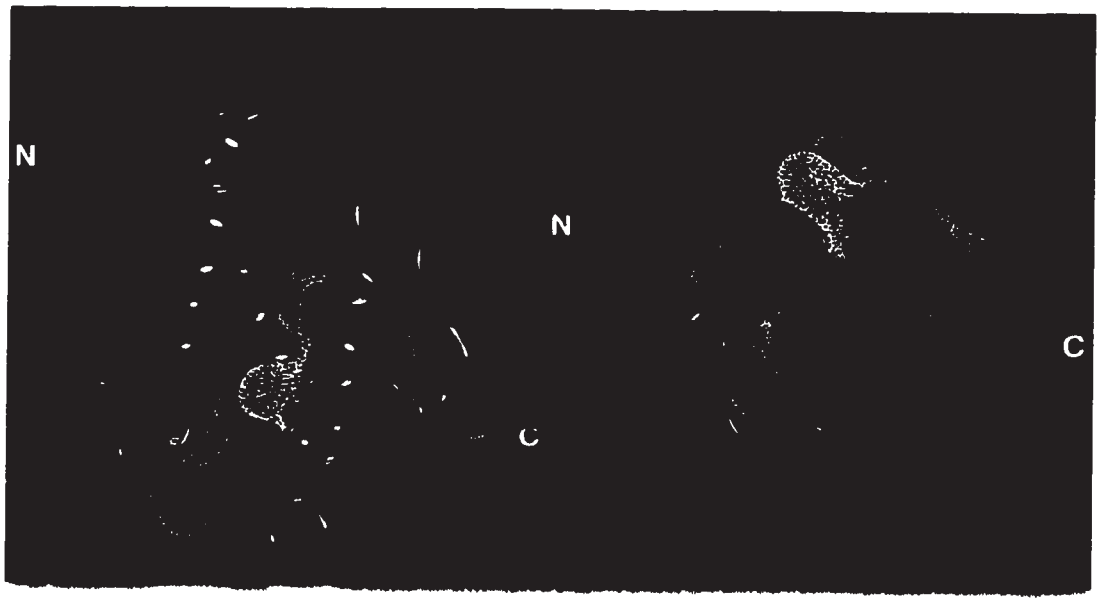


FIG. 3. Ribbon representation of the energy-minimized average structure for residues 139–215 in purple. Connally surfaces are shown for the five mutation-sensitive, charged residues which are conserved between human and yeast TFIIS. When mutated to alanine in human TFIIS, residues K147, R149, R198, and R200 (red/orange) abolish both read-through and transcript cleavage activities of pol II. Mutation of K196 (green/yellow) to alanine in human TFIIS results in the elimination of read-though activity, but it only partially reduces cleavage activity (45, 46).

compact helix bundle, and the β domain of GreA does not bind Zn as does domain III of TFIIS.

We thank Dr. C. Kane for sharing results prior to publication and Dr. Lewis Kay for supplying NMR pulse sequences. This work was supported by grants from the Medical Research Council of Canada and the Human Frontier Science Program to C.H.A. and A.M.E. and by a postdoctoral fellowship from the Human Frontier Science Program (P.M.). C.H.A. and A.M.E. are Medical Research Council of Canada Scholars.

1. Reines, D., Chamberlin, M. J. & Kane, C. M. (1989) *J. Biol. Chem.* **264,** 10799–10809.
2. Bradsher, J. N., Jackson, K. W., Conaway, R. C. & Conaway, J. W. (1993) *J. Biol. Chem.* **268,** 25587–25593.
3. Borukhov, S., Polyakov, A., Nikifovov, V. & Goldfarb, A. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 8899–8902.
4. Borukhov, S., Sagitov, V. & Goldfarb, A. (1993) *Cell* **72,** 459–466.
5. Reines D. & Mote, J., Jr. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 1917–1921.
6. Nakanishi, T., Nakano, A., Nomura, K., Sekimizu, K. & Natori, S. (1992) *J. Biol. Chem.* **267,** 13200–13204.
7. Nakanishi, T., Shimoaraiso, M., Kubo, T. & Natori, S. (1995) *J. Biol. Chem.* **270,** 8991–8995.
8. Aso, T., Lane, W. S., Conaway, J. W. & Conaway, R. C. (1995) *Science* **269,** 1439–1443.
9. Qian, X., Gozani, S. N., Yoon, H. S., Jeon, C. J., Agarwal, K. & Weiss, M. A. (1993) *Biochemistry* **32,** 9944–9959.
10. Qian, X., Jeon, C. J., Yoon, H. S., Agarwal, K. & Weiss, M. A. (1993) *Nature (London)* **365,** 277–279.
11. Christie, K. R., Awrey, D. E., Edwards, A. M. & Kane, C. M. (1994) *J. Biol. Chem.* **269,** 936–943.
12. Kay, L. E., Keifer, P. & Saarinen, T. (1992) *J. Am. Chem. Soc.* **114,** 10663–10665.
13. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J. & Bax, A. (1995) *J. Biol. NMR* **6,** 277–293.
14. Garrett, D. S., Powers, R., Gronenborn, A. M. & Clore, G. M. (1991) *J. Magn. Reson.* **95,** 214–220.
15. Grzesiek S. & Bax, A. (1992) *J. Magn. Reson.* **96,** 432–440.
16. Kay, L. E., Xu, G. Y. & Yamazaki T. (1994) *J. Magn. Reson. Ser. A* **109,** 129.
17. Marion, D., Kay, L. E., Sparks, S. W., Torchia, D. A. & Bax, A. (1989) *J. Am. Chem. Soc.* **111,** 1515–1517.
18. Zuiderweg, E. R. P., Boelens, R. & Kaptein, R. (1985) *Biopolymers* **24,** 601–610.
19. Bax, A., Clore, G. M. & Gronenborn, A. M. (1990) *J. Magn. Reson.* **88,** 425–431.
20. Bodenhausen, G. & Reuben, D. J. (1980) *Chem. Phys. Lett.* **69,** 185–189.
21. Kay, L. E., Xu, G. Y., Singer, A. U., Muhandiram, D. R. & Forman-Kay, J. (1993) *J. Magn. Reson. Ser. B* **101,** 133–136.
22. Vuister, G. W. & Bax, A. (1992) *J. Magn. Reson.* **98,** 428–435.
23. Marion, D., Driscoll, P. C., Kay, L. E., Wingfield, P. T., Bax, A., Gronenborn, A. M. & Clore, G. M. (1989) *Biochemistry* **28,** 6150–6156.
24. Jeener, J., Meier, B. H., Bachmann, P. & Ernst, R. R. (1979) *J. Chem. Phys.* **71,** 4546–4553.
25. Braunschweiler, L. & Ernst, R. R. (1983) *J. Magn. Reson.* **53,** 521–528.
26. Ikura, M., Kay, L. E., Tschudin, R. & Bax, A. (1990) *J. Magn. Reson.* **86,** 204–209.
27. Pascal, S. Muhandiram, T., Yamazaki, T., Forman-Kay, J. D. & Kay, L. E. (1994) *J. Magn. Reson. Ser. B* **101,** 197–201.
28. Kay, L. E. & Bax, A. (1990) *J. Magn. Reson.* **86,** 110.
29. Clore, G. M., Nilges, M., Sukumaran, D. K., Brünger, A. T., Karplus, M. & Gronenborn, A. M. (1986) *EMBO J.* **5,** 2729–2735.
30. Wüthrich, K., Billeter, M. & Braun, W. (1983) *J. Mol. Biol.* **169,** 949–961.
31. Wagner, G., Braun, W., Havel, T. F., Schaumann, T., Go, N. & Wüthrich, K. (1987) *J. Mol. Biol.* **196,** 611–639.
32. Pardi, A., Billeter, M. & Wüthrich, K. (1984) *J. Mol. Biol.* **180,** 741–751.
33. Redfield, C. & Dobson, C. M. (1990) *Biochemistry* **29,** 7201–7214.
34. Smith, L. J., Sutcliffe, M. J., Redfield, C. & Dobson, C. M. (1991) *Biochemistry* **30,** 986–996.
35. Wishart, D. S. & Sykes B. D. (1993) *J. Biol. NMR* **4,** 171–175.
36. Nilges, M., Gronenborn, A. M., Brünger, A. T. & Clore, G. M. (1988) *Protein Eng.* **2,** 27.
37. Brünger, A. T. (1992) X-PLOR Manual (Yale Univ., New Haven, CT), Version 3.0.
38. Brünger, A. T. (1993) X-PLOR Manual (Yale Univ., New Haven, CT), Version 3.1.
39. Bagby, S., Harvey, T. S., Eagle, S., Inouye, S. & Ikura, M. (1994) *Structure* **2,** 107–121.
40. Forman-Kay, J. D., Clore, G. M., Wingfield, P. T. & Gronenborn, A. M. (1991) *Biochemistry* **28,** 7241–7257.
41. Qian, X., Jeon, C. J., Yoon, H. S., Agarwal, K. & Weiss, M. A. (1995) *Nature (London)* **376,** 279.
42. Buratowski, S. & Zhou, H. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 5633–5637.
43. Zhu, W., Zeng, Q., Colangelo, C. M., Lewis, M., Summers, M. F. & Scott, R. A. (1996) *Nat. Struct. Biol.* **3,** 122–124.
44. Woychik, N. A., Lane, W. S. & Young, R. A. (1991) *J. Biol. Chem.* **266,** 19053–19055.
45. Cipres-Palacin, G. & Kane, C. M. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 8087–8091.
46. Cipres-Palacin, G. & Kane, C. M. (1995) *Biochemistry* **34,** 15375–15380.

Appendix III

In vitro characterization of mutant yeast RNA polymerase II with reduced binding for elongation factor TFIIS.

Wu, J., Awrey, D. E., Edwards, A. M., Archambault, J., and Friesen, J. D.(1996). Proceedings of the National Academy of Sciences USA vol. 93, pp. 11552-11557. Reprinted by permission of the Proceedings of the National Academy of Sciences USA.

Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 11552–11557, October 1996
Biochemistry

In vitro characterization of mutant yeast RNA polymerase II with reduced binding for elongation factor TFIIS

JIUSHENG WU*, DONALD E. AWREY†, ALED M. EDWARDS†, JACQUES ARCHAMBAULT*‡, AND JAMES D. FRIESEN*§

*Banting and Best Department of Medical Research, and the Department of Molecular and Medical Genetics, University of Toronto, Toronto, ON Canada M5G 1L6; and †Cancer Research Group, Institute for Molecular Biology and Biotechnology, McMaster University, Hamilton, ON Canada L8N 3Z5

Communicated by Carol A. Gross, University of California, San Francisco, CA, July 11, 1996 (received for review February 20, 1996)

ABSTRACT **We have reported previously the isolation and genetic characterization of mutations in the gene encoding the largest subunit of yeast RNA polymerase II (RNAPII), which lead to 6-azauracil (6AU)-sensitive growth. It was suggested that these mutations affect the functional interaction between RNAPII and transcription-elongation factor TFIIS because the 6AU-sensitive phenotype of the mutant strains was similar to that of a strain defective in the production of TFIIS and can be suppressed by increasing the dosage of the yeast TFIIS-encoding gene, *PPR2*. RNAPIIs were purified and characterized from two independent 6AU-sensitive yeast mutants and from wild-type (wt) cells. *In vitro*, in the absence of TFIIS, the purified wt polymerase and the two mutant polymerases showed similar specific activity in polymerization, readthrough at intrinsic transcriptional arrest sites and nascent RNA cleavage. In contrast to the wt polymerase, both mutant polymerases were not stimulated by the addition of a 3-fold molar excess of TFIIS in assays of promoter-independent transcription, readthrough or cleavage. However, stimulation of the ability of the mutant RNAPIIs to cleave nascent RNA and to read through intrinsic arrest sites was observed at TFIIS:RNAPII molar ratios greater than 600:1. Consistent with these findings, the binding affinity of the mutant polymerases for TFIIS was found to be reduced by more than 50-fold compared with that of the wt enzyme. These studies demonstrate that TFIIS has an important role in the regulation of transcription by yeast RNAPII and identify a possible binding site for TFIIS on RNAPII.**

RNA polymerase II (RNAPII) can be regulated at the transcription-elongation step as well as at the initiation step (reviewed in refs. 1 and 2). Four general transcription-elongation factors, TFIIS, TFIIF, elongin (SIII), and ELL, have been shown *in vitro* to increase the overall elongation rate of transcribing RNAPII. TFIIS (SII, RAP38) stimulates incorporation of nucleotides into RNA in promoter-independent transcription (3) and facilitates the passage of RNAPII through several kinds of transcriptional blocks such as intrinsic arrest sites (4) and certain DNA-bound proteins (5). TFIIS exerts its function by binding directly to RNAPII (6, 7) and stimulating an intrinsic RNAPII endoribonuclease activity (8, 9). The other three elongation factors, TFIIF (RAP30/74), elongin, and ELL, function differently from TFIIS in that they cannot relieve RNAPII molecules that are already arrested. Instead, these factors increase the overall elongation rate of transcribing RNAPII (7, 10–14), probably by preventing arrest of RNAPII at transcriptional blocks.

Little is known about regions of RNAPII that are involved in transcript elongation and in interaction with general elongation factors. In the yeast *Saccharomyces cerevisiae*, deletion of the TFIIS-encoding gene, *PPR2*, renders cells viable but sensitive to the uracil analog, 6-azauracil (6AU) (15). Addition of 6AU to a yeast culture results in a significant depletion of intracellular GTP and UTP pools (15). Since reducing NTP concentrations *in vitro* decreases the elongation rate and increases pause and arrest of RNA polymerase (16–19), it has been hypothesized (15, 20) that a similar effect may be obtained *in vivo* by growth of yeast cells in the presence of 6AU. Thus, the 6AU-sensitivity of the PPR2 deletion mutant (ppr2Δ) cells may be related directly to the well-characterized TFIIS-mediated readthrough of intrinsic blocks to elongation by RNAPII *in vitro*. This reasoning led to the speculation that mutations resulting in 6AU-sensitive growth might also exist in regions of the subunits of RNAPII that are involved in the TFIIS-mediated function.

A genetic search of the gene, *RPO21*, which encodes the largest subunit of RNAPII, revealed seven mutations leading to 6AU-sensitive growth (20). All of them affected amino-acid residues that were located between the highly-conserved regions G and H of Rpo21p (20–22) and all were suppressed by elevated dosage of wild-type (wt) *PPR2* (20), implicating a perturbation of TFIIS-dependent elongation. Two models were proposed to explain this result (20). First, the mutant polymerases themselves might have aberrant elongation properties that result in an increase in the number of mutant polymerase molecules stopped on the DNA template and therefore require a greater number of TFIIS molecules. Second, the alterations in the mutant polymerases may weaken the interaction between RNAPII and TFIIS, demanding more TFIIS to drive the interaction toward the formation of a normal number of active RNAPII-TFIIS complexes.

To distinguish between these models, we have purified two of the seven mutant polymerases and have performed a series of *in vitro* assays. We report here that the alterations in the mutant RNAPIIs weaken the interaction between the enzymes and TFIIS.

MATERIALS AND METHODS

Purification of Proteins. Recombinant yeast TFIIS protein, tagged at the N terminus with six histidine residues (to facilitate purification) and a five-amino acid recognition element for bovine-heart-muscle kinase (to facilitate labeling; ref. 23), was purified to virtual homogeneity from *Escherichia coli* cells essentially as described (24). No significant differences were detected in the stimulation of readthrough and cleavage activities of RNAPII between this preparation of recombinant TFIIS protein and a variety of yeast TFIIS proteins described previously (24). Wt and mutant RNAPIIs were purified to near homogeneity from yeast strains expressing wt RNAPII, *rpo21-18* (I1237TRARV), or *rpo21-24* (E1230K) mutant RNAPII (20) as described (25). Protein concentrations were determined by the method of Bradford (26).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: RNAPII, eukaryotic RNA polymerase II; 6AU, 6-azauracil; ts, temperature sensitive; wt, wild type.

‡Present address: Bio-Mega/Boehringer Ingelheim Research Inc., 2100 Cunard Street, Laval, PQ Canada H7S 2G5.
§To whom reprint requests should be addressed.

***In Vitro* Transcription Assays and RNase-Cleavage Assays.** Promoter-independent transcription assays using poly(rC) or denatured calf-thymus DNA as a template were performed as described (7, 25). TFIIS-stimulated readthrough assays and ternary-complex cleavage assays using poly(dC)-tailed human histone H3.3 DNA as a template were performed as described (24). Briefly, RNA transcripts were pulse-labeled at the 5′ ends by incubating in the presence of 33.3 nM [α-^{32}P]-CTP [10.0 μCi, 3000 Ci/mmol (1 Ci = 37 GBq), DuPont] and 0.8 mM each ATP, GTP, and UTP for 75 sec, followed by elongating to the three intrinsic blocks in the presence of 0.1 mM unlabeled CTP and 100 μg/ml heparin (to prevent reinitiation) for 75 sec. Most of the nascent transcripts derived from this template were displaced from the template DNA strand (27). Ternary complexes stopped at the intrinsic arrest sites were treated in two different ways. The readthrough assay was done as follows. Yeast TFIIS was added, transcription was allowed to proceed for the designated intervals, and the resulting transcripts were collected by ethanol precipitation and were resolved by electrophoresis on a 7% polyacrylamide gel (19:1 acrylamide/bisacrylamide) containing 8.3 M urea, 2 mM EDTA, and buffered with 89 mM Tris·borate. The transcripts were quantified by using PhosphorImager (Molecular Dynamics). The fraction of transcripts arrested was determined as [(TIa cpm) + (TIb cpm) + (TII cpm)]/[(TIa cpm) + (TIb cpm) + (TII cpm) + (run-off cpm)]. The cleavage reaction was done as follows. The stopped ternary complexes were purified from unincorporated nucleotides by two sequential Bio-gel 30 spin columns (Bio-Rad). TFIIS was then added to the purified complexes and at designated times aliquots were removed and the reaction was stopped by the addition of a buffer containing 80% formamide and 0.1% SDS. The transcripts were resolved as described above.

Native PAGE. Labeling of a truncated derivative of TFIIS protein (lacking the first 130-amino acid residues) with [γ-^{32}P]ATP was achieved by following a described procedure (23). This truncated TFIIS was fully active both *in vitro* (24) and *in vivo* (28) and was used because it is more stable than the full-length TFIIS following labeling with [γ-^{32}P]ATP. RNAPII-TFIIS complexes were formed in binding buffer (20 mM Tris·OAc, pH 7.5/10 μM $ZnCl_2$/10 mM DTT/10% glycerol) in a final volume of 10 μl. Following incubation on ice for 30 min, the complexes were resolved on 5% native polyacrylamide gels (37.5:1 acrylamide/bisacrylamide) run at a constant voltage of 100 volts at 4°C for 3 hr in a buffer that contained 50 mM Tris·Cl, 50 mM boric acid, 10 μM ZnSO4, and 1% glycerol (pH 8.3), at 4°C. The amount of TFIIS–RNAPII binary complexes was determined using the Molecular Dynamics PhophorImager with ^{32}P-labeled TFIIS blotted on to filter paper as a standard. The apparent dissociation constant (k_d) was determined using Scatchard analysis.

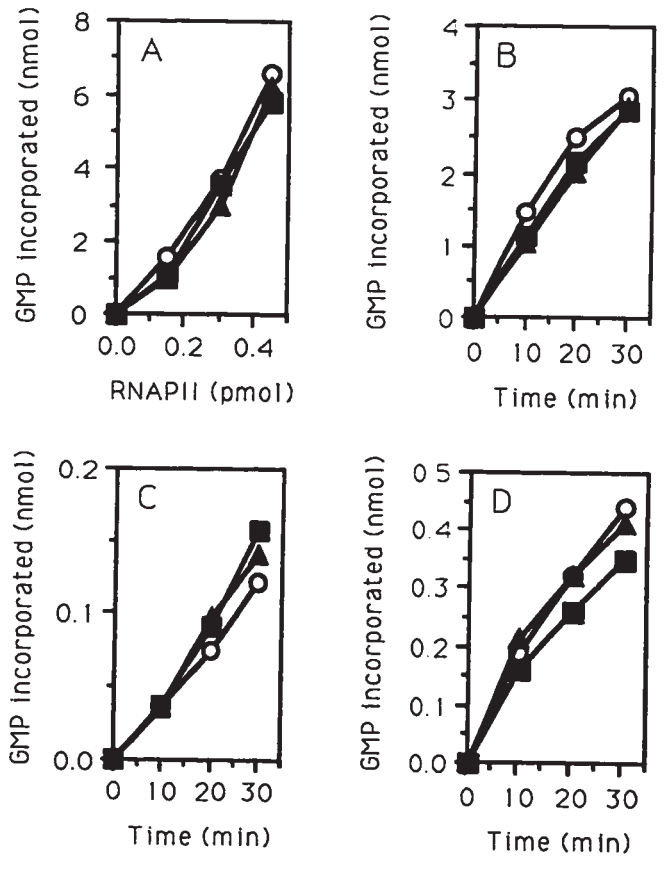


FIG. 1. Promoter-independent transcription assays of polymerases from wt (○), *rpo21-18* (I1237TRARV) (■), and *rpo21-24* (E1230K) (▲) strains using poly(rC) as a template. The polymerization activities of the wt and mutant enzymes were compared under normal reaction conditions (1 mM GTP and incubation at 30°C) (*A* and *B*) and under conditions of reduced GTP concentration (20 μM) (*C*) or elevated temperature (42°C) (*D*). The reactions in *A* were incubated for 30 min; 0.30 pmol of each polymerase was used in *B–D*.

RESULTS

Specific Polymerization Activities of wt and Mutant Polymerases on a Poly(rC) Template Are Comparable. The polymerization activities of the wt and mutant polymerases were compared on a poly(rC) template. On this template, purified RNAPII incorporates GMP in a template-dependent manner (29) under various conditions and produces high radioactive signals (at least 10-fold higher than background) that represent newly synthesized RNAs. Under normal conditions (1 mM GTP and 30°C), specific polymerization activities of wt and mutant polymerases were comparable (Fig. 1 *A* and *B*). The activities were estimated to be in the range of 500–620 nmol of nucleotide incorporated per min per mg of protein at 30°C. Since the mutant strains used in this study are 6AU-sensitive (which is probably related to the depletion of intracellular GTP pool by 6AU; ref. 15) as well as temperature-sensitive for growth, the purified polymerases were also assayed at reduced GTP concentration (20 μM: Fig. 1*C*) and at elevated reaction temperatures (42°C; Fig. 1*D*). Under these two sets of conditions, the polymerization activities of both wt and mutant

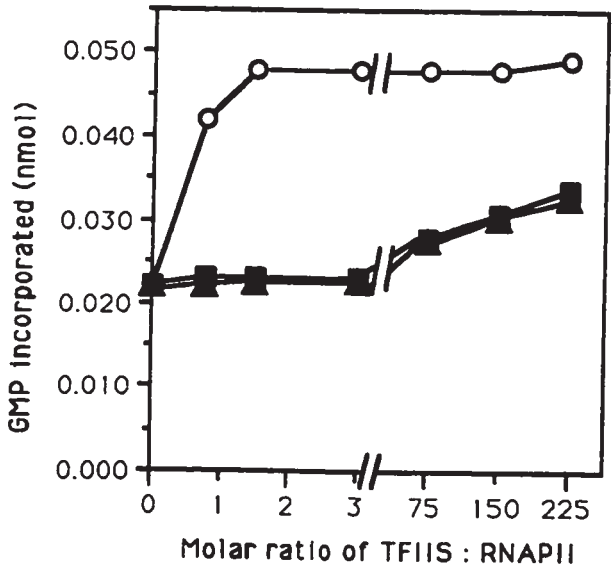


FIG. 2. Stimulation by TFIIS of RNAPII from wt (○), *rpo21-18* (I1237TRARV) (■), and *rpo21-24* (E1230K) (▲) strains. Denatured calf-thymus DNA was used as a template to assay 0.30 pmol of each polymerase. Reactions were incubated at 30°C for 20 min.

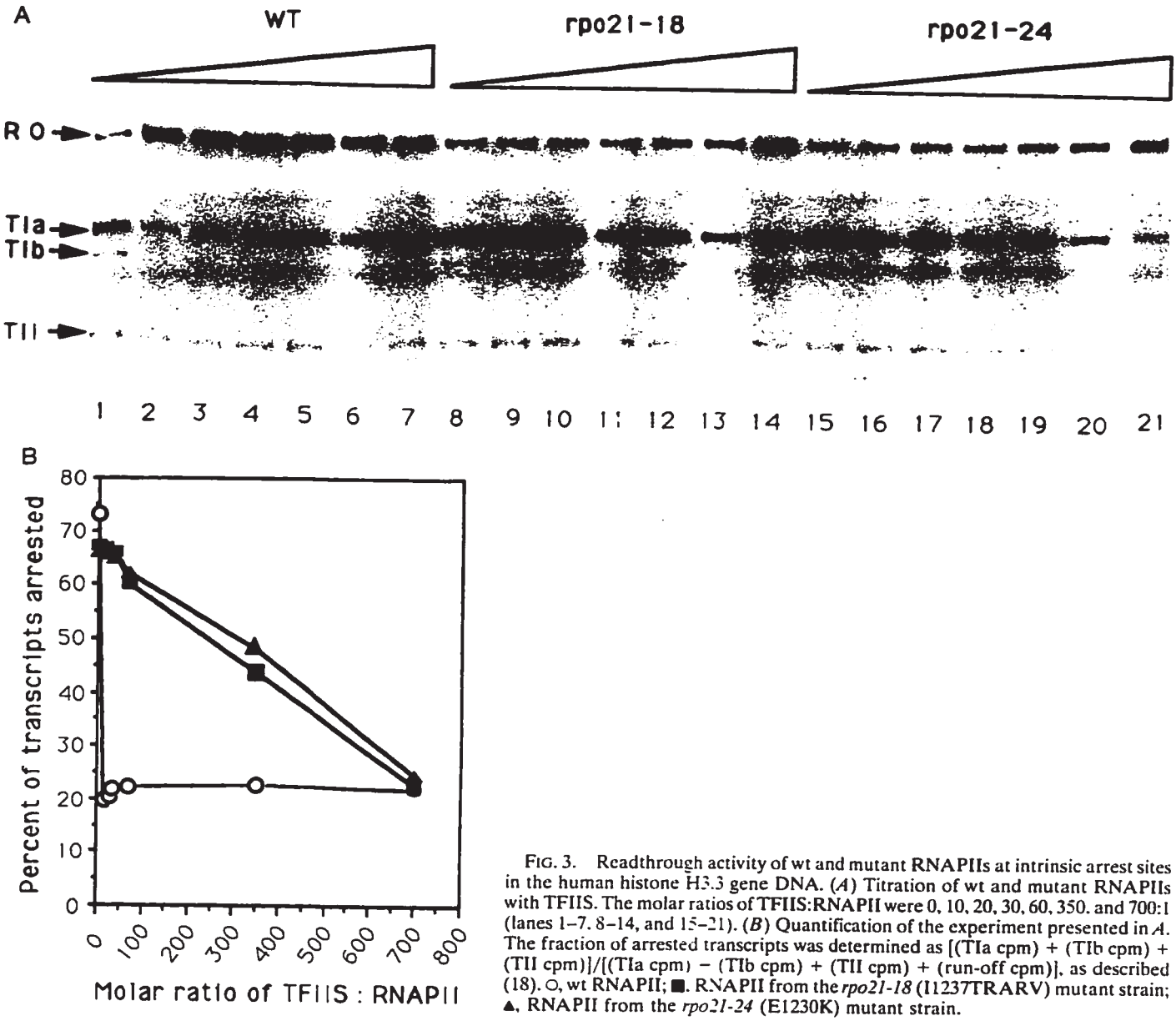


FIG. 3. Readthrough activity of wt and mutant RNAPIIs at intrinsic arrest sites in the human histone H3.3 gene DNA. (*A*) Titration of wt and mutant RNAPIIs with TFIIS. The molar ratios of TFIIS:RNAPII were 0, 10, 20, 30, 60, 350, and 700:1 (lanes 1–7, 8–14, and 15–21). (*B*) Quantification of the experiment presented in *A*. The fraction of arrested transcripts was determined as [(TIa cpm) + (TIb cpm) + (TII cpm)]/[(TIa cpm) − (TIb cpm) + (TII cpm) + (run-off cpm)], as described (18). ○, wt RNAPII; ■, RNAPII from the *rpo21-18* (I1237TRARV) mutant strain; ▲, RNAPII from the *rpo21-24* (E1230K) mutant strain.

polymerases were decreased to a similar extent. Therefore, the two *rpo21* mutations (I1237TRARV and E1230K) do not exacerbate the negative effect of nucleotide-substrate limitation or heat on the polymerization activity of RNAPII.

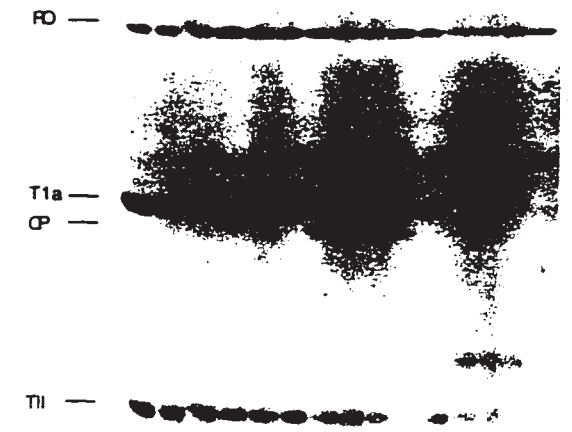
Polymerization Activity of Mutant RNA Polymerases Is Insensitive to Stimulation by TFIIS. Purified RNAPII transcribes denatured duplex DNA at a low efficiency, which can be stimulated 2- to 10-fold by TFIIS (6). Wt and mutant RNAPII preparations were tested for this property by using denatured calf thymus DNA as a template (Fig. 2). In the absence of TFIIS, the activities of wt and mutant polymerases were indistinguishable. Maximum stimulation of wt RNAPII by TFIIS was obtained at a TFIIS/RNAPII molar ratio of ≈1.5:1. In contrast, at the same TFIIS/RNAPII molar ratio, or slightly higher, no significant stimulation of the mutant enzymes was detected. At a molar ratio of at least 75:1, stimulation by TFIIS was observed. Even at the highest molar ratio tested (225:1), only ≈45% of the maximal stimulation of wt polymerase was achieved for the mutant enzymes.

High Concentrations of TFIIS Are Required by the Mutant RNA Polymerases to Approach Levels of wt Readthrough of Intrinsic Arrest Sites on Human Histone H3.3 DNA. Readthrough activity of wt and mutant polymerases at intrinsic blocks to elongation were compared in an assay using poly(dC)-tailed human histone H3.3 gene DNA as a template. It has been shown that purified yeast RNAPII recognizes and stops at the three intrinsic transcriptional arrest sites on the human histone H3.3 DNA but does not terminate (24, 25), and that yeast TFIIS can stimulate RNAPII to read through these sites (24).

Two observations were noteworthy at the outset. First, wt and mutant polymerases stop at the identical intrinsic arrest sites in the human histone H3.3 gene (Fig. 3*A*, lanes 1, 8, and 15); no additional arrested complexes were observed in the case of the two mutant polymerases. Second, comparable percentages of wt and mutant polymerase molecules were arrested at each of the three sites. For example, the proportion of molecules arrested at the TIa site following a 30-min incubation was 64% for the wt enzyme and 58% for both mutant enzymes (Fig. 3*A*, lanes 1, 8, and 15). The time course of the readthrough activity of the wt and mutant enzymes at the three intrinsic arrest sites was examined. In the absence of TFIIS, wt and mutant enzymes behaved indistinguishably following reaction times of 5, 10, and 30 min (data not shown). These observations, as well as those presented above, indicate that the two *rpo21* mutations do not affect the polymerization activity of RNAPII on different templates. Thus, the suggestion that the mutant enzymes are more prone to arrest *in vivo* (see above and ref. 20) is not likely to be correct.

A

Lane	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
TFIIS/ Pol	0					3					600				
Time (hr)	0	.2	1	2	2	0	.2	1	2	2	0	.2	1	2	2
NTPs		-			+		-			+		-			+



B

Lane	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
TFIIS/Pol	0					3					600				
Time (hr)	0	.2	1	2	2	0	.2	1	2	2	0	.2	1	2	2
NTPs		-			+		-			+		-			+

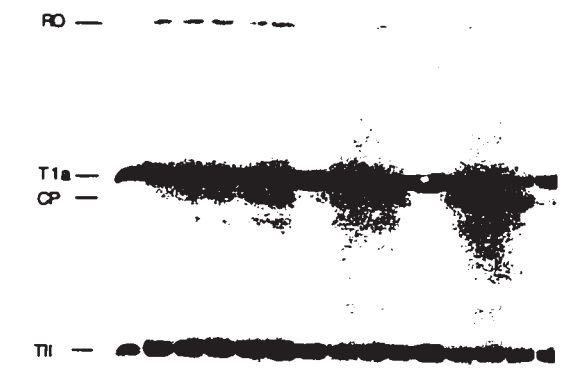


FIG. 4. Transcript cleavage induced by TFIIS. Ternary elongation complexes containing wt (*A*) or rpo21-24 RNAPII (*B*) were purified free of nucleotides and incubated for various times with no TFIIS (*A* and *B*, lanes 1–5), a 3:1 molar ratio of TFIIS/RNAPII (*A* and *B*, lanes 6–10) or a 600:1 molar ratio of TFIIS/RNAPII (*A* and *B*, lanes 11–15). Following a 2-hr incubation period, nucleotides were added to the complexes to allow for RNA-chain extension. The resulting transcripts were resolved by gel electrophoresis and visualized by autoradiography. The mobilities of the run off (RO), the T1a, the TII, and the cleaved product (CP) are indicated on the left.

A striking difference was observed between the wt and the two mutant polymerases in the presence of a 10-fold molar excess of TFIIS (Fig. 3*A*, lanes 2, 9, and 16). The fraction of arrested wt RNAPII was reduced to ≈20%. In contrast, at this molar ratio, the fraction of arrested mutant RNAPIIs remained unchanged, suggesting that at this level TFIIS does not stimulate mutant RNAPIIs to read through the arrest sites.

In an effort to mimic the *in vivo* observation that the *rpo21* mutants can be suppressed by overproduction of TFIIS (20), TFIIS-stimulated readthrough activity of the two mutant polymerases was tested at greatly elevated TFIIS/RNAPII molar ratios. As shown in Fig. 3, only massive addition of TFIIS to the mutant polymerases (over 600:1, TFIIS/RNAPII) resulted in readthrough comparable to that of the wt polymerase.

Lane	1	2	3	4	5		6	7	8	9	10	11	12	13	14	15
Pol	wt Pol II						RPO21-18					RPO21-24				
Time (min)	0	5	15	60	60		0	5	15	60	60	0	5	15	60	60
NTPs		-			+			-			+		-			+

FIG. 5. Intrinsic cleavage properties of wt and mutant polymerases. Ternary elongation complexes containing wt (lanes 1–5), rpo21-18 (I1237TRARV) (lanes 6–10), or rpo21-24 (E1230K) (lanes 11–15) RNAPII were purified free of nucleotides and incubated at pH 9.5 for 0, 5, 15, and 60 min to reveal the intrinsic cleavage activity of polymerase. Following 60 min, nucleotides were added to the complexes to allow for RNA-chain extension. The transcripts were resolved by gel electrophoresis and visualized by autoradiography. The mobilities of the run off (RO), the T1a, the TII and the cleaved product (CP) are indicated on the left.

Stimulation of Intrinsic RNase Activity in the Mutant RNA Polymerases Requires High Concentrations of TFIIS. The RNase activity (both intrinsic and TFIIS-stimulated) of wt and mutant RNAPII preparations was compared. Ternary elongation complexes containing RNA polymerase, DNA template and nascent RNA transcript arrested at the TIa, TIb, and TII sites and run-off transcripts were isolated free of nucleotides by centrifugal gel filtration (24). In the absence of TFIIS, incubation of the ternary complexes for as long as 2 hr revealed only very low levels of intrinsic RNase activity (Fig. 4 *A* and *B*, lanes 2–4). In this assay and under these conditions, the two mutant polymerases behaved identically to each other, thus only the data obtained from the Rpo21-24 (E1230K) enzyme are shown in Fig. 4. In the presence of a 3-fold molar excess of TFIIS, incubation for 0.2, 1, and 2 hr of ternary complexes containing wt RNAPII resulted in cleavage of a majority of the transcripts with 3′-ends at the TIa site (Fig. 4*A*, lanes 7–9). In contrast, incubation of ternary complexes containing the Rpo21-24 (E1230K) mutant polymerase (at 3:1, TFIIS/RNAPII) resulted in no more cleavage than was seen in the absence of TFIIS (Fig. 4*B*, compare lanes 7–9 to lanes 2–4). Therefore, the intrinsic RNase activity of the two mutant enzyme complexes is not stimulated by addition of this amount of TFIIS protein.

When the mutant enzymes were incubated in the presence of a *600*-fold molar excess of TFIIS (Fig. 4*B*, lanes 12–14), the RNA-cleavage ability was stimulated to a level that was similar to that shown by the wt enzyme in the presence of a 3-fold molar excess of TFIIS (compare Fig. 4*B*, lanes 12–14 to Fig. 4*A*, lanes 7–9). Therefore, provided sufficient TFIIS protein is present, the stimulated cleavage activity of the two mutant polymerases can approach normal levels.

Under the above readthrough and cleavage assay conditions (pH 8.0), only low levels of intrinsic cleavage activity of the polymerases were observed. Nevertheless, when the isolated ternary complexes were incubated in buffer at pH 9.5, the intrinsic RNase activity of the polymerase was clearly observ-

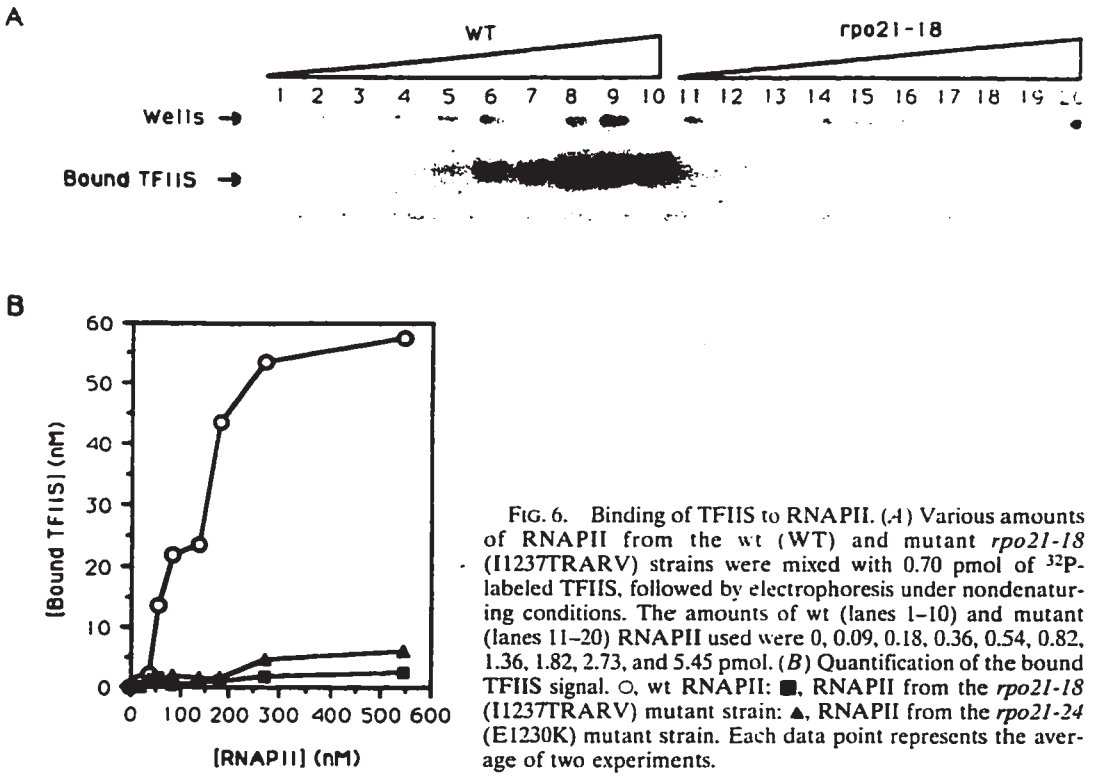


FIG. 6. Binding of TFIIS to RNAPII. (*A*) Various amounts of RNAPII from the wt (WT) and mutant *rpo21-18* (I1237TRARV) strains were mixed with 0.70 pmol of ^{32}P-labeled TFIIS, followed by electrophoresis under nondenaturing conditions. The amounts of wt (lanes 1–10) and mutant (lanes 11–20) RNAPII used were 0, 0.09, 0.18, 0.36, 0.54, 0.82, 1.36, 1.82, 2.73, and 5.45 pmol. (*B*) Quantification of the bound TFIIS signal. ○, wt RNAPII: ■, RNAPII from the *rpo21-18* (I1237TRARV) mutant strain: ▲, RNAPII from the *rpo21-24* (E1230K) mutant strain. Each data point represents the average of two experiments.

able. As shown in Fig. 5, in the absence of TFIIS, incubation for 5, 15, or 60 min of the ternary complexes (at pH 9.5) revealed that the intrinsic RNA-cleavage ability of the Rpo21-18 (I1237TRARV) mutant polymerase was comparable to that of the wt enzyme (Fig. 5, lanes 2–4 and lanes 7–9). The Rpo21-24 (E1230K) mutant polymerase showed a slightly diminished cleavage efficiency (Fig. 5, lanes 12–14). Therefore, we suggest that the *rpo21* mutations do not affect the intrinsic RNA-cleavage ability of RNAPII ternary complexes.

Mutant RNA Polymerases Have Reduced Binding Affinity for TFIIS. The behavior of the mutant enzymes in the *in vitro* readthrough and cleavage assays suggests that the *rpo21* mutations may affect the interaction between RNAPII and TFIIS. Therefore, the physical interaction between TFIIS and wt or mutant RNAPIIs was examined by electrophoresis under nondenaturing conditions (Fig. 6). Various amounts of RNAPII were incubated with ^{32}P-labeled TFIIS to form RNAPII/TFIIS complexes. The complexes were then resolved on a 5% native polyacrylamide gel. Under the electrophoretic conditions used in this experiment (pH 8.3), free TFIIS migrates out of the wells toward the cathode and the RNAPII/TFIIS complex toward the anode (Fig. 6*A*, lanes 1–10). Quantification of the signal corresponding to the bound TFIIS indicated that more TFIIS bound to wt RNAPII than to the two mutant enzymes (Fig. 6*B*), suggesting that the interaction between TFIIS and wt RNAPII was stronger than that between TFIIS and mutant RNAPIIs. The apparent dissociation constant (k_d) for the interaction between TFIIS and wt RNA polymerase was estimated to be in the range of $0.8–1.5 \times 10^{-7}$ M. The apparent dissociation constant for the binding of TFIIS to the two mutant enzymes was estimated to be in the range of $1.0–5.0 \times 10^{-5}$ M. Binding of TFIIS to the mutant polymerases was observed only at high levels of the mutant RNAPs (data not shown). We conclude that the *rpo21* mutations result in a reduction in the binding affinity of the mutant enzymes for TFIIS by more than 50-fold. Thus, they identify a site on RNAPII that is important for TFIIS binding.

DISCUSSION

We have observed that two mutations, which lead to 6AU-sensitive growth and affect amino-acid residues that are located between the highly conserved regions G and H of the largest subunit of RNAPII (Rpo21p; ref. 20), reduce the binding affinity of the mutant enzymes for TFIIS by more than 50-fold. As a consequence, very large amounts of TFIIS protein are required to stimulate the arrested mutant enzymes in cleavage of nascent RNAs and readthrough of intrinsic arrest sites.

These *in vitro* observations offer a plausible explanation for the 6AU-sensitive phenotype of the *rpo21* mutants and the suppression of the 6AU sensitivity by increased gene dosage of *PPR2*. Addition of 6AU to a yeast culture results in more than a 10-fold depletion of the intracellular GTP pool and a 2- to 3-fold depletion of the UTP pool (15). Several investigators have observed that the efficiency with which elongating RNAP becomes paused at intrinsic pause sites is related directly to the elongation rate of transcribing RNA polymerase and to NTP substrate concentration; a reduction in the concentration of one or all of the four NTPs enhances arrest of RNA polymerase (16–19). Therefore, the ultimate effect of 6AU on growing cells might be an increase in the number of transcribing RNAPII molecules that arrest or pause on the DNA template. When this occurs, cell growth becomes dependent on TFIIS. With the help of TFIIS, wt RNAPII might be able to complete a sufficient number of transcripts to permit growth in the presence of 6AU, whereas the mutant enzymes might not because the interactions between TFIIS and the mutant enzymes are weakened. An increase in the gene dosage of wt *PPR2* leads to increased production of yeast TFIIS protein and an elevated intracellular concentration of TFIIS. By driving the assembly of a sufficient number of active RNAPII–TFIIS complexes, an elevated *PPR2* gene dosage compensates for the reduced binding affinity of mutant RNAPII for TFIIS and allows growth.

Six of the seven 6AU-sensitive *rpo21* mutants (including the two described in this study, E1230K and I1237TRARV) are also temperature-sensitive (ts) for growth (20), which might also be related to reduced binding of the mutant enzymes for TFIIS. We have observed that the stopping efficiency of wt RNAPII at intrinsic arrest sites is increased substantially at elevated temperature (unpublished data). For example, at 38°C, the fraction of arrested wt RNAPII at the first and

weaker arrest site in the human histone H3.3 gene, TII, increases to ≈50% from the 15–25% seen at 30°C (data not shown). Therefore, raising the temperature may achieve a similar effect as that generated by adding 6AU to yeast growth medium—i.e., more transcribing RNAPII molecules may stop on the DNA template, rendering cell growth dependent on TFIIS function. However, since the ts phenotype of *rpo21* mutants cannot be suppressed by overproduction of TFIIS (20), one would have to assume that at a high temperature, the interactions between TFIIS and the mutant enzymes are so weak that they cannot be restored. Alternatively, alterations in the G–H region of Rpo21p that confer 6AU-sensitivity could also affect other RNAPII functions, such as promoter-dependent initiation, termination, or assembly of the enzyme. Two observations support this possibility. First, the *ppr2Δ* mutant strain has a near wt growth rate at the permissive temperature of 30°C, whereas one of the 6AU-sensitive *rpo21* strains, *rpo21-7* (T1141TSSSS), has a very slow growth rate at permissive temperature (20). Second, at least one ts mutation in the G–H region of Rpo21p (C1240Y) has been shown to affect the assembly of RNAPII at a nonpermissive temperature (30).

All seven of the *rpo21* mutations leading to 6AU-sensitive growth affect amino acid residues that are clustered between highly conserved regions G and H of Rpo21p (20–22). We suggest that our *rpo21* mutations might weaken the interaction between TFIIS and RNAPII by altering a site on RNAPII that is involved in direct interaction between TFIIS and RNAPII. The nature of the rpo21-24 (E1230K), which changes a glutamic acid to a lysine, suggests that a portion of this TFIIS binding site may be negatively charged. TFIIS might be expected to bind to a negatively charged surface because TFIIS is very basic. The RNAPII/TFIIS interaction would thus be predicted to be partly ionic in nature, which is consistent with the observations that the RNAPII/TFIIS interaction is known to be disrupted by low to moderate salt concentrations. However, we cannot rule out the possibility that the mutations induce conformational changes that reduce indirectly the binding of TFIIS to RNAPII.

Conserved region B of Rpo21p has also been implicated in TFIIS binding (31); a fusion protein containing a fragment of region B, as well as a monoclonal antibody thought to be directed to region B, when added at high concentrations partially inhibit TFIIS-stimulated transcription. However, the interpretation of those experiments is not entirely clear, since no direct interaction between the polymerase domain and TFIIS was measured and the inhibition of TFIIS activity was not reported to be overcome by the addition of excess TFIIS. Nevertheless, if both conserved regions B and G–H do play a role in TFIIS interaction, these two domains might be in close proximity in the RNA polymerase II structure.

We thank David Jansma for critical reading of the manuscript and Deming Xu for help with the preparation of the figures. This work was supported by grants from the Medical Research Council (Canada) to J.D.F. and A.M.E.

1. Aso, T., Conaway, J. W. & Conaway, R. C. (1995) *FASEB J.* **9**, 1419–1428.
2. Kerppola, T. K. & Kane, C. M. (1991) *FASEB J.* **5**, 2833–2842.
3. Sekemizu, K., Kobayashi, N., Mizuno, D. & Natori, S. (1976) *Biochemistry* **15**, 5064–5070.
4. SivaRaman, L., Reines, D. & Kane, C. M. (1990) *J. Biol. Chem.* **265**, 14554–14560.
5. Reines, D. & Mote, Jr., J. (1992) *Proc. Natl. Acad. Sci. USA* **90**, 1917–1921.
6. Agarwal, K., Back, K.-H., Jean, C.-J., Miyamoto, K., Ueno, A. & Yoon, H.-S. (1991) *Biochemistry* **30**, 7842–7851.
7. Sopta, M., Carthew, R. W. & Greenblatt, J. (1985) *J. Biol. Chem.* **260**, 10353–10360.
8. Izban, M. G. & Luse, D. S. (1993) *J. Biol. Chem.* **268**, 12874–12885.
9. Guo, H. & Price, D. H. (1993) *J. Biol. Chem.* **268**, 18762–18770.
10. Price, D. H., Sluder, A. E. & Greenleaf, A. L. (1989) *Mol. Cell. Biol.* **9**, 1465–1475.
11. Bradsher, J. N., Tan, S., McLaury, H.-J., Conaway, J. W. & Conaway, R. C. (1993) *J. Biol. Chem.* **268**, 25594–25603.
12. Conaway, J. W. & Conaway, R. C. (1989) *J. Biol. Chem.* **264**, 2357–2362.
13. Shilatifard, A., Lane, W. S., Jackson, K. W., Conaway, R. C. & Conaway, J. W. (1996) *Science* **271**, 1873–1876.
14. Aso, T., Shilatifard, A., Conaway, J. W. & Conaway, R. C. (1996) *J. Clin. Invest.* **97**, 1561–1569.
15. Exinger, F. & Lacroute, F. (1992) *Curr. Genet.* **22**, 9–11.
16. Reisbig, R. R. & Hearst, J. E. (1981) *Biochemistry* **20**, 1907–1918.
17. Kassavetis, G. A. & Chamberlin, M. J. (1981) *J. Biol. Chem.* **256**, 2777–2786.
18. Kadesch, T. R. & Chamberlin, M. J. (1982) *J. Biol. Chem.* **257**, 5286–5295.
19. Kerppola, T. K. & Kane, C. M. (1990) *Biochemistry* **29**, 269–278.
20. Archambault, J., Lacroute, F., Ruet, A. & Friesen, J. D. (1992) *Mol. Cell. Biol.* **12**, 4142–4152.
21. Allison, L. A., Moyle, M., Shales, M. & Ingles, C. J. (1985) *Cell* **42**, 599–610.
22. Jokerst, R. S., Weeks, J. R., Zehring, W. A. & Greenleaf, A. L. (1989) *Mol. Gen. Genet.* **215**, 266–275.
23. Blanar, M. A. & Rutter, W. J. (1992) *Science* **256**, 1014–1018.
24. Christie, K. R., Awrey, D. E., Edwards, A. M. & Kane, C. M. (1994) *J. Biol. Chem.* **269**, 936–943.
25. Edwards, A. M., Kane, C. M., Young, R. A. & Kornberg, R. D. (1991) *J. Biol. Chem.* **266**, 71–75.
26. Bradford, M. M. (1976) *Anal. Biochem.* **72**, 248–254.
27. Reines, D., Wells, D., Chamberlin, M. J. & Kane, C. M. (1987) *J. Mol. Biol.* **196**, 299–312.
28. Nakanishi, T., Shimoaraiso, M., Kubo, T. & Natori, S. (1995) *J. Biol. Chem.* **270**, 8991–8995.
29. Ruet, A., Sentenac, A. & Fromageot, P. (1978) *Eur. J. Biochem.* **90**, 325–330.
30. Kolodziej, P. A. & Young, R. A. (1991) *Mol. Cell. Biol.* **11**, 4669–4678.
31. Rappaport, J., Cho, K., Saltzman, A., Prenger, J., Golomb, M. & Weinmann, R. (1988) *Mol. Cell. Biol.* **8**, 3136–3142.

Appendix IV

Transcription elongation through DNA arrest sites: A multi-step process involving both RNA polymerase II subunit RPB9 and TFIIS.

Awrey, D. E., Weilbaccher, R.G., Hemming, S. A., Orlicky, S. M., Kane, C. M., and Edwards, A. M. (1997).
Journal of Biological Chemistry, vol. 272, pp. 14747-14754.

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 272, No. 23, Issue of June 6, pp. 14747–14754, 1997
Printed in U.S.A.

Transcription Elongation through DNA Arrest Sites

A MULTISTEP PROCESS INVOLVING BOTH RNA POLYMERASE II SUBUNIT RPB9 AND TFIIS*

(Received for publication, October 24, 1996, and in revised form, April 2, 1997)

Donald E. Awrey, Rodney G. Weilbaecher‡, Sally A. Hemming, Stephen M. Orlicky, Caroline M. Kane‡§, and Aled M. Edwards¶

From the Cancer Research Group, Institute for Molecular Biology and Biotechnology, McMaster University, Hamilton, Ontario L8N 3Z5, Canada and the ‡Department of Molecular and Cell Biology, University of California, Berkeley, California 94720

The role of yeast RNA polymerase II (pol II) subunit RPB9 in transcript elongation was investigated by examining the biochemical properties of pol II lacking RPB9 (pol IIΔ9). The maximal rate of chain elongation was nearly identical for pol II and pol IIΔ9. By contrast, pol IIΔ9 elongated more efficiently through DNA sequences that signal the elongation complex to pause or arrest. The addition of purified recombinant RPB9 to pol IIΔ9 restored the elongation properties of the mutant polymerase to those of the wild-type enzyme. Arrested pol IIΔ9 complexes were refractory to levels of TFIIS that promoted maximal read-through with pol II. However, both pol II and pol IIΔ9 complexes stimulated with TFIIS undergo transcript cleavage, confirming that transcript cleavage and read-through activities can be uncoupled. Our observations suggest that both TFIIS and RPB9 are required to stimulate the release of RNA polymerase II from the arrested state.

The mRNA transcription machinery in eukaryotes is a complex of more than 30 different polypeptides of which 12 polypeptides are tightly associated with RNA polymerase (pol)[1] II. The largest two subunits of pol II, which are thought to harbor catalytic activity, are related to the largest subunits of the other nuclear RNA polymerases (1, 2). Each of the remaining pol II subunits is conserved throughout the *Eukarya* and more than half are also conserved in the *Archae* (3).

In *Saccharomyces cerevisiae*, the genes encoding the twelve subunits of pol II have been cloned and sequenced. Five of the ten small subunits of yeast pol II (RPB5, RPB6, RPB8, RPB10, and RPB12) are common to all three nuclear RNA polymerases (4, 5). Several remaining pol II subunits have homologues in pol I and pol III. RPB11 (6) is related to AC19, a subunit shared by pols I and III (7). RPB7 is similar to the pol III subunit C25 (8). RPB9 is related to the pol I subunit, A12.2 (9). At least six pol II small subunits from *S. cerevisiae* are functionally interchangeable with human subunits (RPB6, RPB7, RPB8, RPB9, RPB10, and RPB12) (10–12). Genetic analysis indicates that only two yeast pol II subunits, RPB4 and RPB9, are not essential for cell viability (13, 14).

In yeast, deletion of RPB9 results in mild temperature sensitivity and relatively normal levels of transcription *in vivo*. However, for most genes examined, the selectivity of the site of transcription initiation is altered, with new start sites shifted upstream relative to wild-type sites (15–17). The transcription initiation phenotype can be recapitulated *in vitro*, and addition of recombinant purified RPB9 (rRPB9) restores wild-type start site selection (15). RPB9 from *S. cerevisiae* is a 122-amino acid polypeptide that contains two zinc binding domains (14). The COOH-terminal zinc binding domain shares 25% sequence identity with that of the general transcript elongation factor TFIIS (18) and is predicted to adopt a zinc ribbon fold (17–19). This domain is required for the function of RPB9 in start site selection[2] and is required within TFIIS for elongation stimulation (20).

The homology of RPB9 to TFIIS suggested that RPB9 may play a role in transcript elongation. In addition, the altered start site selection of pol II lacking RPB9 (pol IIΔ9) suggested some alteration in positioning the catalytic center of the polymerase, and a mobile catalytic center is a prominent feature of several recent models for transcript elongation (21–24). Thus the elongation properties of highly purified pol II and pol IIΔ9 were compared. The availability of purified rRPB9 and TFIIS was exploited to characterize the functional interactions among pol II and these factors during the transcription elongation process.

MATERIALS AND METHODS

Purification of Yeast pol IIΔ9 and pol II—RNA polymerase II lacking RPB9 (pol IIΔ9) was purified from the *S. cerevisiae* strain WY9 (14). pol II and pol IIΔ9 were purified as described previously (25), except that 10 μM $ZnCl_2$ was included in all buffers. The polymerase preparations were stored at −80 °C in 10 mM Tris-HCl, pH 7.9, 40 mM ammonium sulfate, 10% glycerol, 2.5 mM DTT, 100 μM EDTA, and 10 μM $ZnCl_2$.

Expression and Purification of Recombinant RPB9—The RPB9 coding region (14) was subcloned using polymerase chain reaction amplification to generate the coding region with an *Nde*I restriction site at the 5′ end of the sequence and a *Bam*HI site at the 3′ end. The amplified insert was then cloned between the *Nde*I and the *Bam*HI sites of the pET15b bacterial expression vector (Novagen, Madison, WI) to generate a fusion protein between RPB9 and an NH_2-terminal hexahistidine tag. BL21(DE3) cells containing the pET15bRPB9 plasmid were grown to A_{600} of 0.8 at 37 °C. The RPB9 protein expression was induced at 30 °C by the addition of 0.5 mM isopropyl-1-thio-β-D-galactopyranoside. The cells were harvested 3 h postinduction by centrifugation, resuspended in 50 mM Hepes, pH 7.5, 10% sucrose, 100 mM NaCl, 10 μM $ZnCl_2$, 1 mM PMSF, and 1 mM benzamidine (2 ml/g of cells), frozen, thawed, and then lysed by sonication at 4 °C. All subsequent steps were performed at 4 °C. The lysate was clarified by centrifugation at 55,000 × *g* for 30 min in a Beckman SW28 rotor, and then NaCl was added to a final concentration of 500 mM. The supernatant was passed through a DE52 column (2.5 × 5 cm; Whatman, Maidstone, UK) and loaded directly onto His-

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

§ Supported by the National Science Foundation.

¶ Supported by the Medical Research Council of Canada. Research Scholar of the Medical Research Council.

[1] The abbreviations used are: pol, polymerase; rRPB, recombinant RPB9; DTT, dithiothreitol; PMSF, phenylmethylsulfonyl fluoride; GST, glutathione *S*-transferase; CAPS, 3-(cyclohexylamino)propanesulfonic acid; HMK, heart myosin kinase.

[2] M. Hampsey, personal communication.

bind resin (2.5 × 2.5 cm; Novagen, Madison, WI), previously charged with 200 mM $NiSO_4$ and equilibrated in purification buffer (50 mM Hepes, pH 7.5, 10 μM $ZnCl_2$, 10% glycerol, 1 mM PMSF, and 1 mM benzamidine) containing 5 mM imidazole. The His-bind column was washed with 5 column volumes of purification buffer containing 500 mM NaCl and 50 mM imidazole, pH 7.5. RPB9 was eluted from the column with purification buffer containing 500 mM NaCl and 300 mM imidazole, pH 7.5. The eluate was dialyzed for 8 h against purification buffer containing 100 mM NaCl and 10 mM DTT but lacking PMSF and benzamidine. The dialysate was treated with bovine thrombin (3 μg of thrombin/mg of RPB9 fusion) for 4 h at room temperature, diluted 2-fold with purification buffer containing 10 mM DTT, passed through a Poros S column (0.75 × 10 cm; PerSeptive Biosystems, Cambridge, MA), and loaded onto a 2-ml hydroxyapatite fast protein liquid chromatography column (Bio-Rad). The column was developed with a 5 mM to 150 mM gradient of sodium phosphate, pH 7.5 in 50 mM NaCl, 10 mM DTT, 10 μM $ZnCl_2$, 1 mM PMSF, and 1 mM benzamidine, and RPB9 eluted at 50 mM phosphate. The purified protein was dialyzed against 5 mM Hepes, pH 7.5, 100 mM NaCl, 10 μM $ZnCl_2$, and 10 mM DTT and stored frozen at −70 °C.

Expression and Purification of Recombinant GST-RPB9 Fusion—To generate a fusion protein of GST and RPB9, the RPB9 coding region (14) was subcloned using polymerase chain reaction amplification to generate the coding region with a *Bam*HI restriction site at both the 5′ and 3′ ends. The amplified insert was then cloned into the *Bam*HI site of the pGEX-2T bacterial expression vector (Pharmacia Biotech Inc.). The GST-RPB9 fusion protein was expressed as described above for recombinant RPB9 using HB101 cells. The cells were harvested and lysed as described previously, with the exception that the NaCl concentration was maintained at 100 mM. The supernatant was passed through a DE52 column (2.5 × 2 cm; Whatman) and loaded directly onto a glutathione-Sepharose-4B column (2.5 × 3 cm; Pharmacia) previously equilibrated with Buffer A (10 mM sodium phosphate, 1.8 mM potassium phosphate, 2.7 mM KCl, 140 mM NaCl 10 mM DTT, 1 mM PMSF, and 1 mM benzamidine). The glutathione column was washed with 10 column volumes of Buffer A followed by 5 column volumes of Tris buffer (50 mM Tris, pH 8.0, 10 mM DTT, 1 mM PMSF, and 1 mM benzamidine). The GST-RPB9 fusion protein was eluted from the column with Tris buffer containing 10 mM glutathione. The purified protein was dialyzed against 5 mM Hepes, pH 7.5, 100 mM NaCl, 10% glycerol, 10 μM $ZnCl_2$, and 10 mM DTT and stored frozen at −70 °C.

Purification of Yeast TFIIS—Yeast TFIIS was cloned and expressed as a fusion protein containing an amino-terminal hexahistidine tag in bacterial cells (26), and the expressed protein was purified as described below. All procedures were performed at 4 °C. Bacterial cells (2L) expressing the yeast TFIIS protein were lysed using a French pressure cell in Buffer A (20 mM Hepes, pH 7.5, 10% glycerol, 10 μM $ZnCl_2$, 1 mM PMSF, and 1 mM benzamidine) containing 100 mM NaCl, 10 mM DTT, and 1 mM EDTA. The lysate was clarified by centrifugation at 55,000 × *g* for 30 min in a Beckman SW28 rotor, and the supernatant was loaded onto a DE52 column (2.5 × 2.5 cm; Whatman). The flow through was loaded onto an SP Fast Flow column (1 × 8 cm; Pharmacia), washed with Buffer A containing 100 mM NaCl, and eluted with Buffer A containing 500 mM NaCl. The eluate was loaded directly onto a Poros MC column (0.75 × 10 cm; PerSeptive Biosystems, Cambridge, MA) equilibrated with $NiSO_4$. The column was washed with Buffer A containing 500 mM NaCl and 5 mM imidazole. Yeast TFIIS was eluted with Buffer A containing 500 mM NaCl and 500 mM imidazole, immediately dialyzed against Buffer A containing 50 mM NaCl and 10 mM DTT, and stored frozen at −70 °C.

Transcription Assays: Read-through and Nucleolytic Cleavage by Arrested Ternary Complexes—These procedures were carried out as described previously (26). Briefly, transcription by purified pol II was initiated from a 3′ deoxycytidine tailed template containing a human histone H3.3 gene fragment, which contains well characterized blocks to elongation (27). The RNA transcript was pulse-labeled at the 5′ end by incubating in the presence of 20 μCi of [α-^{32}P]CTP (3000 Ci/mmol, DuPont NEN), 0.8 mM GTP, ATP, and UTP for 75 s, followed by elongation to the blocks to elongation in the presence of 0.1 mM unlabeled CTP for 75 s. Ternary complexes stalled at these sites were treated in two different ways. For the read-through assay, yeast TFIIS was added, transcription was allowed to proceed for the designated intervals, and the resulting transcripts were collected by ethanol precipitation and resolved by electrophoresis on a 6% polyacrylamide (19:1 acrylamide:bisacrylamide), 8.3 M urea, Tris borate-EDTA gel. The transcripts were quantified by phosphorimaging (Molecular Dynamics, Sunnyvale, CA). For the cleavage reaction, the stalled ternary complexes were purified from unincorporated nucleotides by two sequential

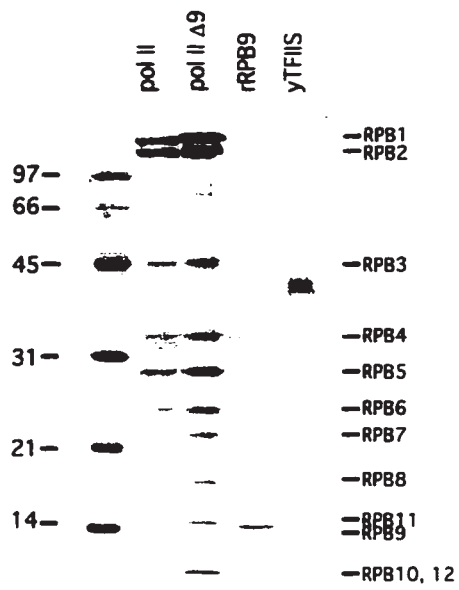


FIG. 1. **Purified wild-type pol II, pol IIΔ9, recombinant RPB9, and recombinant yeast TFIIS.** 5 μg of pol II and pol IIΔ9 and 1 μg of recombinant RPB9 and recombinant yeast TFIIS (*yTFIIS*) were resolved by denaturing gel electrophoresis in a 16% polyacrylamide gel and stained with Coomassie Blue. The size of the molecular mass standards (kDa) are indicated on the *left*, and the mobilities of the pol II subunits are shown on the *right*.

Bio-gel 30 spin columns (Bio-Rad). TFIIS was then added to the purified complexes, and incubation was continued at 30 °C. At designated time points, aliquots were removed, and the reaction was stopped by the addition of stopping buffer. The transcripts were resolved and quantified as described above.

For analysis of intrinsic cleavage, the standard cleavage assay described above was modified by including a third Bio-gel 30 spin column equilibrated with 70 mM CAPS, K^+, pH 9.5, 100 mM KCl, 5 mM $MgCl_2$, and 10% glycerol. Reactions were incubated at 30 °C, and aliquots were removed to an equal volume of stop buffer and resolved as described above.

Binding Assay between Yeast TFIIS and RNA Polymerase II—An oligonucleotide encoding the heart myosin kinase (HMK) consensus phosphorylation sequence (RRASVDF) was inserted into the *Nde*I site of the plasmid pET15bPPR2 (26), between the sequences encoding the thrombin cleavage site and TFIIS. The resulting TFIIS fusion protein was expressed and purified as described for the purification of yeast TFIIS. This protein was phosphorylated at the HMK consensus site using HMK (Sigma) and [γ-^{32}P]ATP according to the manufacturer's instructions. The labeled protein was dialyzed against Buffer A containing 10 mM DTT and 50 mM NaCl to remove the free radiolabel. The ^{32}P-labeled TFIIS and unphosphorylated TFIIS were equally active in assays of read-through and transcript cleavage. For the binding assay, 0.85 pmol of ^{32}P-HMK-TFIIS were incubated in a 5-μl volume for a minimum of 20 min with 0.1–4.5 pmol of pol II in buffer containing 50 mM Tris, pH 7.5, 10% glycerol, and 5 mM DTT and a final salt concentration between 40 and 50 mM KOAc. The binary TFIIS-polymerase complex was resolved from free yeast TFIIS by electrophoresis for 2–3 h at 100 V at 4 °C on a 5% polyacrylamide (30:0.8 acrylamide:bisacrylamide) gel containing 50 mM Tris, 50 mM borate, 10 μM $ZnSO_4$, and 1% glycerol. The amount of bound ^{32}P-HMK-TFIIS was quantified by phosphorimaging.

RESULTS

Purification of pol IIΔ9 and Recombinant RPB9—pol IIΔ9 was isolated using a combination of heparin, immunoaffinity, and anion exchange chromatography. This method yielded a purified enzyme devoid of RPB9 but containing all of the remaining subunits (Fig. 1). The yield of pol IIΔ9 enzyme was similar to that of pol II, approximately 0.8 mg/100 g of packed yeast. Purified pol IIΔ9 consistently demonstrated a 2–3-fold higher specific activity than pol II in assays of promoter-independent RNA chain elongation on denatured DNA templates.

However, the maximal elongation rate was indistinguishable between the two polymerases. The enzymes were also similar in the rate and extent of pyrophosphorolysis and the temperature dependence of promoter-independent assays of RNA synthesis (data not shown).

rRPB9 was expressed and purified from *Escherichia coli*. The yield of rRPB9 was 5 mg/liter of bacterial culture. We judged rRPB9 to be functional on the basis of two observations. First, the zinc content of rRPB9 was determined by atomic absorption spectroscopy to be 2.4 ± 0.4 mol of zinc/mol of rRPB9, in keeping with the predicted stoichiometry of zinc binding by RPB9.[3] Second, this RPB9 preparation restores accurate transcription start site utilization in nuclear extracts prepared from yeast cells lacking RPB9.[4]

Although rRPB9 was transcriptionally active, it readily formed aggregates. Dynamic light scattering studies of rRPB9 at 3 mg/ml revealed a polydisperse preparation whose average size corresponded to a pentamer. In an attempt to reduce the propensity to aggregate, another recombinant version of RPB9 was prepared, in the form of a fusion to GST. The purified GST-RPB9 was dimeric (in keeping with the dimeric nature of native GST), monodisperse in solution, and also active in assays of promoter-dependent transcription. However, if the GST moiety was removed from the purified GST-RPB9, the RPB9 that was generated once again aggregated. We could not prepare a version of RPB9 that was unaggregated. Therefore we used RPB9 fusion proteins for reconstitution experiments. In each such experiment, the source of the RPB9 is indicated, and the corresponding state of aggregation was therefore known.

pol IIΔ9 Elongates More Efficiently than pol II through an Intrinsic Block to Elongation—The observation that pol II and pol IIΔ9 have the same maximal elongation rate yet pol IIΔ9 synthesizes more RNA in chain elongation assays suggested that pol IIΔ9 pauses less frequently during elongation. To test this idea, we examined the behavior of pol IIΔ9 at well characterized blocks to elongation. Purified pol II stops *in vitro* at several sites within the human histone H3.3 first intron (27). The strongest block, TIa, has been used extensively to study the elongation properties of pol II from yeast and mammalian cells (26, 28–33). We observed that pol IIΔ9 transcribes through TIa with higher efficiency than pol II; 50–65% of pol IIΔ9 elongates past TIa compared with 15–30% for pol II. The differences in elongation efficiency are moderate but reproducible; these values reflect the range of TIa read-through in more than a dozen independent experiments with at least two different preparations of each enzyme.

To exclude the possibility that the purification of pol IIΔ9 for some reason results in a more efficiently elongating enzyme and that the elongation differences were indeed a property of RPB9, we tested whether purified RPB9 could restore the wild-type elongation properties of pol IIΔ9. The addition of increasing amounts of rRPB9 or GST-RPB9 to pol IIΔ9 resulted in decreased (wild-type) levels of read-through at the TIa intrinsic block to elongation (Fig. 2*A*). The decrease in transcription likely results from the reconstitution of wild-type polymerase activity by RPB9 and pol IIΔ9 because addition of greater than 200-fold molar excess of rRPB9 to either pol IIΔ9 or pol II did not decrease read-through below that observed with pol II alone. Wild-type elongation levels were restored to pol IIΔ9 by a 3–5-fold molar excess of GST-RPB9 and a 10-fold molar excess of rRPB9.

The addition of purified RPB9 to pol IIΔ9 not only decreased read-through activity but also decreased the total level of transcription from the dC-tailed templates (Fig. 2*B*). We observed that pol IIΔ9 synthesized 3–15 times more transcripts from the dC-tailed template than did either pol IIΔ9 + RPB9 or the wild-type enzyme, even though pol IIΔ9 was only 1.5–2-fold more active than the wild-type enzyme in the normal assays for promoter-independent chain elongation. RPB9 must inhibit transcription from these templates at an early stage in the reaction because the transcripts are pulse-labeled during the formation of the first 20 or so nucleotides. Using native gel electrophoretic mobility shift assays, we have shown that the two enzymes bind the tailed template with equal affinity (data not shown). Therefore, the inhibitory effect of RPB9 takes place after binding of the polymerase and before the synthesis of the first 20 nucleotides. The mechanism of inhibition remains to be established.

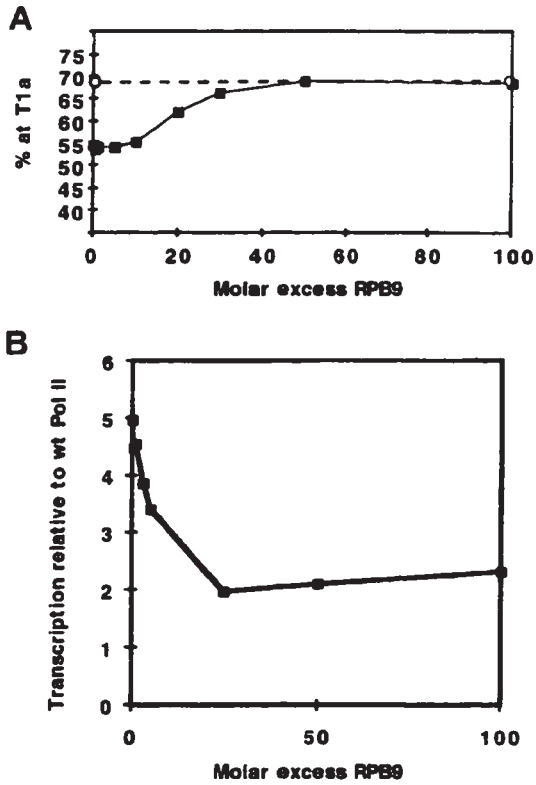


FIG. 2. **The effect of RPB9 on the efficiency of arrest by pol II[/dela]9 at the TIa intrinsic arrest site.** *A*, pol IIΔ9 and increasing amounts of RPB9 were preincubated for 20 min prior to transcription on the dC-tailed template containing the TIa arrest site. Reactions were stopped after 10 min, purified by ethanol precipitation, and resolved on a 6% acrylamide, 8.3 M urea, TBE gel (89 mM Tris-Cl, 89 mM borate, 2 mM EDTA). Transcripts were visualized by autoradiography and quantified with a Molecular Dynamics PhosphorImager. The efficiency of transcription arrest by the wild-type pol II (*dashed line*) and pol IIΔ9 (*solid line*) of a representative experiment are shown. *B*, the absolute level of transcription, as measured by the amount of radioactivity in TIa and run-off transcripts, was determined with increasing concentrations of rRPB9 for wild-type pol II and pol IIΔ9. The *solid line* represents the relative efficiency of transcription by pol IIΔ9 compared with wild-type pol II.

pol IIΔ9 Ternary Complexes Are Not Stimulated to Read-through Blocks to Elongation by the Elongation Factor TFIIS—The addition of rRPB9 to pol IIΔ9 causes more polymerase to arrest at TIa. The effect of this subunit apparently contrasts with that of the elongation factor TFIIS, which stimulates transcription through TIa, as well as several other elongation blocks. To explore the functional interactions between RPB9 and TFIIS, we compared read-through stimulation by TFIIS with both pol II and pol IIΔ9-containing arrested complexes. For the wild-type enzyme, a 3-fold molar excess of TFIIS maximally stimulates read-through of TIa by 30 min, at which time

[3] I. Donaldson, personal communication.
[4] M. W. Hull and N. A. Woychik, personal communication.

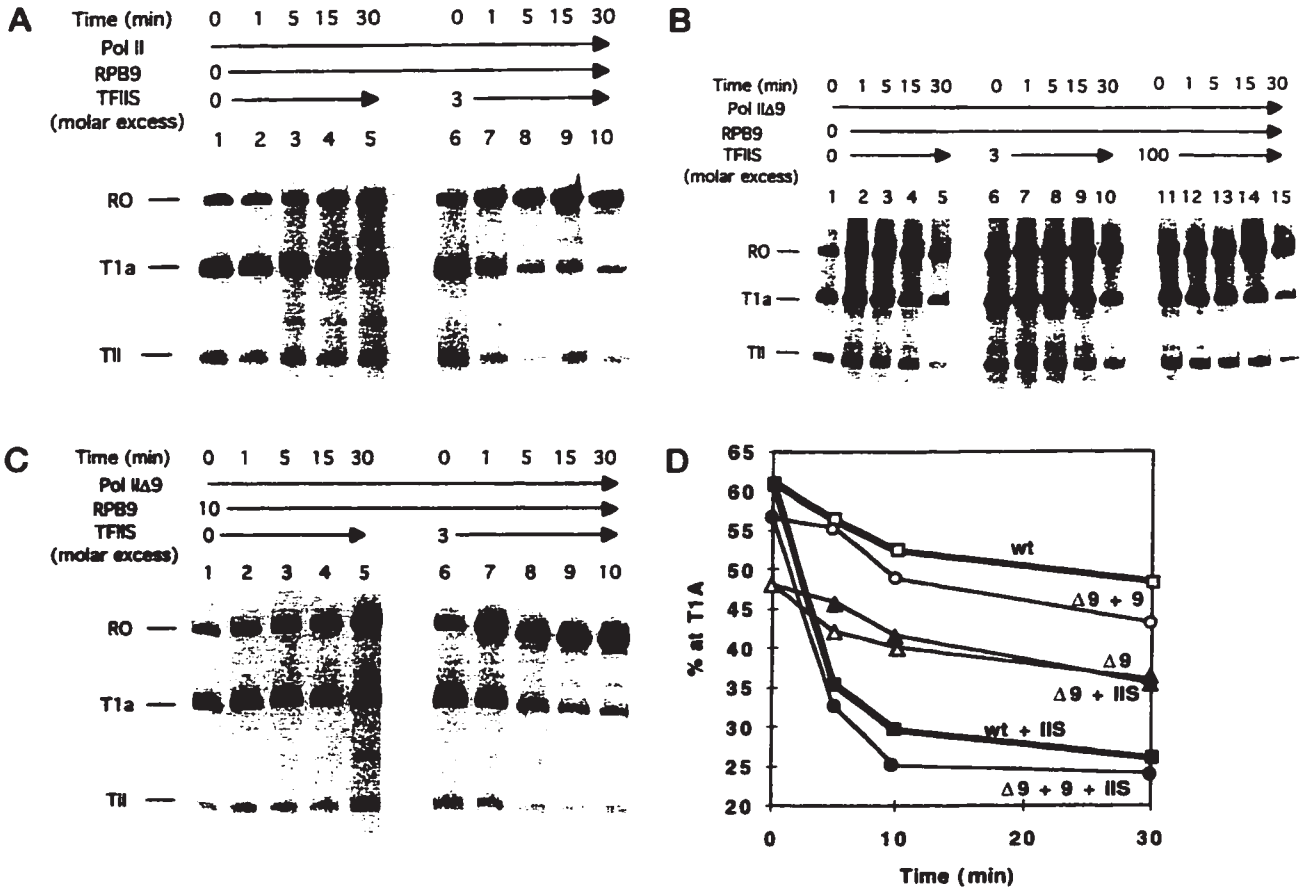


FIG. 3. **Stimulation of elongation through arrest sites by TFIIS.** Ternary elongation complexes were formed using pol II (*A*), pol IIΔ9 (*B*), or pol IIΔ9 reconstituted with a 10-fold molar excess of recombinant RPB9 (*C*) and incubated in the presence or the absence of the indicated amounts of TFIIS for 0, 1, 5, 15, and 30 min. The transcription products were resolved by gel electrophoresis and visualized by autoradiography. The relative mobilities of the TII, TIa, and run-off (*RO*) transcripts are indicated on the *left* of each panel. *D*, the percentage of transcripts arrested at TIa were determined for pol II (*squares*, *thick line*), pol IIΔ9 (*triangles*), and Δ9 RNA pol II reconstituted with recombinant RPB9 (*circles*) in the absence (*open symbols*) and the presence of TFIIS (*closed symbols*). The average of three reactions is shown. *wt*, wild type.

nearly two-thirds of the initially arrested complexes are reactivated (Fig. 3*A*). By contrast, arrested pol IIΔ9 ternary complexes were completely unresponsive to this level of TFIIS during a 30-min incubation (Fig. 3*B*, *lanes 6–10*; also Fig. 3*D* for quantification). The apparent decrease in the levels of the TIa transcript in Fig. 3*B* (*lane 5*) reflects a decreased level of radioactivity in the entire lane and not a TFIIS-independent read-through event.

Once again, to ensure that the inability of pol IIΔ9 to respond to these levels of TFIIS was caused by the lack of RPB9, we attempted to restore wild-type TFIIS responsiveness by the addition of purified RPB9. A 3-fold molar excess of TFIIS was added to pol IIΔ9 ternary complexes formed in the presence of rRPB9. The reconstituted polymerase (pol IIΔ9 + RPB9) responded to TFIIS with the same kinetics of read-through as did pol II (Fig. 3*C*, *lanes 6–10*). It was not necessary to preincubate the RPB9 with pol IIΔ9 to observe reconstitution of activity; identical results were obtained if RPB9 was added with TFIIS.

We tested if the pol IIΔ9 arrested complexes were completely refractory to TFIIS treatment or whether a large molar excess of TFIIS was able to promote read-through by pol IIΔ9-containing ternary complexes. pol IIΔ9 arrested complexes were incubated with a 100-fold molar excess of TFIIS, and at these levels the arrested complexes could be marginally stimulated to read-through the TIa arrest site, as determined by PhosphorImager analysis (Fig. 3*B*, *lanes 11–15*).

TFIIS-induced Transcript Cleavage in Arrested pol IIΔ9 Ternary Complexes—TFIIS-stimulated read-through requires a nascent transcript cleavage event (34–36). Because pol IIΔ9 appears to be deficient in responding to TFIIS, we tested whether pol IIΔ9 displayed aberrant cleavage activity by comparing cleavage by pol II and pol IIΔ9 complexes in response to TFIIS. As had been observed previously (26), the addition of a 3-fold molar excess of TFIIS to arrested wild-type elongation complexes rapidly generated shortened transcripts (C1 and C2) that were stably associated with active ternary complexes as shown by their ability to elongate upon the addition of nucleotides (Fig. 4*A*). The cleaved transcripts appear to be generated sequentially. The first cleaved product (C1) appeared within 1 min (Fig. 4*A*, *lane 3*), and over the course of 30 min a smaller product (C2) appeared, followed then by a range of shorter RNA species. Such a pattern of digestion has been observed previously for both yeast and mammalian enzymes (26, 34).

The pattern and kinetics of TFIIS-stimulated transcript cleavage by pol IIΔ9 differ significantly from that of pol II. Like pol II, pol IIΔ9 complexes also cleave the nascent RNA chain to C1 in the presence of a 3-fold molar excess of TFIIS. The pol IIΔ9-generated C1 cleavage products appear more slowly than do the pol II-generated C1 cleavage products (compare Fig. 4*A*, *lanes 3–6*, and Fig. 4*B*, *lanes 3–7*). In contrast to the pol II complexes, pol IIΔ9 complexes did not generate the second and subsequent cleavage products, even after 30 min or longer in the presence of TFIIS. To ensure that the differences in the cleavage reaction were due to the presence or the absence of RPB9, we tested if purified RPB9 would restore the wild-type cleavage pattern and kinetics. A 10-fold molar excess of rRPB9 was added to pol IIΔ9 complexes either prior or subsequent to arrest complex formation. The resulting enzyme (pol IIΔ9 +

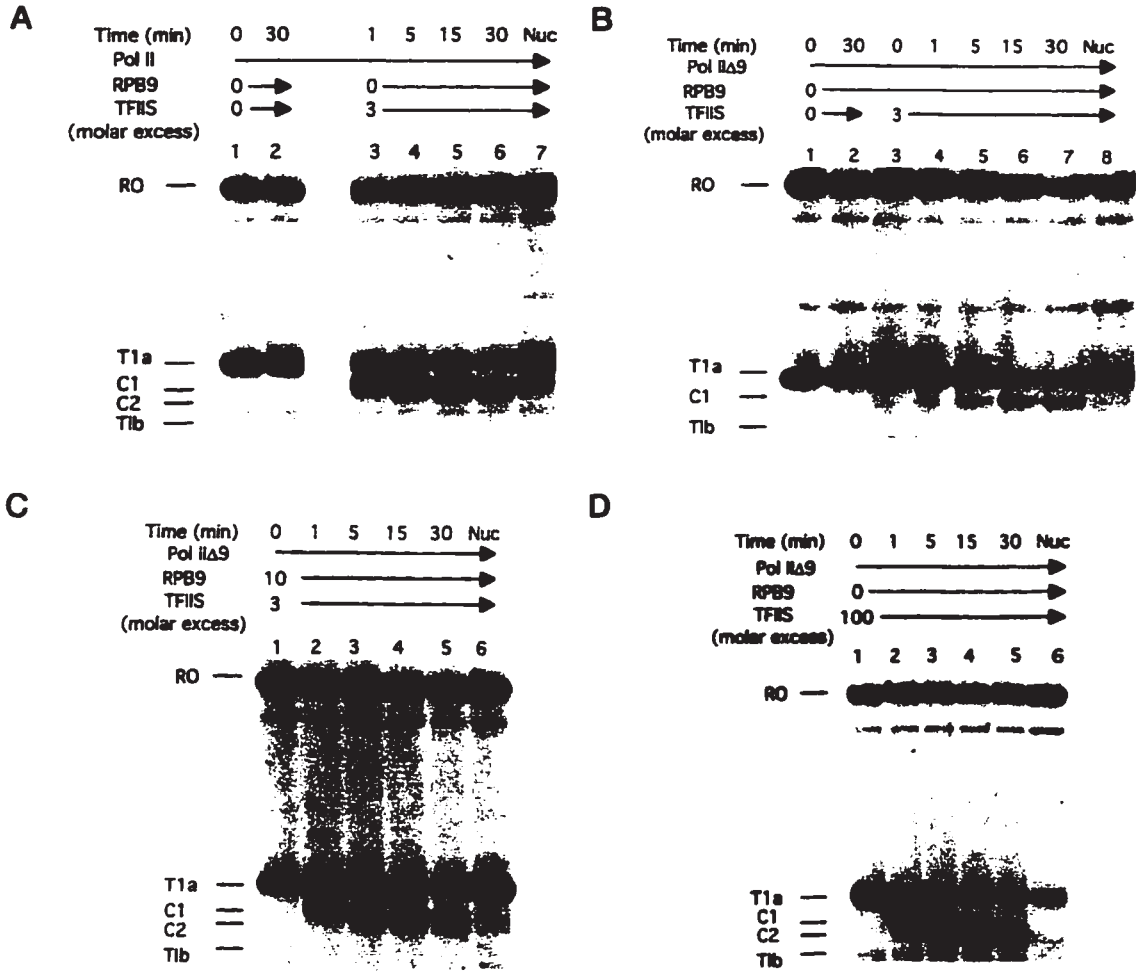


FIG. 4. **Time course and titration of cleavage activity by TFIIS.** Stalled ternary complexes containing pol II (*A*), pol IIΔ9 (*B* and *D*), or pol IIΔ9 reconstituted with rRPB9 (*C*) were incubated with a 3-fold (*A*, *B*, and *C*) or a 100-fold molar excess (*D*) of TFIIS for 0, 1, 5, 15, and 30 min. The transcription products were resolved by gel electrophoresis and visualized by autoradiography. The relative mobilities of the TIa, TII, and run-off (*RO*) transcripts and the first two cleavage products, C1 and C2, are indicated. The first two lanes in each of *panels A* and *B* demonstrate that no transcript cleavage occurs after 30 min in both pol II and pol IIΔ9 complexes in the absence of TFIIS. The lanes designated Nuc show the products of a 30-min cleavage reaction after the addition of nucleotides for an additional 10 min.

RPB9) displayed an identical response to TFIIS compared with pol II (Fig. 4*C*, *lanes 1–5*). The requirement for RPB9 in read-through could be diminished by the addition of a 100-fold molar excess of TFIIS (Fig. 3). Similarly, the addition of a 100-fold molar excess of TFIIS to arrested pol IIΔ9 complexes returned the cleavage patterns and kinetics to that seen in complexes formed with the wild-type enzyme (Fig. 4*D*). Therefore, in the absence of RPB9, the polymerase requires substantially higher TFIIS levels to effect cleavage and read-through.

The observed results would be explained if pol IIΔ9 was defective in catalyzing transcript cleavage. Thus the intrinsic cleavage activities of pol II and pol IIΔ9 were compared. We found that pol IIΔ9 did not appear to have any intrinsic defect in catalyzing transcript cleavage. Moderate elevation in pH can stimulate intrinsic cleavage by pol II,[5] similar to the intrinsic cleavage reaction observed with *E. coli* RNA polymerase (37). At basic pH, yeast pol II complexes halted at TIa cleave their transcripts in the absence of TFIIS. Although such intrinsic cleavage can occur at physiologic pH under some solution conditions, the conditions used to assay TFIIS-stimulated cleavage demonstrated no intrinsic cleavage over the course of 30 min (Fig. 4, *A* and *B*, *lanes 1* and *2*). In comparing intrinsic cleavage by pol II and pol IIΔ9 under mildly alkaline conditions, we observed no difference in the size of the cleavage products nor in the rate of transcript cleavage (Fig. 5). The intrinsic cleavage products correspond approximately to the mobilities of the TFIIS-stimulated cleavage products; however, direct determination of the cleavage increment has not been established. Extension of the intrinsic shortened transcripts at pH 9.5 can result in synthesis back to the original TIa site. Some intrinsic shortened transcripts are elongated slightly past the TIa site (see Fig. 5, *Nuc lane*); the mechanism of this elongation event remains to be established, and further characterization of the intrinsic cleavage activity of pol II will be presented elsewhere.[5] In summary, the similarity of pol II and pol IIΔ9 in the intrinsic cleavage reaction suggests that pol IIΔ9 is deficient in some other aspect of TFIIS-mediated read-through.

pol II and pol IIΔ9 Have Equivalent Affinity for Yeast TFIIS—A trivial explanation for our observations would be that the reduced response of pol IIΔ9 to TFIIS arises from a reduced affinity of pol IIΔ9 for TFIIS. Therefore, the affinity of pol II and pol IIΔ9 for TFIIS was examined using an electrophoretic mobility shift assay that measures the association of radiolabeled TFIIS with polymerase in nondenaturing polyacrylamide gels (38). In this assay, both pol II and pol IIΔ9 bind TFIIS with equivalent affinity (approximately 80 nM) (Fig. 6). Binding of TFIIS to either polymerase saturates at a 1:1 molar ratio and is competed by the addition of unlabeled TFIIS (data not shown). In this gel system, pol II and the pol II-TFIIS

[5] Weilbaecher, R. G., Awrey, D. E., Edwards, A. M., and Kane, C. M., in preparation.

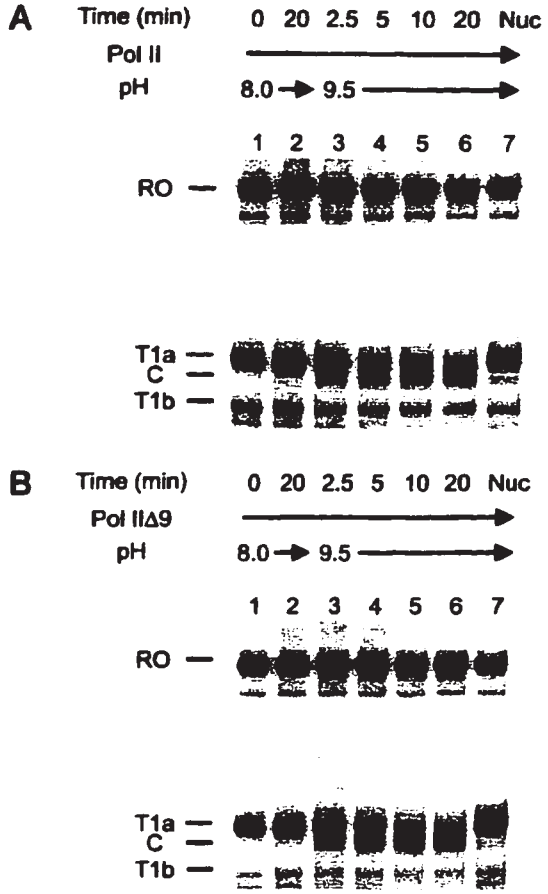


FIG. 5. **Intrinsic cleavage activity of pol II and pol IIΔ9.** Purified TIa ternary complexes containing pol II (*A*) or pol IIΔ9 (*B*) were incubated at pH 9.5 for 2.5, 5, 10, and 20 min (*lanes 3–6*), or pH 8.0 for 0 and 20 min (*lanes 1* and *2*). After 20 min of incubation, an aliquot was adjusted to 800 μM rNTPs and incubated further for 5 min at 30 °C in pH 9.5 buffer (*lane Nuc*). The transcription products were resolved by gel electrophoresis and visualized by autoradiography.

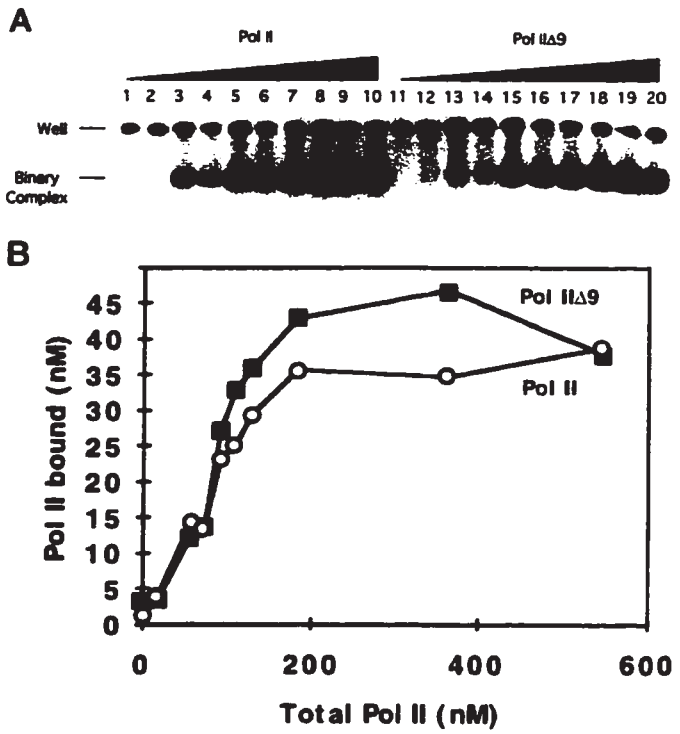


FIG. 6. **Binding of TFIIS to pol II and pol IIΔ9.** *A*, 33 ng of radiolabeled TFIIS was incubated with either pol II (*lanes 1–10*) or pol IIΔ9 (*lanes 11–20*) for 20 min as described under "Materials and Methods." The complexes were resolved by nondenaturing gel electrophoresis and visualized by autoradiography. *Lanes 1–10* contain 0, 50, 150, 200, 250, 300, 350, 500, 1000, and 1500 ng of polymerase, respectively. The mobility of the binary complex of TFIIS and polymerase, as well as the position of the well, is shown on the *left*. The free TFIIS migrated toward the cathode and was not visualized. *B*, quantification of the average of 10 different binding assays. The absolute amount of radiolabeled TFIIS bound to pol II was determined by spotting known amounts on the dried gel prior to phosphorimaging.

complex enter the gel, whereas free TFIIS migrates toward the cathode.

DISCUSSION

Wild-type RNA polymerase II and the enzyme lacking RPB9 have interesting differences in elongation properties, and these differences suggest that this subunit is important for transmitting signals to the elongating ternary complex. The behavior of ternary elongation complexes containing pol IIΔ9 and the wild-type enzyme differ in two ways. First, pol IIΔ9 does not arrest at intrinsic blocks to elongation with the frequency of the wild-type enzyme. Second, when pol IIΔ9 does arrest during elongation, the arrested pol IIΔ9 complex is much less responsive than wild-type complexes to the read-through and cleavage stimulatory factor, TFIIS. These altered properties have also been reported for two different mutants in the *Drosophila* RNA polymerase II largest subunit (39). The association of these properties with a single subunit in yeast RNA polymerase II sheds additional light on the mechanism of arrest site utilization.

The process of arrest is thought to occur in at least two steps. First, the elongating RNA polymerase pauses transiently at blocks to elongation. Blocks can be either DNA sequences or physical blocks such as bound proteins. Pausing defines a state of the elongating enzyme from which there are three outcomes. Many paused complexes, after some period, continue chain elongation. A proportion terminates transcription (*i.e.* dissociate from the template and release the RNA), and the remainder converts to an arrested state, in which the ternary complex is intact but incapable of further transcription in the absence of accessory factors. The amount of the complexes undergoing arrest also can be correlated with the length of time the complexes remain blocked during elongation (40–42).

The decreased propensity of pol IIΔ9 to arrest at intrinsic blocks to elongation might be explained if pol IIΔ9 complexes either (i) had an overall elongation rate faster then the wild-type enzyme or (ii) had reduced ability to recognize signals that lead to an intrinsic block to elongation (31). Our experiments suggest that the latter explanation is the case, namely that pol IIΔ9 complexes do not recognize intrinsic blocks to elongation or other pause sites as efficiently as the wild-type complexes. Paradoxically, although pol IIΔ9 complexes do not appear to respond to DNA sequences that trigger transcription arrest, the pol IIΔ9 complexes that do arrest appear to be more stable as judged by their reduced responsiveness to TFIIS. These results highlight the complexity of the behavior of RNA polymerases at arrest sites.

RNA polymerases likely adopt many different conformations during the elongation process (21, 39). If arrest site recognition may be mediated by only a subset of the RNA polymerase conformations (arrest-competent), then the efficiency of transcription arrest would be governed by the proportion of the elongating enzymes in an arrest-competent conformation. The relative amounts of each conformation might be regulated by accessory factors, substrate levels and composition, solution conditions, and polymerase subunits. The behavior of pol IIΔ9 is consistent with a role for RPB9 in increasing the amount of RNA polymerase in an arrest competent conformation. In the

absence of RPB9, the read-through conformation might be favored, and as a result a smaller proportion of the elongating enzymes arrest. In the presence of RPB9, the arrest competent conformation would be favored, and consequently a greater proportion of wild-type complexes arrest compared with pol IIΔ9. This proposed role for RPB9 in regulating the interaction of pol II with the nucleic acid is consistent with the location of the RNA polymerase I RPB9 homologue (A12.2) (9) in the three-dimensional structure of yeast RNA polymerase I. Subunit A12.2 is located near the region thought to include the sites for DNA interaction (43) and in a region proposed to envelope the DNA template in an elongation complex. If RPB9 occupies an analogous position in RNA polymerase II, then the subunit is situated perfectly to modulate the structure of the elongation complex.

pol II and pol IIΔ9 elongation complexes also differ in the ability to respond to TFIIS, which promotes reactivation of transcription after arrest. Arrested pol IIΔ9 complexes fail to read-through the TIa site in response to levels of TFIIS that completely stimulate read-through by arrested wild-type complexes. We showed that the defect in pol IIΔ9 complexes is not in binding to TFIIS or in intrinsic cleavage activity. Rather, in pol IIΔ9 complexes, TFIIS stimulated read-through occurs much more slowly than with the wild-type enzyme. The reactivation of arrested pol IIΔ9 complexes is therefore affected at a step after TFIIS binding and transcript cleavage. Perhaps RPB9 facilitates the conversion of an arrest-competent conformation to a read-through competent conformation, and this interconversion is a obligatory step in the read-through process. Thus, our observations define at least a three-step process for the release from the arrested state; binding to TFIIS, induction of transcript cleavage, and then a reactivation step(s) that likely includes a conformational change. This multistep model is consistent with the observations of Ciprés-Palacín and Kane, who also showed, using mutants of TFIIS, that transcript cleavage could be uncoupled from read-through (44).

There is a sequential appearance of transcript cleavage products in both mammalian (34) and yeast RNA polymerase II arrest complexes at the T1a site. For the mammalian enzyme, only cleavage to the first site was required for read-through (35). The cleavage to C1 can occur in pol IIΔ9 complexes in response to low levels of TFIIS, although the kinetics are slower than with the wild-type enzyme. However, at these levels of TFIIS, the subsequent transcript cleavage events are not detected in the absence of RPB9, nor is read-through detected even when the first cleavage event has occurred in the absence of RPB9. Our observations raise the possibility that transcript cleavage by purified pol II at the TIa site can be distinguished into two steps. The first cleavage event, to C1, is not sufficient for read-through. Rather, cleavage to the C2 site, and perhaps an accompanying conformational change, is correlated with the read-through process.

A Possible Connection between Start Site Selection and Elongation—Transcription arrest, according to the model described above, is mediated by a particular conformation(s) of the polymerase that interacts specifically with the DNA in arrest sites. Elongating polymerase is proposed to contain two DNA binding domains, one upstream and one downstream of the position of the active site (21, 45). The downstream DNA binding domain likely plays a significant role in recognition of the arrest site because in many instances the sequences that trigger arrest lie downstream of the catalytic site in regions yet to be transcribed (23). The downstream DNA binding activity of RNA polymerase may also play a role in the selection of the transcription start site. In yeast, start site selection is thought to involve specific recognition of DNA sequences ahead of the actual start site (46) because the distance from the TATA box to the start site in yeast can vary anywhere from 40 to 120 bases (47). The most frequently used start site(s) therefore cannot be chosen based on a set distance from the TATA box. Rather, the DNA sequences encompassing the start site are likely selected by pol II, perhaps via a scanning mechanism (48, 49).

In cells lacking RPB9, the selection of the start site appears less stringent than in wild-type cells. In these cells, pol IIΔ9 uses a wider range of start sites, which are located upstream of the major sites used in wild-type cells (15, 17). The pol IIΔ9 enzyme also uses these altered start sites *in vitro* (15). Perhaps this start site alteration is related to the altered elongation properties of the pol IIΔ9 enzyme. RNA polymerase, once recruited to the promoter, may need to recognize and bind specific sequences to position the catalytic site for accurate initiation. pol IIΔ9 may be less capable of effective recognition and binding to the start site because an increased proportion of the enzyme is in the read-through conformation. In other words, the pol IIΔ9 enzyme does not efficiently arrest to allow the correct positioning of the catalytic site at the transcription start site. RPB9 would allow pol II to adopt a conformation compatible with selective binding to the sequences that constitute the start sites of transcription.

A necessary feature of this model is that promoter sequences and arrest sites share common features recognized by RNA polymerase. Although sequence similarities are not readily apparent, promoter release and release from intrinsic arrest sites share many characteristics. First, the release of polymerase from both a promoter and intrinsic arrest sites appears to be unfavorable in that the polymerase utilizes accessory factors to facilitate each process (50–52). Second, the polymerase is catalytically active within both initiation and elongation complexes and is able to perform abortive synthesis of small transcripts while remaining in place. For example, when bound to a promoter, RNA polymerases catalyze the formation of small RNAs that are abortively terminated and ejected from the complex (53–55). Similarly in arrested ternary complexes, the enzyme can also catalyze several cycles of transcript cleavage and re-extension with ejection of the small cleaved RNAs (29, 34, 36, 56). Third, for bacterial RNA polymerase, escape from both promoter sequences and the movement through blocks to elongation is accompanied by a downstream movement of the forward edge of the enzyme, as detected by a significant increase in the size of nuclease footprint ahead of the position of the catalytic center of the enzyme (24, 53, 57). Fourth, the biochemical analysis of yeast RNA polymerase III harboring a point mutation in the C160 subunit has also revealed functional connections between promoter clearance, arrest site recognition, and transcript cleavage (58). The C160 mutant polymerase more efficiently recognized arrest sites, less efficiently cleared the promoter, and showed increased transcript cleavage. Fifth, recent studies with *E. coli* RNA polymerase demonstrate that the bacterial elongation factors, GreA and GreB, also stimulate the conversion from an abortive synthesis complex at a promoter to an elongation competent ternary complex (55).

RNA polymerase lacking subunit RPB9 clearly has altered initiation and elongation properties. The ability to reconstitute wild-type pol II activity from purified pol IIΔ9 and RPB9 will facilitate the comparison of the *in vivo* and *in vitro* behavior of RPB9 mutants. Finally, the stability of the pol IIΔ9 arrested complexes, even in the presence of TFIIS, provides a means to derive structural information about the complex of an elongating RNA polymerase bound to TFIIS.

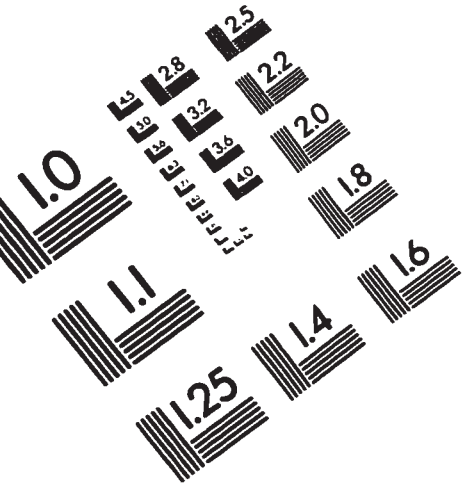
Acknowledgments—We thank Drs. M. Hull and N. Woychik for pro-

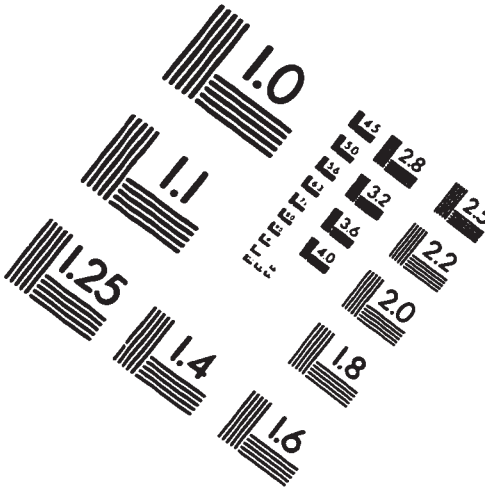
viding the WY9 strain and D. Reines and M. Chamberlin for helpful discussions.

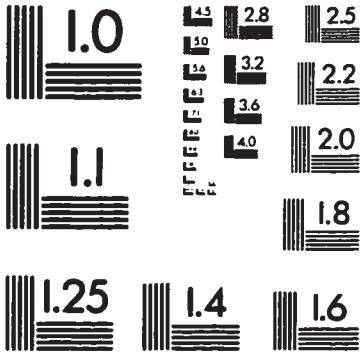
REFERENCES

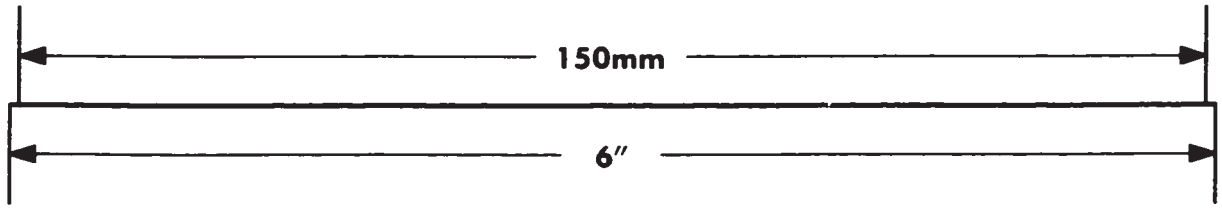
1. Langer, D., Hain, J., Thuriaux, P., and Zillig, W. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92,** 5768–5772
2. Sentenac, A., Riva, M., Thuriaux, P., Buhler, J.-M., Treich, I., Carles, C., Werner, M., Ruet, A., Huet, J., Mann, C., Chiannilkulchai, N., Stettler, S., and Mariotte, S. (1992) *Transcriptional Regulation* (Conaway, R. C., and Conaway, J. W., eds) pp. 27–54, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
3. Built, C. J., White, O., Olsen, G. J., and Venter, J. C. (1996) *Science* **273,** 1058–1073
4. Woychik, N. A., Liao, S.-M., Kolodziej, P. A., and Young, R. A. (1990) *Genes & Dev.* **4,** 313–323
5. Treich, I., Carles, C., Riva, M., and Sentenac, A. (1992) *Gene Exp.* **2,** 31–37
6. Woychik, N., and Young, R. A. (1993) *Gene Exp.* **3,** 77–82
7. Dequard-Chablat, M., Riva, M., Carles, C., and Sentenac, A. (1991) *J. Biol. Chem.* **266,** 15300–15307
8. Sadhale, P. P., and Woychik, N. A. (1994) *Mol. Cell. Biol.* **14,** 6164–6170
9. Nogi, Y., Yano, R., Dodd, J., Carles, C., and Nomura, M. (1993) *Mol. Cell. Biol.* **13,** 114–122
10. Khazak, V., Sadhale, P., Woychik, N., Brent, R., and Golemis, E. (1995) *Mol. Cell. Biol.* **6,** 759–775
11. McKune, K., Moore, P. A., Hull, M. W., and Woychik, N. A. (1995) *Mol. Cell. Biol.* **15,** 6895–6900
12. Shpakovski, G., Acker, J., Wintzerith, M., Lacroix, J., Thuriaux, P., and Vigneron (1995) *Mol. Cell. Biol.* **15,** 4702–4710
13. Woychik, N. A., and Young, R. A. (1989) *Mol. Cell. Biol.* **9,** 2854–2859
14. Woychik, R. A., Lane, W. S., and Young, R. A. (1991) *J. Biol. Chem.* **266,** 19053–19055
15. Hull, M. W., McKune, K., and Woychik, N. A. (1995) *Genes & Dev.* **9,** 481–490
16. Furter-Graves, E. M., Furter, R., and Hall, B. D. (1991) *Mol. Cell. Biol* **11,** 4121–4127
17. Furter-Graves, E., Hall, B., and Furter, R. (1994) *Nucleic Acids Res.* **22,** 4932–4936
18. Kaine, B. P., Mehr, I. J., and Woese, C. R. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91,** 3854–3856
19. Qian, X., Jeon, C., Yoon, H., Agarwal, K., and Weiss, M. (1993) *Nature* **365,** 277–279
20. Agarwal, K., Baek, K., Jeon, C., Miyamoto, K., Ueno, A., and Yoon, H. (1991) *Biochemistry* **30,** 7842–7851
21. Chamberlin, M. J. (1995) *Harvey Lect.* **88,** 1–21
22. Gu, W., and Reines, D. (1995) *J. Biol. Chem.* **270,** 30441–30447
23. Nudler, E., Goldfarb, A., and Kashlev, M. (1994) *Science* **265,** 793–799
24. Wang, D., Meier, T., Chan, C., Feng, G., Lee, D., and Landick, R. (1995) *Cell* **81,** 341–350
25. Edwards, A. M., Darst, S. A., Feaver, W. J., Thompson, N. E., Burgess, R. R., and Kornberg, R. D. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87,** 2122–2126
26. Christie, K. R., Awrey, D. E., Edwards, A. M., and Kane, C. M. (1994) *J. Biol. Chem.* **269,** 936–943
27. Reines, D., Wells, D., Chamberlin, M. J., and Kane, C. M. (1987) *J. Mol. Biol.* **196,** 299–312
28. Reines, D., Chamberlin, M. J., and Kane, C. M. (1989) *J. Biol. Chem.* **264,** 10799–10809
29. Reines, D., Ghanouni, P., Li, Q., and Mote, J., Jr. (1992) *J. Biol. Chem.* **267,** 15516–15522
30. SivaRaman, L., Reines, D., and Kane, C. M. (1990) *J. Biol. Chem.* **265,** 14554–14560
31. Kerppola, T. K., and Kane, C. M. (1990) *Biochemistry* **29,** 269–278
32. Nakanishi, T., Shimoaraiso, M., Kubo, T., and Natori, S. (1995) *J. Biol. Chem.* **270,** 8991–8995
33. Ciprés-Palacín, G., and Kane, C. M. (1995) *Biochemistry* **34,** 15375–15380
34. Reines, D. (1992) *J. Biol. Chem.* **267,** 3795–3800
35. Gu, W., Powell, W., Mote, J., Jr., and Reines, D. (1993) *J. Biol. Chem.* **268,** 25604–25616
36. Guo, H., and Price, D. H. (1993) *J. Biol. Chem.* **268,** 18762–18770
37. Orlova, M., Newlands, J., Das, A., Goldfarb, A., and Borukhov, S. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92,** 4596–4600
38. Wu, J., Awrey, D., Edwards, A. M., Archambeault, J., and Friesen, J. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93,** 11552–11557
39. Chen, Y., Chafin, D., Price, D. H., and Greenleaf, A. L. (1996) *J. Biol. Chem* **271,** 5993–5999
40. Powell, W., and Reines, D. (1996) *J. Biol. Chem.* **271,** 6866–6873
41. Hawley, D. K., Wiest, D. K., Holtz, M. S., and Wang, D. (1993) *Cell Mol Biol Res* **39,** 339–348
42. Kane, C. M., and Edwards, A. M. (1996) *Methods Enzymol.* **274,** 419–436
43. Klinger, C., Huet, J., Song, D., Petersen, G., Riva, M., Bautz, E. K. F., Sentenac, A., Oudet, P., and Schultz, P. (1996) *EMBO J.* **15,** 4643–4653
44. Ciprés-Palacín, G., and Kane, C. M. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91,** 8087–8091
45. Nudler, E., Ekaterina, A., Markovtsov, V., and Goldfarb, A. (1996) *Science* **273,** 211–217
46. Maicas, E., and Friesen, J. (1990) *Nucleic Acids Res.* **18,** 3387–3393
47. Struhl, K. (1989) *Annu. Rev. Biochem.* **58,** 1051–1077
48. Hekmatpanah, D. S., and Young, R. A. (1991) *Mol. Cell. Biol.* **11,** 5781–5791
49. Giardina, C., and Lis, J. (1993) *Science* **261,** 759–762
50. Reines, D. (1994) *Transcription: Mechanism and Regulation* (Conaway, R. C., and Conaway, J. W., eds) pp. 263–278, Raven Press, Ltd., New York
51. Kane, C. (1994) *Transcription: Mechanism and Regulation* (Conaway, R. C., and Conaway, J. W., eds) pp. 279–296, Raven Press, Ltd., New York
52. Goodrich, J., and Tjian, R. (1994) *Cell* **77,** 145–156
53. Krummel, B., and Chamberlin, M. (1989) *Biochemistry* **28,** 7829–7842
54. Luse, D. S., and Jacob, G. A. (1987) *J. Biol. Chem.* **262,** 14990–14997
55. Hsu, L., Vo, N., and Chamberlin, M. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92,** 11588–11592
56. Izban, M. G., and Luse, D. S. (1992) *Genes & Dev.* **6,** 1342–1356
57. Nudler, E., Kashlev, M., Nikiforov, V., and Goldfarb, A. (1995) *Cell* **81,** 351–357
58. Thuillier, V., Brun, I., Sentenac, A., and Werner, M. (1996) *EMBO J.* **15,** 618–629

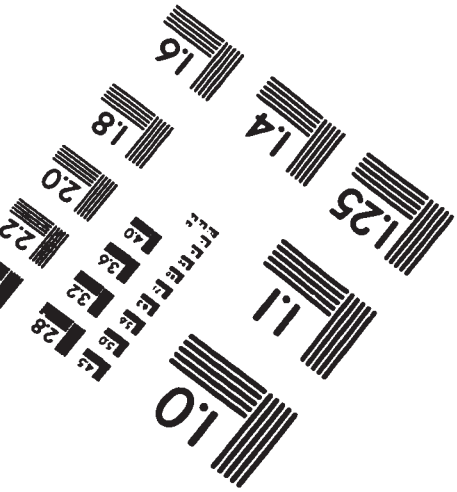
IMAGE EVALUATION
TEST TARGET (QA-3)

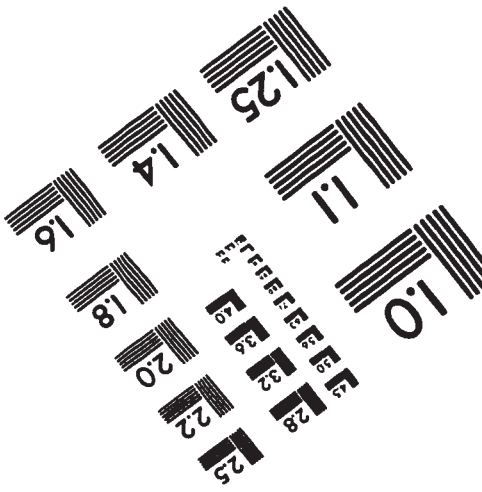












APPLIED IMAGE, Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989